ASTES

Advances in Science, Technology & Engineering Systems Journal

VOLUME 7-ISSUE 4|JUL-AUG 2022

www.astesj.com
ISSN: 2415-6698

Editorial

In this issue, we proudly present a collection of 19 cutting-edge accepted research papers spanning diverse domains. These contributions showcase the latest advancements in fields such as artificial intelligence, transportation, renewable energy, computer vision, geophysics, cybersecurity, and more. Each paper presents novel insights and solutions, contributing to the ever-expanding landscape of scientific knowledge. Let's delve into the details of each paper, with reference numbers included for your convenience.

The first paper delves into the realm of artificial intelligence and medical imaging, specifically focusing on the detection of lung cancer tumors through improved CT images [1]. The study employs advanced image processing techniques and a one-stage detector, achieving impressive results in sensitivity, precision, and F1-score rates.

Expanding on previous research presented at ICECCME2021, this paper discusses the development of a solitary wave track circuit with added functionality, such as insulation deterioration diagnosis [2]. The enhancements contribute to improved reliability, availability, maintainability, and safety in railway operations.

Addressing economic challenges faced by traditional low-income retail stores, this paper introduces a low-cost smart basket designed using ARM system on chip architecture [3]. The innovative basket supports local micro-businesses and promotes social distancing during the COVID-19 pandemic.

Traffic management takes a leap forward with a market-based control approach for real-time intelligent speed adaptation in road networks [4]. The paper presents a solution to optimize traffic flow using dynamic market-based control, addressing the challenges posed by communication delays in connected autonomous vehicles.

The study investigates the application of infrared radiation, microwave antennas, and metamaterials-based sensors for detecting red palm weevils in date palm trees [5]. The experimental results demonstrate the effectiveness of these sensing modalities in predicting the presence of pests.

Contributing to hydrocarbon exploration in the Doba Basin, Chad, this paper integrates seismic and well log data to characterize and analyze petroleum reservoirs [6]. The 3D static reservoir model provides valuable insights into reservoir properties, facilitating predictions of performance and production behaviour.

Addressing the challenges of modern autonomous driving, this paper presents a computer vision radar system with a road line lane detection approach based on the histogram of grayscale images [7]. The method is compared with other computer vision techniques, showcasing its real-time effectiveness.

Introducing a modified simulation tool using Minecraft and ARAIG, this paper provides researchers with an optimized search space and egress path [8]. The integration of the ARAIG haptic suit enhances the user's experience, demonstrating the adaptability of virtual environments in research projects.

Focusing on distribution grid applications, this article presents a comprehensive roadmap for micro-Phasor Measurement Unit (μPMU) hardware and software design [9]. The proposed device ensures high performance, robustness, and accurate measurements in distribution grids.

Recognizing the challenges of interpretability in machine learning models, this paper introduces model selection methods that strike a balance between accuracy and interpretability [10]. The results showcase significant improvements in interpretability with minimal trade-offs in accuracy.

This paper addresses the critical issue of energy demand prediction, presenting a one-year-ahead estimation for Turkey using metaheuristic algorithms [11]. The proposed approaches, especially the M4 model, demonstrate superior estimation capabilities compared to existing models.

Focusing on filter design, this paper introduces compact bandpass filters using innovative structures [12]. The triple bandpass filter, designed with stepped impedance microstrip lines and T-shaped stubs, shows promise for applications in GPS, WLAN, WiMAX, and radar systems.

In the realm of solar photovoltaic modules, this paper introduces a performance adjustment factor to address output power variations caused by factors such as solar irradiation and sun position [13]. The proposed factor ensures optimal performance during different seasons, enhancing the efficiency of solar PV systems.

Examining cyber security reports for Central European countries, this article critically evaluates the state of cyber security, threats, and common attack types [14]. The study emphasizes the impact of COVID-19 on cyber security and proposes measures to enhance defense against phishing attacks.

Addressing the challenges of the false nearest neighbors method, this study proposes a robust method to estimate the minimum embedding dimension without relying on an arbitrary threshold [15]. The results demonstrate the accuracy and reliability of the proposed approach.

Exploring the properties of the Radon transform on convex shapes, this work extends previous findings on its discontinuity [16]. The study reveals that the regularity in the Radon space is determined by the regularity of the shape's points, providing insights into the continuity conditions for line detection methods.

In the realm of wireless networks, this paper introduces a deep learning algorithm for joint source-channel coding, aiming to improve Bit Error Rate (BER) performance [17]. The results demonstrate the superiority of the deep learning autoencoder model over conventional coding systems.

Examining the scalability of optical switches, this paper presents a nested Mach-Zehnder interferometer (MZI) configuration with phase generating couplers [18]. The multi-stage switch exhibits low crosstalk over a broad wavelength range, showcasing its potential for high-speed optical switching.

In addressing the societal issue of alcohol consumption, this paper explores the use of accelerometer data for detecting over-consumption [19]. The comparative analysis of five supervised machine learning methods reveals that "Decision Tree Learning" is the most suitable for accurate sobriety classification using mobile devices.

In conclusion, this special issue offers a diverse array of innovative research papers, spanning fields from medical imaging and transportation to renewable energy and cybersecurity. The contributions showcase advancements in artificial intelligence, geophysics, computer vision, and more, reflecting the ongoing evolution of scientific inquiry. From the improved detection of lung cancer tumors to the development of low-cost smart baskets and the exploration of deep learning algorithms for enhanced wireless communication, each paper adds a valuable piece to the puzzle of contemporary scientific knowledge. The integration of advanced technologies, such as market-based control for traffic management and the utilization of metamaterials in insect detection, demonstrates the interdisciplinary nature of modern research. As we navigate through these insightful papers, it becomes evident that the pursuit of knowledge continues to drive breakthroughs, fostering a dynamic landscape of innovation across various domains.

Reference:

[1] Y.-J. Park, H.-S. Cho, "Lung Cancer Tumor Detection Method Using Improved CT Images on a One-stage Detector," Advances in Science, Technology and Engineering Systems Journal, **7**(4), 1–8, 2022, doi:10.25046/aj070401.

[2] T. Terada, H. Mochizuki, H. Nakamura, "Maintainability Improving Effects such as Insulation Deterioration Diagnosis in Solitary Wave Track Circuit," Advances in Science, Technology and Engineering Systems Journal, **7**(4), 9–14, 2022, doi:10.25046/aj070402.

[3] S. Prongnuch, S. Sitjongsataporn, P. Sang-Aroon, "Low-cost Smart Basket Based on ARM System on Chip Architecture: Design and Implementation," Advances in Science, Technology and Engineering Systems Journal, **7**(4), 15–23, 2022, doi:10.25046/aj070403.

[4] J. Raiyn, "Using Dynamic Market-Based Control for Real-Time Intelligent Speed Adaptation Road Networks," Advances in Science, Technology and Engineering Systems Journal, **7**(4), 24–27, 2022, doi:10.25046/aj070404.

[5] M.M. Bait-Suwailam, N. Al-Nassri, F. Al-Khanbashi, "Assessment of Electromagnetic-Based Sensing Modalities for Red Palm Weevil Detection in Palm Trees," Advances in Science, Technology and Engineering Systems Journal, **7**(4), 28–33, 2022, doi:10.25046/aj070405.

[6] D.A. Diad, D.K. Janvier, A. Boukar, V. Oyoa, "The use of Integrated Geophysical Methods to Assess the Petroleum Reservoir in Doba Basin, Chad," Advances in Science, Technology and Engineering Systems Journal, **7**(4), 34–41, 2022, doi:10.25046/aj070406.

[7] H. Facoiti, A. Boumezzough, S. Safi, "Computer Vision Radar for Autonomous Driving using Histogram Method," Advances in Science, Technology and Engineering Systems Journal, **7**(4), 42–48, 2022, doi:10.25046/aj070407.

[8] C.F. Laffan, R.V. Kozin, J.E. Coleshill, A. Ferworn, M. Stanfield, B. Stanfield, "ARAIG and Minecraft: A Modified Simulation Tool," Advances in Science, Technology and Engineering Systems Journal, **7**(4), 49–58, 2022, doi:10.25046/aj070408.

[9] A.A. Elsayed, M.A. Abdellah, M.A. Mohamed, M.A.E. Nayel, "uPMU Hardware and Software Design Consideration and Implementation for Distribution Grid Applications," Advances in Science, Technology and Engineering Systems Journal, **7**(4), 59–71, 2022, doi:10.25046/aj070409.

[10] Z. Nazir, T. Zarymkanov, J.-G. Park, "A Machine Learning Model Selection Considering Tradeoffs between Accuracy and Interpretability," Advances in Science, Technology and Engineering Systems Journal, **7**(4), 72–78, 2022, doi:10.25046/aj070410.

[11] B. Jamil, L. Serrano-Luján, J.M. Colmenar, "On the Prediction of One-Year Ahead Energy Demand in Turkey using Metaheuristic Algorithms," Advances in Science, Technology and Engineering Systems Journal, **7**(4), 79–91, 2022, doi:10.25046/aj070411.

[12] A. Sowjanya, D. Vakula, "Metamaterial-Inspired Compact Single and Multiband Filters," Advances in Science, Technology and Engineering Systems Journal, **7**(4), 92–97, 2022, doi:10.25046/aj070412.

[13] K. Tsamaase, J. Sakala, K. Motshidisi, E. Rakgati, I. Zibani, E. Matlotse, “Performance Adjustment Factor for Fixed Solar PV Module,” Advances in Science, Technology and Engineering Systems Journal, **7**(4), 98–104, 2022, doi:10.25046/aj070413.

[14] K. Halouzka, L. Burita, A. Coufalikova, P. Kozak, P. Františ, “A Comparison of Cyber Security Reports for 2020 of Central European Countries,” Advances in Science, Technology and Engineering Systems Journal, **7**(4), 105–113, 2022, doi:10.25046/aj070414.

[15] K. Nakane, A. Sugiura, H. Takada, “Estimating a Minimum Embedding Dimension by False Nearest Neighbors Method without an Arbitrary Threshold,” Advances in Science, Technology and Engineering Systems Journal, **7**(4), 114–120, 2022, doi:10.25046/aj070415.

[16] P. Vatiwutipong, “Regularity of Radon Transform on a Convex Shape,” Advances in Science, Technology and Engineering Systems Journal, **7**(4), 121–126, 2022, doi:10.25046/aj070416.

[17] N.O. Chikezie, U.C. Femi, O.O. Ozioma, A.E. Oluwatomisin, A.-U. Chukwuebuka, N.E. Onyekachi, G.C. Kalejaiye, “BER Performance Evaluation Using Deep Learning Algorithm for Joint Source Channel Coding in Wireless Networks,” Advances in Science, Technology and Engineering Systems Journal, **7**(4), 127–139, 2022, doi:10.25046/aj070417.

[18] M. Kawasako, T. Watanabe, T. Nagayama, S. Fukushima, “Scalability of Multi-Stage Nested Mach-Zehnder Interferometer Optical Switch with Phase Generating Couplers,” Advances in Science, Technology and Engineering Systems Journal, **7**(4), 140–146, 2022, doi:10.25046/aj070418.

[19] D. Kumar, A. Thanikkal, P. Krishnamurthy, X. Chen, P. Zhang, “Analysis of Different Supervised Machine Learning Methods for Accelerometer-Based Alcohol Consumption Detection from Physical Activity,” Advances in Science, Technology and Engineering Systems Journal, **7**(4), 147–154, 2022, doi:10.25046/aj070419.

Editor-in-chief

Prof. Passerini Kazmersk

ADVANCES IN SCIENCE, TECHNOLOGY AND ENGINEERING SYSTEMS JOURNAL

Volume 7 Issue 4 July-August 2022

CONTENTS

Lung Cancer Tumor Detection Method Using Improved CT Images on a One-stage Detector 01
Young-Jin Park, Hui-Sup Cho

Maintainability Improving Effects such as Insulation Deterioration Diagnosis in Solitary Wave Track Circuit 09
Takayuki Terada, Hiroshi Mochizuki, Hideo Nakamura

Low-cost Smart Basket Based on ARM System on Chip Architecture: Design and Implementation 15
Sethakarn Prongnuch, Suchada Sitjongsataporn, Patinya Sang-Aroon

Using Dynamic Market-Based Control for Real-Time Intelligent Speed Adaptation Road Networks 24
Jamal Raiyn

Assessment of Electromagnetic-Based Sensing Modalities for Red Palm Weevil Detection in Palm Trees 28
Mohammed M. Bait-Suwailam, Nassr Al-Nassri, Fahd Al-Khanbashi

The use of Integrated Geophysical Methods to Assess the Petroleum Reservoir in Doba Basin, Chad 34
Diad Ahmad Diad, Domra Kana Janvier, Abdelhakim Boukar, Valentin Oyoa

Computer Vision Radar for Autonomous Driving using Histogram Method 42
Hassan Facoiti, Ahmed Boumezzough, Said Safi

ARAIG and Minecraft: A Modified Simulation Tool 49
Cassandra Frances Laffan, Robert Viktor Kozin, James Elliott Coleshill, Alexander Ferworn, Michael Stanfield, Brodie Stanfield

μPMU *Hardware and Software Design Consideration and Implementation for Distribution Grid Applications* 59
Ahmed Abdelaziz Elsayed, Mohamed Ahmed Abdellah, Mansour Ahmed Mohamed, Mohamed Abd Elazim Nayel

A Machine Learning Model Selection Considering Tradeoffs between Accuracy and Interpretability 72
Zhumakhan Nazir, Temirlan Zarymkanov, Jurn-Guy Park

On the Prediction of One-Year Ahead Energy Demand in Turkey using Metaheuristic Algorithms 79
Basharat Jamil, Lucía Serrano-Luján, José Manuel Colmenar

Metamaterial-Inspired Compact Single and Multiband Filters 92
Ampavathina Sowjanya, Damera Vakula

Performance Adjustment Factor for Fixed Solar PV Module 98
Kelebaone Tsamaase, Japhet Sakala, Kagiso Motshidisi, Edward Rakgati, Ishmael Zibani, Edwin Matlotse

A Comparison of Cyber Security Reports for 2020 of Central European Countries 105
Kamil Halouzka, Ladislav Burita, Aneta Coufalikova, Pavel Kozak, Petr Františ

Estimating a Minimum Embedding Dimension by False Nearest Neighbors Method without an Arbitrary Threshold 114
Kohki Nakane, Akihiro Sugiura, Hiroki Takada

Regularity of Radon Transform on a Convex Shape 121
Pat Vatiwutipong

BER Performance Evaluation Using Deep Learning Algorithm for Joint Source Channel Coding in Wireless Networks 127
Nosiri Onyebuchi Chikezie, Umanah Cyril Femi, Okechukwu Olivia Ozioma, Ajayi Emmanuel Oluwatomisin, Akwiwu-Uzoma Chukwuebuka, Njoku Elvis Onyekachi, Gbenga Christopher Kalejaiye

Scalability of Multi-Stage Nested Mach-Zehnder Interferometer Optical Switch with Phase Generating Couplers 140
Masayuki Kawasako, Toshio Watanabe, Tsutomu Nagayama, Seiji Fukushima

Analysis of Different Supervised Machine Learning Methods for Accelerometer-Based Alcohol Consumption Detection from Physical Activity 147
Deeptaanshu Kumar, Ajmal Thanikkal, Prithvi Krishnamurthy, Xinlei Chen, Pei Zhang

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 7, No. 4, 01-08 (2022)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Lung Cancer Tumor Detection Method Using Improved CT Images on a One-stage Detector

Young-Jin Park, Hui-Sup Cho*

Division of Electronics and Information System, DGIST, Daegu, 42988, Republic of Korea

ARTICLE INFO

Article history:
Received: 26 May, 2022
Accepted: 23 June, 2022
Online: 18 July, 2022

Keywords:
Lung Cancer Tumor Detection
DICOM
CAD system

ABSTRACT

Owing to the recent development of AI technology, various studies on computer-aided diagnosis systems for CT image interpretation are being conducted. In particular, studies on the detection of lung cancer which is leading the death rate are being conducted in image processing and artificial intelligence fields. In this study, to improve the anatomical interpretation ability of CT images, the lung, soft tissue, and bone were set as regions of interest and configured in each channel. The purpose of this study is to select a detector with optimal performance by improving the quality of CT images to detect lung cancer tumors. Considering the dataset construction phase, pixel arrays with Hounsfield units applied to the regions of interest (lung, soft tissue, and bone region) were configured as three-channeled, and a histogram processing the technique was applied to create a dataset with an enhanced contrast. Regarding the deep learning phase, the one-stage detector (RetinaNet) performs deep learning on the dataset created in the previous phase, and the detector with the best performance is used in the CAD system. In the evaluation stage, the original dataset without any processing was used as the reference dataset, and a two-stage detector (Faster R-CNN) was used as the reference detector. Because of the performance evaluation of the developed detector, a sensitivity, precision, and F1-score rates of 94.90%, 96.70%, and 95.56%, respectively, were achieved. The experiment reveals that an image with improved anatomical interpretation ability improves the detection performance of deep learning and human vision.

1. Introduction

Lung cancer is the leading cause of cancer-related deaths (18.0% of the total cancer deaths), followed by colorectal (9.4%), liver (8.3%), stomach (7.7%), and female breast (6.9%) cancers [1]. Early diagnosis and treatment may save lives. Although computerized tomography (CT) scan imaging is the best imaging technique in the medical field, it is difficult for doctors to interpret and identify cancer using CT scan images [2]. In addition, because lung cancer detection can increase the detection time and error rate depending on the skill of the doctor, computer-aided diagnosis (CAD) studies to passively assist the detection are on image segmentation, denoising, and 3D image processing using image processing [3] and neural network optimization [4–6].

To improve the anatomical interpretation ability of CT images, this study is designed to improve the cognitive ability of detectors by setting lung, soft tissue, and bone as the regions of interest and utilizing a dataset which is visually easy to distinguish between each region's features in deep learning.

In the dataset construction phase, the dataset was constructed by preprocessing the Digital Imaging and Communications in Medicine (DICOM) files provided by the Lung-PET-CT-Dx dataset [7,8]. The region of interest to which Hounsfield Unit (HU) windowing is applied is composed of a fundamental three-channel dataset (3ch-ORI) to generate an image, and the characteristics of each region can be recognized in one image. Because this process can improve the cognitive ability to visually classify the regions of interest, it is relevant to feature extraction through anatomical analysis in the training process of a neural network, mimicking the human brain and lung cancer detection process. In addition, to enhance the contrast of the 3ch-ORI, a detector with optimal performance was selected by comparing the results of deep learning on a dataset (3ch-CLAHE) to that which the contrast-limited adaptive histogram equalization (CLAHE) was applied. Considering the deep learning phase, deep learning was performed using reference datasets and a reference detector to determine the optimal train customization settings.

The Lung-PET-CT-Dx dataset used in this study provides version 1 (release date: June 1, 2020) to version 5 datasets (release

*Corresponding Author: Hui-Sup Cho, mozart73@dgist.ac.kr

https://dx.doi.org/10.25046/aj070401

date: December 22, 2020); nonetheless, it has been released recently and the related studies [9,10] are insufficient. Therefore, in the evaluation stage, the raw original dataset (1ch-ORI) was used as a reference for comparison. Regarding performance evaluation, the intersection of union (IoU) was calculated to achieve high-level results with a sensitivity, precision, and F1-score rates of 94.90%, 96.70%, and 95.56%, respectively.

2. Related studies

This chapter describes the existing studies that have used methods such as structural separation, noise removal, and three-dimensional (3D) technology for visualization of CT images using the Lung Image Database Consortium and Image Database Resource Initiative dataset (LIDC-IDRI). A novel objective evaluation framework for nodule detection algorithms using the largest publicly available LIDC-IDRI dataset or subset lung nodule analysis 2016 (LUNA16) is a challenge [11]. This set of additional nodules for further development of the IDRI-IDRI dataset that was initiated by the National Cancer Institute (NCI) [12,13] have been released.

In [14], they proposed a novel pulmonary nodule detection CAD system and developed to detect nodule candidates using improved Faster R-CNN. They have archived sensitivity of 94.6%. In [15], the noise present in the CT image was removed by applying the weighted mean histogram equalization (WMHE) method, and the quality of the image was improved using the improved profit clustering technique. Consequently, minimum classification errors of 0.038% and 98.42% accuracies were obtained. The method [16] using the modified gravity search algorithm (MGSA) for the classification and identification of lung cancer in CT images achieved a sensitivity, specificity, and accuracy of 96.2%, 94.2%, and 94.56%, respectively, owing to the application of the optimal deep neural network (ODNN). Furthermore, a threshold-based technique for separating the nodules of lung CT images from other structures (e.g., bronchioles and blood vessels) was proposed [17], and from the evaluation, a sensitivity of 93.75% was achieved. The 3D region segmentation of the nodule in each lung CT image achieved 83.98% [18] because of image reconstruction using the sparse field method. In many other studies, many methods for detection and classification using image processing and deep learning have been proposed, and their performance is quite high. These studies aimed at assisting medical staff with visualization based on image processing. Therefore, various methods need to be continuously studied for CAD systems, where even a 0.01% performance improvement is significant. In this study, the improved CT image is used for deep learning to improve the anatomical analysis ability of each region of interest in the CT image. The dataset aimed at enhancing the quality of the CT image in the pre-processing of the DICOM file without using a complicated image processing method to achieve a high level of result. If the improved CT image is applied to the method proposed in previous studies, a better performance is expected.

3. Materials and Methods

3.1. Dataset construction phase

The preprocessing step of this study includes obtaining a purified CT image through structural analysis of the Lung-PET-CT-Dx dataset, pixel range normalization of the DICOM file, and HU windowing for each region of interest. The Lung-PET-CT-Dx dataset consists of CT and PET-CT DICOM images of lung cancer subjects with XML annotation files that indicate tumor location with bounding boxes. The subjects were grouped according to tissue histopathological diagnosis. Patients with names/IDs containing letter 'A' were diagnosed with Adenocarcinoma, 'B' corresponded to Small Cell Carcinoma, 'E' indicated Large Cell Carcinoma, and 'G' corresponded to Squamous Cell Carcinoma [19].

Object detection performs classification and localization to obtain detection and classification results for each class. However, because this study focuses on the performance of detecting lung cancer, one class was evaluated using only the adenocarcinoma class, without using a dataset with a different number for each class. Therefore, the results of the classification are meaningless, and only the results of localization are used to evaluate the performance. The adenocarcinoma class consists of sub-directories divided for each slice in 265 main directories, and 21 main directories that do not have annotation files or do not match the annotation xml information are excluded from the dataset configuration. In addition, the DICOM files existing in each directory were merged into one directory for easy management and quick data access. The annotation files matching 1:1 with the DICOM file were stored in one common csv file, and after xml parsing, they were stored in the DICOM file. Unlike greyscale images which are in a range of 0 to 255, DICOM files are converted to 12-bit pixel arrays, and DICOM files are composed of HU [20] (a unit that expresses the degree of attenuation of X-rays when penetrating the body). Therefore, as shown in Figure 1, the DICOM file can be viewed more clearly by normalization and HU windowing.

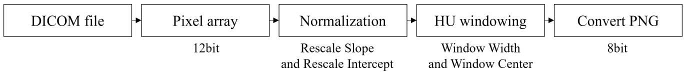


Figure 1: Process steps of the DICOM file

First, the 12-bit (4096 level) pixel array extracted from the DICOM image was normalized according to the unit defined in the HU. Depending on the CT equipment, the pixel range was stored as 0 to 4095 or -2048 to 2047. Using Equation (1), linear transformation was applied to the ‘Rescale slope’ and ‘Rescale intercept’ fields to remap the image pixel:

$$\textit{Out pixel} = \textit{rescale slope} * \textit{input pixel} + \textit{rescale intercept} \quad (1)$$

For example, as shown in the figure, in the 0-to-4095-pixel range, the rescale intercept has a value of − -2048, and the rescale slope has a value of 1; therefore Equation (1) is used to convert it to a value in the range of − -2048 to 2047. In contrast, the rescale intercept of the DICOM file stored in the pixel range of -2048 to 2047 is 0; hence, there is no change even if the above formula is used. Therefore, normalization is applied to the range shown in Figure 2, and all DICOM files are placed within the same pixel range. Considering the reference, dataset, the rescale slope and rescale intercept attributes do not exist in the properties of the DICOM file, they are excluded from the dataset configuration.

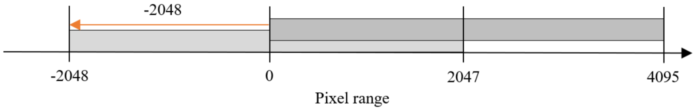


Figure 2: Brightness settings for DICOM image

Subsequently, the normalized pixel array performs the HU windowing process by applying the window width and center properties to each region of interest as shown in Table 1.

Table 1: Window setting using the Hounsfield Unit

Window	Lung	Soft tissue	Bone
Window Center	-700	40	500
Window Width	1400	350	2000

The overall flow of the data construction phase that processes the DICOM file and composes each dataset is shown in Figure 3.

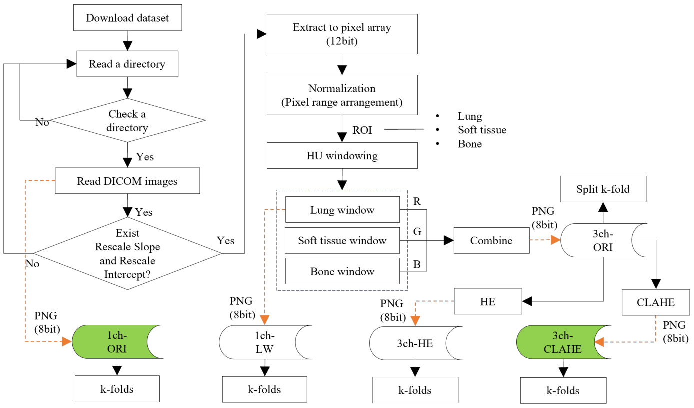


Figure 3: Flow of the dataset construction phase

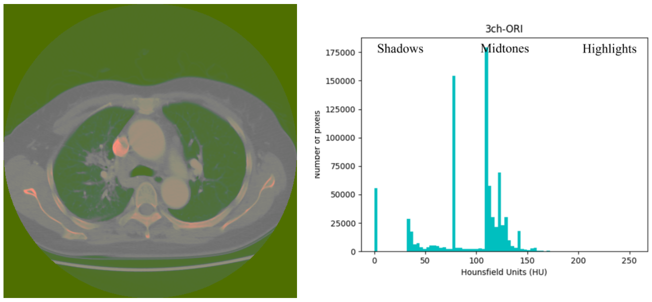
In dataset composition step, a method of composing an image in three-channel and generating images with enhanced contrast by applying CLAHE is described. The CLAHE is an algorithm that uniformly divides an image and distributes pixels of a specific height to each area. After setting the clip limit (the threshold), the height of the histogram was limited.

Table 2 lists the composition and use of the dataset employed. The 1ch-ORI is a dataset consisting of images converted directly into a PNG format from the original DICOM image without any processing and is used as a reference in the experiment. On the contrary, the 3ch-HE, which applies histogram equalization (HE) equally to all pixels, is used as a reference dataset for comparison with the 3ch-CLAHE. In addition, the 3ch-ORI is a fundamental three-channel dataset before the application of equalization.

Table 2: Datasets used for the experiment

Name	Configuration	Usage
1ch-ORI	Raw dataset	Reference dataset
3ch-ORI	3-channel dataset before equalization	Fundamental dataset of 3-channel
3ch-CLAHE	3-channel dataset after CALHE	Proposed dataset
3ch-HE	3-channel dataset after HE	Reference dataset for equalization

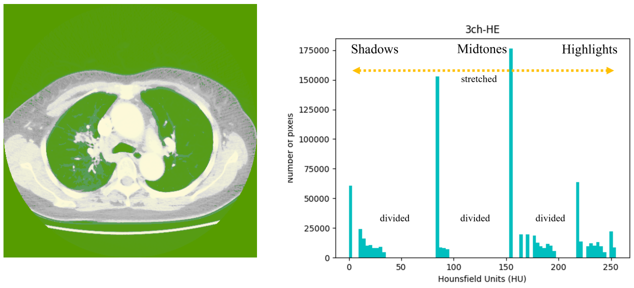
Contrast enhancement using image processing can acquire more detailed information by improving visual recognition ability; thereby, increasing the analysis ability of CT images in the process of human visual and feature recognitions in deep learning. Because methods such as linear combination [21] used for contrast enhancement use multiple images for one-channel, it may affect the contrast range when observing the HU-applied window. Contrast enhancement is a specific characteristic enhancement of image enhancement processing. Histogram equalization is a popular method for image contrast enhancement [22]. Therefore, in this study, the histogram processing technique is applied to the 3ch-ORI, in which the window of interest is set as the data for each channel. However, histogram equalization is not the best method for contrast enhancement because the mean brightness of the output image is significantly different from that of the input image [23]. Brightness is used as an important feature along with shape information when configuring channels, and it is difficult to distinguish between regions because CT images that are not pre-processed may be dark and noise may exist as shown in Figure 4-(A). The histogram can be divided into left, right, and midtones. In Figure 4-(B), where equalization is not applied, the highlighted area is empty; hence, it is difficult to distinguish each area as dark as shown in Figure 4-(A).



(a) 3ch-ORI image b) Histogram

Figure 4: Exemplary image and histogram of 3ch-ORI

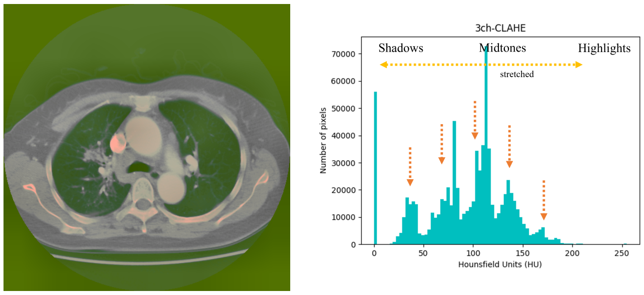
If HE is applied to all pixels at once, equalization is performed indiscriminately. This may cause noise in extremely dark or bright areas or loss of necessary information. Considering Figure 5-(B), where histogram equalization is applied, the number of pixels in the highlighted area is increased (yellow dotted arrows), and it can be seen that Figure 5-(A) affects the increase in pixel intensity and brightness. However, the midtone area decreased, resulting in the spread of shadow and highlight directions. Because it is more difficult to classify each area owing to the addition or loss of information in a specific area, the CLAHE method is used in this study to prevent noise overamplification.



(a) 3ch-HE image (b) Histogram

Figure 5: Example image and histogram of 3ch-HE

In Figure 6, to which CLAHE is applied, it can be observed that the level of pixel areas is spread around the midtone area (yellow dotted arrows), the number of pixels is evenly distributed, and the average value is decreased (orange dotted arrows), enhancing the contrast.



(a) 3ch-CLAHE image (b) Histogram

Figure 6: Exemplary image and histogram of the 3ch-CLAHE

In Figure 4, the contrast is too low to distinguish each region using human eyes; therefore, Figures 5 and 6 with enhanced contrast are used for the experiment. Considering a human point of view, the characteristics of each area in Figure 6 can be distinguished better than in Figure 5; nevertheless, an accurate judgment is made by comparing the deep learning results. Equalization of the histogram using brightness rather than the color of the image was applied after configuration as three three-channels because if three three-channels were configured by applying them to each gray image, different contrasts could be applied to each channel. Therefore, because it is different from the intended image when combined with the color model, three-channel, unintended results of anatomical organs, structures, or artifacts in the human body can have a significant impact on CT image analysis.

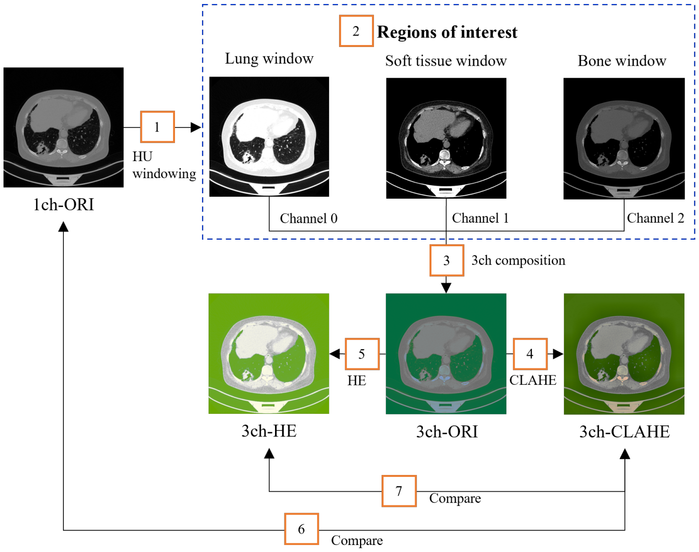


Figure 7: Dataset construction flow

The configuration flow of the dataset is shown in Figure7. The three types of areas of interest (number 2 in the orange box), lung, soft tissue, and bone window, appear clearly after HU windowing. However, it is difficult to anatomically distinguish each area owing to the addition or loss of the specific areas.

After applying each method in the 12-bit pixel array, each dataset was converted to an 8-bit PNG format and stored on a disk. A total of 13,233 images were divided into train- and test-sets in a ratio of 8:2 (10,586:2,647). In addition, because data bias in each dataset can affect the evaluation results of the deep learning model, five-fold cross-validation was performed as shown in Figure 8 to select the optimal fold to be used in the experiment.

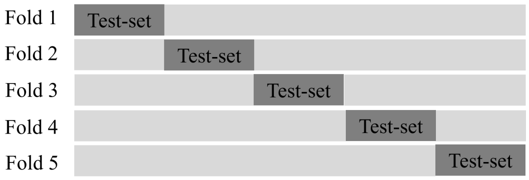


Figure 8: Dataset composition of the cross-validation

3.2. Deep Learning phase

In general, object detection is categorized into one- and two-stage detectors as shown in Table 3. One-stage detectors perform classification and localization concurrently. Therefore, they are fast; however, they are low in accuracy. The two-stage detectors use the Legion Proposal Network (RPN) to select candidate areas where objects are expected to be detected, making them slow; nonetheless, they are high in accuracy. The most recent one-stage detectors exceed the accuracy of two-stage detectors; hence, classification according to accuracy is less meaningful.

In this study, we compared RetinaNet [24] using two-dimensional (2D) image-based anchor-based detectors and Faster R-CNN [25–27] as a reference, modified and utilized Faster R-CNN [28,29] and RetinaNet [30] cloned from the GitHub repository to create a model.

Table 3: Comparison of one- and two-stage detectors

Detector	Anchor based	Detector (Deep learning algorithm)
one-stage	O	YOLO-v1, v2, v3 (2016), SSD (2016) – RetinaNet (2017)
	X	CornerNet (2018) – ExtremeNet (2019) – CenterNet (2019)
two-stage	O	R-CNN (2013) – Fast R-CNN (2015) – Faster R-CNN (2015) – Mask RCNN (2017)

In deep learning, train customization is an important experimental step for selecting the optimal dataset configuration and detector. In this study, ResNet-{50, 101, 152} pretrained with ImageNet [31] was used for transfer learning, and random flip and data shuffle were applied for data augmentation. Moreover, for stable optimization, Adam [32] was used as the optimizer, and the learning rate was set to 1e-5. Considering the reference, Faster R-CNN, which is divided into two stages of RPN and classifier, sets the customization of both elements similarly. Furthermore, the epoch size was set to 500, and the batch size was set to four or eight. Experiments were performed in the environment of Python 3.7, Cuda-10, and a GPU on a 64bit Ubuntu18.04LTS operating system.

After generating a model using the training set and evaluating it using the test set, the overall flow of the deep learning phase (the process of determining the final performance) is shown in Figure 9.

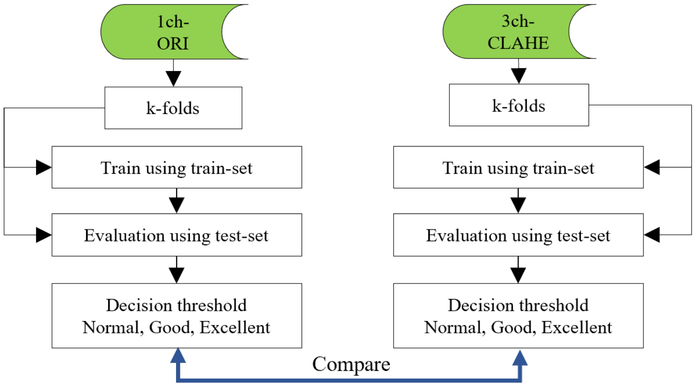


Figure 9: Flows of the deep learning phase

Object detection performs both classification, which classifies objects in the bounding box, and localization, which is a regression process for finding the bounding box. However, because the classification performance of this study is recognized as only one class using a single class, it is not reflected in the evaluation, and the IoU of the bounding box detected by localization and the ground truth (GT) box, which is the annotation information, is calculated. Considering the reference dataset, when multiple bounding boxes are detected in the test-set image, the bounding box with the highest IoU is selected as the IoU of the image. In many cases, natural scene images can be judged by predicting low-level detection results of objects such as people or automobiles with the human eye. However, lung cancer tumors have a non-standard shape; thereby, requiring a higher performance. In this study, to increase the reliability of the detection performance as shown in Table 4, the decision thresholds for each image are set to be narrower than those of the natural scene image for final judgment. However, the narrower the threshold setting range is, the higher the reliability and the lower the statistical evaluation results. Because the achievement result can be relatively decreased, it must be carefully set according to the field of use

Table 4: Setting of the decision threshold

Decision	Natural scene	Proposed
Normal	>=0.50	>=0.60
Good	>=0.70	>=0.75
Excellent	>=0.90	>=0.90

Table 5: Confusion matrix

Name	Threshold	Description
TP	>= 0.6	Lung cancer exists, detected correctly
TN	No use	No lung cancer exists, identified correctly
FP	< 0.6	No lung cancer exists, detected incorrectly
FN	-	Lung cancer exists, missed

The evaluation of models using the test-set images uses the outcomes of four kinds of statistical confusion matrices as shown in Table 5. True Positive (TP) is determined to correctly detect lung cancer tumors when the IoU is 0.6 or higher, and false positive (FP) is determined to be incorrectly detected when the IoU value of the detected bounding box is 0.6 or less. Considering the false negatives (FN), because there is no detected bounding box, the IoU for the GT box cannot be calculated; therefore, a value excluding TP from the total dataset is used. Regarding the reference dataset, negative (TN) is used when lung cancer tumors do not exist and is not used in the field of object detection for a dataset consisting of one class in which the GT box exists in all the test-set images. The total test-set image length and number of GT boxes were the same.

The experimental results were evaluated using statistical performance measurement methods such as sensitivity, precision, and F1-score. Sensitivity represents the predicted positive among all positives and is calculated using Equation 2:

$$Sensitivity\ (or\ Recall) = TP / (TP + FN) \qquad (2)$$

Precision is the proportion of true positives among the predicted positives, calculated using Equation 3:

$$Precision = TP / (TP + FP) \qquad (3)$$

Because the indicators of sensitivity and precision are inversely proportional, it cannot be concluded that a high value of one of them has good performance. Finally, the F1-score is a harmonic mean that considers both sensitivity and precision. The optimal value is one, and the higher the value is, the better the performance, and it is calculated using Equation 4:

$$F1\text{-}score = 2 * (Precision * Recall) / (Precision + Recall) \quad (4)$$

3.3. Results and Discussion

In this chapter, experiments are conducted using each evaluation element, as shown in Figure 10, and the results are discussed.

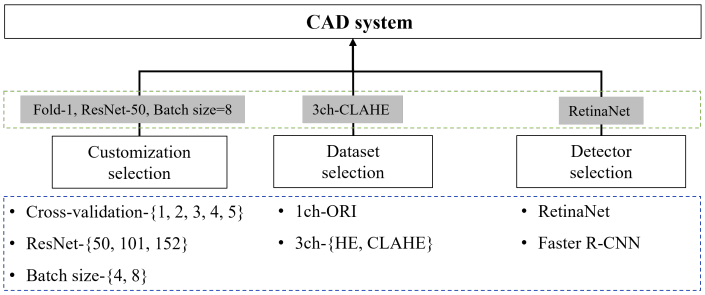


Figure 10: Evaluation factors and flows.

Train customization selection (blue dotted box) through cross-validation experiments, ResNet-depth, and batch size. The data selection and detector selection steps describe the process of selecting a detector with the best performance (green dotted box) through evaluation and using it as a CAD system. First, the experimental results of train customization according to the conditions of cross validation, ResNet-depth, and batch size for the selection of a detector to be used in the CAD system are shown in Table 6. When the Fold-Num was set to one and the batch size was set to eight, the highest result was a sensitivity of 94.9% and precision of 96.7% in ResNet-50, and the F1-score result was the highest in ResNet-101 with 95.8%. Considering the reference dataset, the result of the F1-score in ResNet-50 was 95.6%, being the second highest result.

Table 6: Comparison of ResNet-50 that changed the conditions of customization and ResNet-{101, 152}

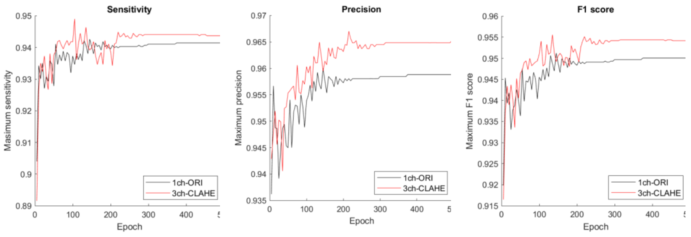
Fold	Depth	Batch Size	Sensitivity (%)			Precision (%)			F1-score (%)		
			Normal	Good	Excellent	Normal	Good	Excellent	Normal	Good	Excellent
1	50	8	94.9	72.3	14.5	96.7	73.9	14.8	95.6	73.1	14.6
2	50	8	93.7	72.9	13.8	96.3	75.6	14.3	94.7	74.2	14.1
3	50	8	93.8	70.3	13.4	96.4	73.1	13.8	94.8	71.5	13.6
4	50	8	93.1	72.0	13.9	95.3	73.7	14.3	94.2	72.9	14.1
5	50	8	93.0	72.4	13.9	96.0	74.8	14.4	94.4	73.6	14.1
1	101	8	94.8	72.5	13.7	96.7	74.0	14.0	95.8	73.2	13.8
1	152	8	94.6	72.8	14.3	96.5	74.6	14.7	95.4	73.7	14.5
1	50	4	94.4	72.5	14.8	96.6	74.7	15.3	95.6	73.6	15.0
1	101	4	94.5	72.6	14.9	96.6	74.5	15.3	95.2	73.5	15.1
1	152	4	94.6	72.8	14.3	96.5	74.6	14.7	95.4	73.7	14.5

Considering the experimental results, Fold-1, ResNet-50, and batch size: eight (which show the best overall performance), were selected the customization setting values of the detector. Since the performance was the best when using the Fold-1 dataset trained using the ResNet-50 neural network, the experiments using the ResNet-101 and ResNet-152 neural networks were compared with the ResNet-50 using only the Fold-1 dataset.

Table 7 shows a comparison between the 3ch-CLAHE and reference datasets using customization setting values in ResNet-50. First, comparing the results with 1ch-ORI (an unprocessed CT image) showed a performance improvement in the sensitivity (+0.74%), precision (+0.70%), and F1-score (+0.44%). Additionally, comparing the result with 3ch-HE, which applied HE to all pixels, showed a performance improvement in the sensitivity (+0.94%), precision (+0.33%), and F1-score (+0.60%). Therefore, it was found that the dataset in which each ROI was composed of three three-channels and CLAHE applied for contrast enhancement had a significant effect on the performance improvement of deep learning.

Table 7: Comparison of proposed dataset with reference datasets

Dataset	Sensitivity (%)	Precision (%)	F1-score (%)
3ch-CLAHE	94.90	96.70	95.56
1ch-ORI	94.26	96.00	95.12
3ch-HE	93.96	96.37	94.96

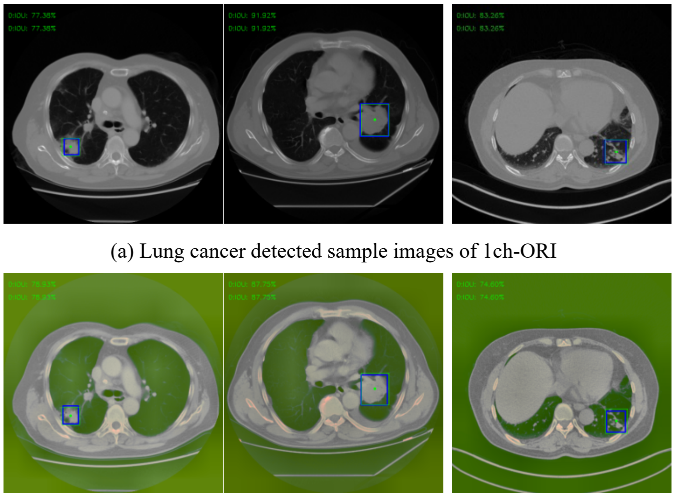


(a) Sensitivity (b) Precision (c) F1-score

Figure 11: Comparison of the proposed dataset (3ch-CLAHE) with the reference dataset (1ch-ORI)

The dataset with the best performance can be obtained from the experimental results in the table; however, visual performance analysis using a graph as shown in Figure 11 can be used as a tool for selecting an appropriate model and determining when to stop learning at the highest performance. The maximum measurement value (y-axis) of each epoch (x-axis) for the comparison datasets 1ch-ORI and 3ch-CLAHE appeared before approximately 200 epochs; nonetheless, stable learning results appeared after approximately 250 epochs. Two hundred and fifty epochs indicate that the learning efficiency is the best.

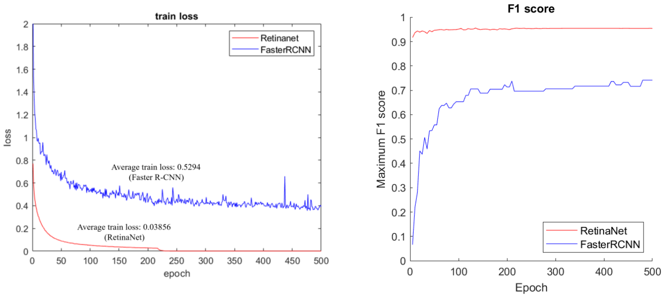
Figure 12 shows only a few cases among the actual detection results using the test set of 1ch-ORI and 3ch-CLAHE. Using the coordinate values of the GT box (blue box) and bounding box (green box) shown in each image, IoU was calculated and used for the evaluation.



(a) Lung cancer detected sample images of 1ch-ORI

(b) Lung cancer detected sample images of 3ch-CLAHE

Figure 12: Examples of detected result images



(a) Train loss (b) F1 score.

Figure 13: Comparison of RetinaNet with the Faster R-CNN

Finally, the performance was compared to Faster R-CNN, which was used as a reference detector to select the final detector. As shown in Figure 13-(A), which compares the average train loss, RetinaNet shows a result of 0.03856. This a big difference from the Faster R-CNN, which shows a result of 0.5294, and the difference in performance and stability during the training process is reflected in the evaluation result. As shown in Figure 13-(B), RetinaNet (red line) shows a high F1-score of 0.9 or higher. However, the train loss of Faster R-CNN (blue line) is unstable, and the F1-score shows a result between 0.7 and 0.8 in Figure 13-(B).

4. Conclusion

In this study, to detect lung cancer quickly and accurately, we attempted to improve the detection performance by improving the image quality. Novel lung cancer detection methods using image processing provide a high level of accuracy by applying noise removal from CT images, segmentation techniques, and methods using 3D images for deep learning. The segmentation technique, which is mainly used to find a small nodule, has the advantage of concentrating on the area. Nonetheless, it also consumes a lot of application time and resources and has the disadvantage that the shape and boundary line may appear irregularly for each image. Because the 3D visualization method of CT images can represent the lungs more realistically, it consumes a lot of resources compared to the other methods although it is used to detect the shape of the lesion with more details.

In this study, we propose a CLAHE-based three-channel dataset construction method that automatically detects lung cancer tumors. Although this method processes CT images in a relatively simple way compared to the novel lung cancer detection methods, high performance has been confirmed through several comparative experiments, and a better performance is expected when applied to the methods of other studies. In addition, the customization of the deep learning process is as important as the configuration of the dataset, and the experimental results reveal that the CT image with improved human visual perception is important for the neural network that mimics the human brain. However, owing to the lack of reference studies, the study was conducted with the goal of improving the performance of the original dataset, and achieved a sensitivity, precision and F1-score rates of 94.90%, 96.70%, and 95.56%. In the results of this study, the one-stage detector showed better performance in train stability and object detection rate than the two-stage detector. Since the images used in this study are medium or small size objects, different results may appear when big size objects are detected using a natural scene dataset, etc.

In addition, although this study cannot be directly compared to studies using the popular public dataset, it serves as a prior study using a dataset that has insufficient comparative studies.

5. Data Availability

The CT scan images used to support the findings of this study have been collected from the Cancer Imaging Archive (TCIA) (link: https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=70224216)

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This study was supported by the Technology Innovation Program (20011875, Development of AI-Based Diagnostic Technology for Medical Imaging Devices) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea).

References

[1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," CA: A Cancer Journal for Clinicians, **71**(3), 209–249, 2021, doi:10.3322/caac.21660.

[2] S. Makaju, P.W.C. Prasad, A. Alsadoon, A.K. Singh, A. Elchouemi, "Lung Cancer Detection using CT Scan Images," Procedia Computer Science, **125**, 107–114, 2018, doi:10.1016/J.PROCS.2017.12.016.

[3] D. Sharma, G. Jindal, "Identifying lung cancer using image processing techniques," in International Conference on Computational Techniques and Artificial Intelligence (ICCTAI), Citeseer: 872–880, 2011.

[4] W. Sun, B. Zheng, W. Qian, "Computer aided lung cancer diagnosis with deep learning algorithms," in Medical imaging 2016: computer-aided diagnosis, SPIE: 241–248, 2016.

[5] A. El-Baz, G.M. Beache, G. Gimel'farb, K. Suzuki, K. Okada, A. Elnakib, A. Soliman, B. Abdollahi, "Computer-aided diagnosis systems for lung cancer: challenges and methodologies," International Journal of Biomedical Imaging, 2013, 2013.

[6] Y. Abe, K. Hanai, M. Nakano, Y. Ohkubo, T. Hasizume, T. Kakizaki, M. Nakamura, N. Niki, K. Eguchi, T. Fujino, N. Moriyama, "A computer-aided diagnosis (CAD) system in lung cancer screening with computed tomography," Anticancer Research, **25**(1 B), 2005.

[7] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, L. Tarbox, F. Prior, "The cancer imaging archive (TCIA): Maintaining and operating a public information repository," Journal of Digital Imaging, **26**(6), 2013, doi:10.1007/s10278-013-9622-7.

[8] P.. W.S.. L.T.. L.J.. H.Y.. & W.D. Li, A Large-Scale CT and PET/CT Dataset for Lung Cancer Diagnosis, doi:https://doi.org/10.7937/TCIA.2020.NNC2-0461.

[9] S. Mazza, D. Patel, I. Viola, "Homomorphic-encrypted volume rendering," IEEE Transactions on Visualization and Computer Graphics, **27**(2), 2021, doi:10.1109/TVCG.2020.3030436.

[10] D. Gu, G. Liu, Z. Xue, "On the performance of lung nodule detection, segmentation and classification," Computerized Medical Imaging and Graphics, 89, 2021, doi:10.1016/j.compmedimag.2021.101886.

[11] A.A.A. Setio, A. Traverso, T. de Bel, M.S.N. Berens, C. van den Bogaard, P. Cerello, H. Chen, Q. Dou, M.E. Fantacci, B. Geurts, R. van der Gugten, P.A. Heng, B. Jansen, M.M.J. de Kaste, V. Kotov, J.Y.H. Lin, J.T.M.C. Manders, A. Sóñora-Mengana, J.C. García-Naranjo, E. Papavasileiou, M. Prokop, M. Saletta, C.M. Schaefer-Prokop, E.T. Scholten, L. Scholten, M.M. Snoeren, E.L. Torres, J. Vandemeulebroucke, N. Walasek, et al., "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge," Medical Image Analysis, **42**, 2017, doi:10.1016/j.media.2017.06.015.

[12] LIDC, The Lung Image Database Consortium image collection, Https://Wiki.Cancerimagingarchive.Net/Display/Public/LIDC-IDRI,.

[13] I.S.G. Armato, H. MacMahon, R.M. Engelmann, R.Y. Roberts, A. Starkey, P. Caligiuri, G. McLennan, L. Bidaut, D.P.Y. Qing, M.F. McNitt-Gray, D.R. Aberle, M.S. Brown, R.C. Pais, P. Batra, C.M. Jude, I. Petkovska, C.R. Meyer, A.P. Reeves, A.M. Biancardi, B. Zhao, C.I. Henschke, D. Yankelevitz, D. Max, A. Farooqi, E.A. Hoffman, E.J.R. Van Beek, A.R. Smith, E.A. Kazerooni, G.W. Gladish, et al., "The Lung Image Database Consortium ({LIDC}) and Image Database Resource Initiative ({IDRI}): A completed reference database of lung nodules on {CT} scans," Medical Physics, **38**(2), 2011.

[14] J. Ding, A. Li, Z. Hu, L. Wang, "Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2017, doi:10.1007/978-3-319-66179-7_64.

[15] P.M. Shakeel, M.A. Burhanuddin, M.I. Desa, "Lung cancer detection from CT image using improved profuse clustering and deep learning instantaneously trained neural networks," Measurement: Journal of the International Measurement Confederation, **145**, 2019, doi:10.1016/j.measurement.2019.05.027.

[16] S.K. Lakshmanaprabu, S.N. Mohanty, K. Shankar, N. Arunkumar, G. Ramirez, "Optimal deep learning model for classification of lung cancer on CT images," Future Generation Computer Systems, **92**, 2019, doi:10.1016/j.future.2018.10.009.

[17] N. Khehrah, M.S. Farid, S. Bilal, M.H. Khan, "Lung nodule detection in CT images using statistical and shape-based features," Journal of Imaging, **6**(2), 2020, doi:10.3390/jimaging6020006.

[18] S. Saien, H.A. Moghaddam, M. Fathian, "A unified methodology based on sparse field level sets and boosting algorithms for false positives reduction in lung nodules detection," International Journal of Computer Assisted Radiology and Surgery, **13**(3), 2018, doi:10.1007/s11548-017-1656-8.

[19] TCIA, A Large-Scale CT and PET/CT Dataset for Lung Cancer Diagnosis (Lung-PET-CT-Dx) , Https://Wiki.Cancerimagingarchive.Net/Pages/Viewpage.Action?PageId=70224216,.

[20] S.J. DenOtter TD, Hounsfield Unit, StatPearls Publishing, 2020, doi:10.32388/aavabi.

[21] S. Ullman, R. Basri, "Recognition by Linear Combinations of Models," IEEE Transactions on Pattern Analysis and Machine Intelligence, **13**(10), 1991, doi:10.1109/34.99234.

[22] Scott E Umbaugh, Computer Vision and Image Processing, Prentice Hall:

New Jersey 1998, 1988.

[23] O. Patel, Y. P. S. Maravi, S. Sharma, "A Comparative Study of Histogram Equalization Based Image Enhancement Techniques for Brightness Preservation and Contrast Enhancement," Signal & Image Processing : An International Journal, **4**(5), 2013, doi:10.5121/sipij.2013.4502.

[24] T.Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollar, "Focal Loss for Dense Object Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, 42(2), 2020, doi:10.1109/TPAMI.2018.2858826.

[25] R. Girshick, "Fast R-CNN," Proceedings of the IEEE International Conference on Computer Vision, 2015 Inter, 2015.

[26] R. Girshick, J. Donahue, T. Darrell, J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2014, doi:10.1109/CVPR.2014.81.

[27] S. Ren, K. He, R. Girshick, J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, **39**(6), 2017, doi:10.1109/TPAMI.2016.2577031.

[28] Kentaro Yoshioka, FRCNN, Https://Github.Com/Kentaroy47/Frcnn-from-Scratch-with-Keras,.

[29] Young-Jin Kim, FRCNN, Https://Github.Com/You359/Keras-FasterRCNN,.

[30] Yann Henon, pytorch-retinanet, Https://Github.Com/Yhenon/Pytorch-Retinanet,.

[31] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2010, doi:10.1109/cvpr.2009.5206848.

[32] D.P. Kingma, J.L. Ba, "Adam: A method for stochastic optimization," in 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, 2015.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 7, No. 4, 09-14 (2022)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Maintainability Improving Effects such as Insulation Deterioration Diagnosis in Solitary Wave Track Circuit

Takayuki Terada[1, *], Hiroshi Mochizuki[2], Hideo Nakamura [2]

[1]*Technology Production Headquarters, Daido Signal Co., Ltd., Tokyo, 105-8650, Japan*

[2]*College of Science and Technology, Nihon University, Chiba, 274-8501, Japan*

ARTICLE INFO

Article history:
Received: 02 June, 2022
Accepted: 01 July, 2022
Online: 18 July, 2022

Keywords:
Track circuit
Solitary wave
Maintainability

ABSTRACT

This paper is an extended version of the journal presented at ICECCME2021. In ICECCME2021, the authors presented that we have developed a solitary wave track circuit (SW-TC), and it is energy-saving compared to existing track circuits. Furthermore, we also explained that it can realize advanced train control at a low cost, equivalent to digital automatic train control. After that, we have conducted research to improve preventive maintenance, which is a problem of existing track circuits, by using SW-TC. In this extended paper, we explain that we can further expand the functions of SW-TC, added new functions such as insulation deterioration diagnosis of the track circuit. With these new functions, the SW-TC can improve reliability, availability, maintainability, and safety. Especially, because of the effect of the insulation deterioration diagnosis function, so railway operators can significantly reduce the time required to identify the cause, when a track circuit failure occurs.

1. Introduction

This paper is an extended version of the journal presented at ICECCME2021 [1]. A track circuit is used as a train detection sensor in railway signals, and consists of a transmitter, receiver, and rails connecting them. In 1872, William Robinson invented a track circuit that detects trains using rails as a circuit. The advent of track circuits enabled automatic signaling systems and contributed to the modernization of train control. Since 1904, various track circuits have been introduced in Japan, and they have contributed to ensuring the safety of railway signals for more than 100 years [2-5]. Since the latter half of the 1990s, research on digital track circuits using microcomputers for track circuits in station premises has been conducted with the aim of saving energy, reducing hardware, and improving maintenance performance, and sending micro-electronics track circuit (SMET) was developed [6-8]. An SMET is capable of time-division processing using digital processing by a microprocessor, can reduce energy and hardware.

On the other hand, in the track circuit between the stations where the distance of the track circuit is long, when processing is performed by one device such as an SMET, the amount of cable is large and the cost is high, so the existing track relay has been used for a long time. Therefore, there were problems such as reduction of cables and energy saving.

To improve these problems, the authors have been developing a new track circuit method, and named it as a solitary wave track circuit (SW-TC). We presented in ICECCME2021 that significant energy savings can be achieved comparing with existing track circuits [1,9]. Furthermore, we explained that it can realize advanced train protection control equivalent to D-ATC [10-18].

After that, we have conducted research to improve preventive maintenance, which is a problem of existing track circuits, by using SW-TC. Currently, railway operators regularly drive inspection vehicles to check the condition of rails in order to prevent track circuit failure. However, there are still many cases of rail breakage and etc., and if adverse conditions overlap, it may lead to dangerous accidents such as derailment. In addition, due to cost issues, inspections using track inspection vehicles are carried out only several times a year, and many small and medium-sized railway operators have not been able to perform sufficient track inspections.

When the track circuit is unexpectedly cut off due to a rail failure, the railway operator walks on the rail and investigates the failure site using a dedicated measuring instrument such as a search

[*]Corresponding Author: Takayuki Terada, terada@daido-signal.co.jp

https://dx.doi.org/10.25046/aj070402

coil. However, it takes a considerable amount of time to investigate the cause and recover.

To address these problems, we have expanded the functions of SW-TC, added new functions such as diagnosing deterioration of the insulation of the track circuit. Hence, it can realize preventive maintenance of the track circuit. In this paper, we explain that SW-TC can improve not only maintainability but also reliability, availability, and safety by expanding the functions, compared to existing track circuit.

2. Materials and Method

In this chapter, we explain the basic principles of SW-TC.

2.1. Solitary wave

A solitary wave is a waveform obtained by cutting out a part of a continuous signal wave, and refers to a signal having only one wavelength of a certain frequency. Figure 1 shows an example of a solitary wave and its interval (no current), respectively. SW-TC can save energy by transmitting a few wave sources (WSs) within a cycle instead of continuous alternating current (AC) signals as in existing track circuits.

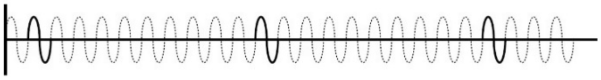


Figure 1: Example of solitary waves

2.2. Example of information allocation

We decided to use the interval between WSs of SW-TC as information. Figure 2 shows an example of information allocation of the SW-TC. Here we prepare 2 WSs and set the WS to a 25Hz sinusoidal waveform for the sake of clarity.

In Figure 2, there are 23 non-current spaces (broken dotted line) in one cycle excluding the 2 WSs (solid line). In the example at the third of Figure 2, there are non-current spaces of 3 waves and 20 waves between the 2 WSs, which is defined as (3:20) or (20: 3). However, (3:20) and (20:3) are considered the same allocation. The small intervals distance between the 2 WSs is defined as signal number (Signal No.), and in the above example, it is defined as Signal No.3. As a result, 11 types of information can be acquired in Figure 2.

If we set the WS to 3 WSs, the information quantity will increase significantly from 11 types to 70 types. In this way, the information quantity can be expanded by the number of WSs, and can be further expanded by increasing the frequency from 25Hz to 50Hz or 100Hz. There is no restriction on the shape of the WS, and it is possible to use a triangle waveform instead of a sinusoidal waveform.

Figure 3 shows an example of the Signal No. transmitted to the SW-TC of each track circuit. First, the state where the train exists on the track circuit and the signal current cannot be received is defined as Signal No.0. Next, the case where a train exists on the front track circuit is defined as Signal No.1, and Signal No. corresponding to the position of the front train is transmitted to the rear track circuit, as shown in Figure 3. Signal No.11 is transmitted to the track circuit that is more than 11 tracks away from the front train.

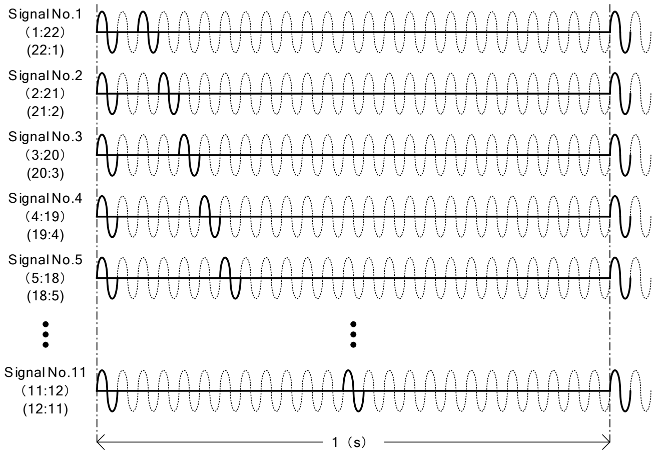


Figure 2: Example of information allocation

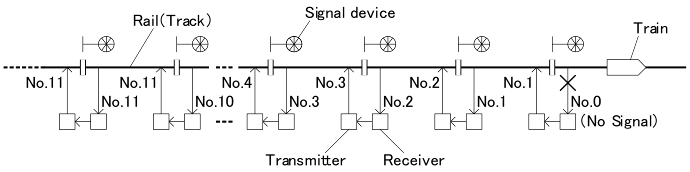


Figure 3: Example of information allocation

2.3. Enhancement by defining solitary wave frame

We have defined the solitary wave frame (SWF) so that SW-TC can be equipped with various functions except for the train position. The structure of the SWF is shown in Figure 4. In this paper, we assume that one SWF is configured at the position of 25WSs.

First, we defined 2 consecutive WSs as a starting element (SE) and placed it at the beginning of the SWF, and set the start position of the frame. If space is secured before and after the SE and one WS is assigned in the remaining positions, 21 types of information can be acquired. If 2 WSs are assigned, the information quantity increases to 190 types, and because of the effect of the SE, the information quantity of SW-TC increases significantly, and the function of SW-TC can be expanded.

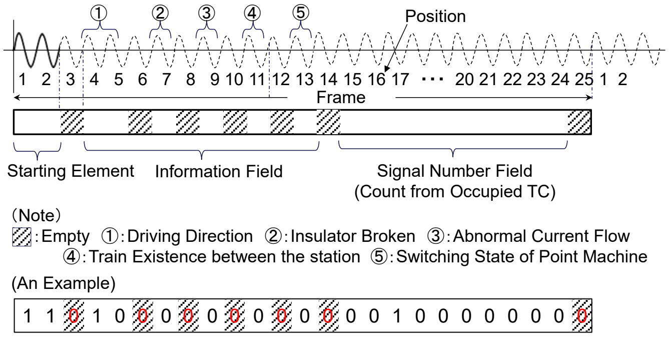


Figure 4: Structure of an SWF

Next, we defined the 4th-13th positions on the SWF as the information field. SW-TC can expand the function by assigning each position of the information field a function, however the details will be explained in the following chapters.

Finally, we defined the remaining 15th-24th positions on the SWF as the signal number field. When the state where the train exists on the track circuit, the 15th-24th positions are assigned 0. When the case where a train exists on the front track circuit, the 15th position is set. When the case where a train exists on the track circuit that is more than 9 tracks away, the 23th position is set. When a train is not exist on the track circuit at the station premises and a route related to that SW-TC is not set , the 24th position is set. Furthermore, in the SWF, there is a restriction that the WS of 2 consecutive waves are assigned only to the SE, and are not assigned in the information field and the signal number field.

3. Current Issue

3.1. Periodic inspection by track inspection car

The track circuit has an important role in detecting trains to ensure safe train operation. Thus, railway operators run inspection cars such as track inspection cars to regularly verify the condition of the tracks [19]. However, despite the inspection, there are still many railroad damages such as rail breakage, and if adverse conditions overlap, it may lead to dangerous accidents such as derailment. Furthermore, since inspection by a dedicated track inspection vehicle is expensive, and there are many small and medium-sized railway operators that cannot perform sufficient track inspections.

3.2. Condition monitoring by track circuit monitor

In the existing track circuit, preventive maintenance has been achieved by introducing a track circuit monitor, a condition monitoring system, etc., and constantly measuring the transmission/reception level of the track circuit. These track monitoring system must work in different weather conditions. Figure 5 shows an example of the screen of the track circuit monitor for the SMET (SMET monitor) of the train detection device for station premises.

The SMET monitor can display the reception level of each track circuit accumulated in the past, the leakage voltage of the adjacent track circuit, etc., which is effective in identifying the cause of a track circuit failure [20]. However, fluctuations in the transmission/reception level vary by track circuit or weather, and the technology for detecting signs of track circuit failure from the tendency of level fluctuations is not completely developed.

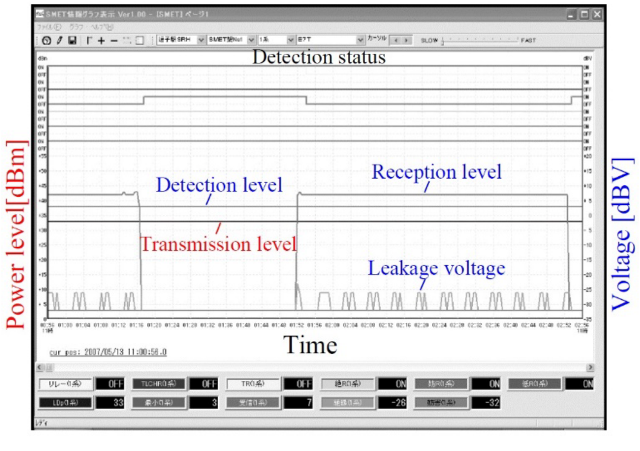


Figure 5: Sample of SMET monitor screen

3.3. Investing the cause with a dedicated measuring instrument

Currently, when an unexpected interruption in the track circuit occurs owing to a failure caused by the rail, in addition to measuring the voltage, a dedicated measuring instrument such as a search coil is used to measure the signal current flowing through the rail by walking on the site. Because there is a high possibility that some type of failure has occurred near the place where the current is changing, the investigation is conducted while grasping the current distribution, but it takes a considerable time to identify the cause and recover the track.

Furthermore, the signal and return currents flow in the track circuit, however, if a large return current flows in the left or right rails, the signal current is affected, and the track circuit becomes unbalanced. When track circuit becomes unbalanced, railway operators investigate the cause using a dedicated measuring instrument such as a return current measuring device. This device measures the return current flowing through the rail with a clamp-type current sensor and measuring device, as shown in Figure 6.

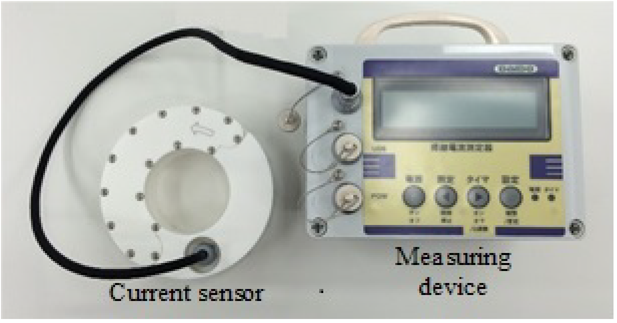


Figure 6: Return current measuring device

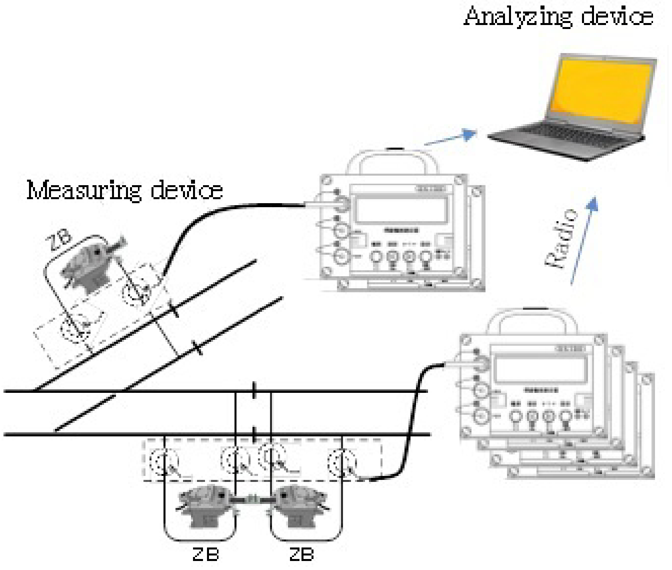


Figure 7: On-site measurement method

Figure 7 shows the on-site measurement method [21]. A current sensor and measuring device are installed on-site to measure the current flowing through the left and right rails. The measurement data are stored in a plurality of measuring devices, and the analyzing devices wirelessly collects the measurement data from the measuring devices in a batch, analyzes the data, and identifies the cause. However, it takes time and effort to investigate, such as installing devices at the site after a track circuit failure occurs, measuring data for a long time, and removing the devices from the site after measurements are completed.

4. Results and Discussion

In this chapter, we present that SW-TC can diagnose the insulation deterioration of the track circuit and detect the deterioration of the reception level. As a result, we explain that railway operators can identify signs of track circuit failure and improve maintainability using SW-TC. Furthermore, we also explain that SW-TC can also improve reliability, availability, and safety, compared to existing track circuits.

4.1. Insulation deterioration diagnosis of the track circuit

Ordinally, track circuits other than non-insulated track circuits have track insulation inserted to separate their boundaries [22]. When the insulation breakdown occurs at the boundary of the track circuit, if the track circuit current is short-circuited because of the train approaching, the current of the rear track circuit may wrap around and the track relay will operate, preventing the detection of the train. As a countermeasure, an orbital current with the phase inverted is passed through the rear track circuit, and when there is insulation breakdown regardless of the state of the train, the track relay is dropped, and the block signal is stopped. However, it is not easy to identify the faulty part in the field because the track circuit failure has various factors such as the insulation breakdown of the track circuit, failure of the receiving equipment, rail breakage, and failure of the track circuit transmission equipment. Figure 8 shows the track circuit boundary and track circuit transmission/reception equipment (TC-TRE). Ordinally, there is a dead section between the bonding points where the impedance bond is connected across the rail insulation, and no current flows. When insulation breakdown occurs at the boundary of the track circuit, leakage current flows in the dead section at the point where the track insulation is broken, and the faulty part can be detected by performing on-site measurements using a dedicated measuring instrument. However, during that time, the train cannot operate and maintenance personnel needs to move to the site.

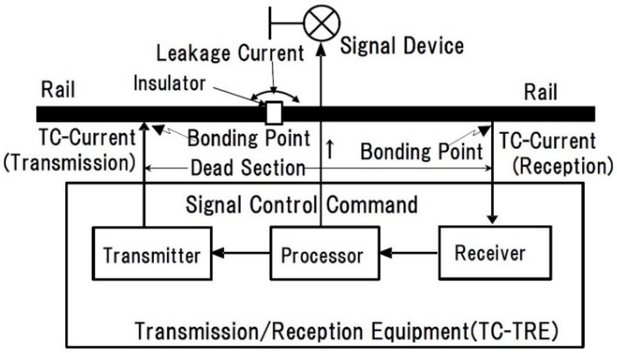


Figure 8: On-site measurement method

As SW-TC assigns SE to the beginning of SWF, if the track circuit insulation is broken, the rear track circuit current wraps around and is superimposed. As a result, the SE generation interval is disturbed in the receiving equipment of the track circuit. At this time, if the current transmission of the transmission equipment to the rear SW-TC is stopped, the wraparound and SE interval disturbance in the SWF are eliminated. The SW-TC can clearly detect the breakdown of the track circuit insulation even during operation by using this mechanism.

Figure 9 shows a deterioration diagnosis method for track insulation of the SW-TC using the principle of the track circuit boundary insulation breakdown detection. Ordinally, the received current input to the TC-TRE has a waveform of [Line No.1], as shown in Figure 9. For the transmission signal to the rear track circuit, an SWF of [Line No.2] is generated, indicating that the Signal No. shifts by one position. In fact, SW-TC transmits SWF at different timings after a certain offset time, as in [Line No.3].

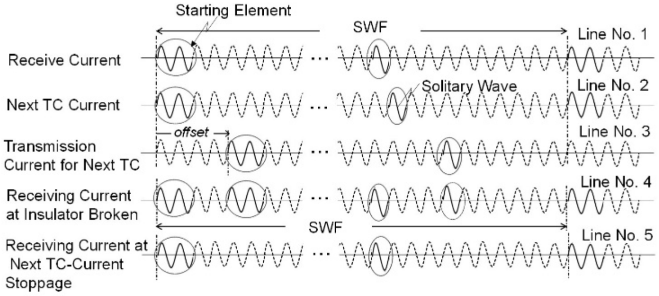


Figure 9: Fault detecting method at a broken insulator

If insulation breakdown occurs at the boundary of the track circuit in this state, SW-TC receives a signal that is a mixture of [Line No.1] and [Line No.3], in which two SEs exist in the SWF. This signal has a waveform similar to that of [Line No.4], and, as explained in section3 of chapter2, TC-TRE can detect insulation breakdown because only SE has two consecutive SWs in the SWF. At this time, if the transmission to the rear track circuit is temporarily stopped, the track circuit current waveform becomes [Line No.5], which is the same as [Line No.1], and the SW-TC can diagnose that the waveform disturbance is due to insulation breakdown. As a result, SW-TC can significantly reduce the time required to identify the insulation breakdown location of the track circuit, and improve maintainability.

Especially in the existing track circuit, railway operators need to measure the received signal level of the track circuit on a regular basis. On the other hand, the SW-TC can automatically performs maintenance measurement for each track circuit and request maintenance, so maintainability of the track circuit can be significantly improved.

4.2. Output of preventive maintenance information

The SW-TC obtains track signals based on the position of WS in SWF; therefore, even when the reception level drops, processing can be continued if the required solitary wave can be detected. However, it is thought that the decrease in the reception level eventually shifts to a state in which the solitary wave itself cannot be detected. Therefore, it is possible to request maintenance at an early stage by outputting the state in which the reception level has dropped as preventive maintenance information. As a result, SW-TC can automatically perform preventive maintenance measurements at the track circuit level and request maintenance while maintaining its function even when the reception level drops, and maintainability of the track circuit can be improved.

4.3. Transmission method

When the SW-TC detects failure sign information of insulation breakdown of the track circuit or deterioration of reception level, as shown in Figure 4, ② (7position: Insulator Broken) or ③ (9position: Abnormal Current Flow) of the in the SWF is set. With this track signal, it is possible to transmit preventive maintenance information via the crew, for example, by blinking the aspect of the block signal.

If the on-board device is mounted on the train, it can detect insulation breakdown or a decrease in the reception level of the track circuit by decoding the failure bit of the information field in the received SWF. With this information, for example, it is possible to turn on the LED of the on-board device and request maintenance.

4.4. Improved reliability

In the existing track circuit, in order to control block signal of multiple aspects, it was necessary to lay a cable and obtain the information of the forward block signal. On the other hand, SW-TC can obtain the necessary train location information with cableless, and can control block signal of multiple aspects. Furthermore, SW-TC is a simple configuration that eliminates the need for transformers, resistors, phase adjusters, etc., which are required depending on the type of existing track circuit.

With the existing D-ATC, it is necessary to install a large-scale ground equipment to generate digital telegrams and detect trains. On the other hand, SW-TC does not require ground equipment and can realize advanced train protection control equivalent to D-ATC, so significant cost reduction can be expected. Therefore, SW-TC can realize the same functions as the existing track circuit with a simple configuration, and thus the reliability is improved.

4.5. Improved availability and safety

In the field environment of an actual railway signal, the reception level of the track circuit signal fluctuates due to the influence of electric rolling stock current flowing in the track circuit and rainfall. Therefore, in the SMET, which is an existing digital track circuit, a highly reliable measure that follows changes in environmental conditions, such as automatic tracking of threshold levels, have been adopted and are effective. SW-TC judges reception based on the digital sampling data obtained by analog-to-digital (A/D) conversion. Therefore, SW-TC records the received waveform when adjusted at the time of installation as a template, and discriminates WS from the correlation between the received waveform and the waveform of the template. Furthermore, SW-TC can separate from noise by confirming the validity of the number of WS in one cycle.

SW-TC can filter noise judging the track circuit information based on the input waveform itself obtained through A/D conversion, not just the level, and confirming the validity such as the shape check of the isolated wave. As a result, it can improve the noise resistance performance, and reduce dangerous accidents due to disturbing waves, as compared with the existing track circuit method. In this way, the SW-TC can reduce the probability that the train will stop due to track circuit failure due to noise, and improves availability. Furthermore, safety can be improved by reducing dangerous accidents.

5. Conclusion

In this paper, we explained the principle of the SW-TC, and introduced application examples to realize preventive maintenance. The proposed SW-TC scheme can not only realize advanced train protection control equivalent to D-ATC, but also significantly improve maintainability compared to the existing track circuits by detecting failure sign information such as insulation breakdown and deterioration of the reception level. As a result, it can significantly reduce the cause identification and recovery time for railway operators when track circuit failures occur.

Track circuits are a proven technology that has ensured the safety for many years, however, in recent years, new train control signal systems that utilize radio have increased, and the systems tend not to use the track circuits. However, it is difficult for small and medium-sized railway operators to introduce a wireless train control system in terms of costs. The proposed SW-TC method is an extension of the existing track circuit, however, cableless and energy saving compared to the existing track circuit, and can realize advanced train protection control at a low cost. Furthermore, it has excellent reliability, availability, maintainability, and safety, its introduction effect is high, and it is a cheaper and more manageable system for small and medium-sized railway operators. SW-TC has the potential to continue to develop its functions, and we expect that the track circuit can be regenerated by using this method.

Conflict of Interest

The authors declare no conflicts of interest.

Acknowledgment

This research has been funded with a grant from Daido Signal co., Ltd. I would like to thank the three anonymous reviewers for their valuable suggestions for improvement.

References

[1] T. Terada, H. Mochizuki, H. Nakamura, “Development of New Track Circuits for Energy Conservation and Signal Control Innovation," 2021 International Conference on Electrical, Computer, Communications and Mechatronics Enginnering (ICECCME2021), 1-5, 2021, doi:10.1109/ICECCME52200.2021.9591127.

[2] E. Itakura, Kidoukairo, Signal Safety Association of signal technology series of Japan, 1971.

[3] S. Egusa, Shingou niokeru system kaihatsu, Railway and Electrical Engineering of Japan, 2016.

[4] T. Kawano, M. Fukuda, Atarashii tougougata kidoukairo no kaihatsu, Japan Railway Engineer's Association of Japan, 2012.

[5] Y. Hirao, “Safety Technologies on Railway Signalling and Functional Safety," Fundamentals Review of Japan, 7(2), 124-132, 2013, doi:10.1587/essfr.7.124.

[6] S. Masutani, H. Utsumi, A. Minami, Scanning siki kidoukairo nitsuite,

Railway and Electrical Engineering of Japan, 2011.

[7] A. Minami, Y. Youda, T. Suga, Ressya kenti souti (SMETgata), Daido signal corporation quarterly publication of Japan, 2003.

[8] T. Mizuno, T. Yamamoto, Shingoukiki no shou energy ka no torikumi, Railway and Electrical Engineering of Japan, 2016.

[9] T. Terada, Y. Matsuwaki, T. Fuse, A. Minami, H. Mochizuki, H. Nakamura, "Regeneration of Track Circuit and a Proposal for a New Signal Control System," The transactions of the Institute of Electrical Engineers of Japan.D, **141**(3), 206-211, 2021, doi:10.1541/ieejias.141.206.

[10] N. Terada, Digital ATC, Railway and Electrical Engineering of Japan, 1998.

[11] T. Igarashi, K. Tashiro, Digital ATC system niokeru RAMS kikaku heno taiou, Railway and Electrical Engineering of Japan, 2004.

[12] T. Takashige, kiki bunsangata ATC (digital ATC), Railway and Electrical Engineering of Japan, 1993.

[13] H. Nakamura, Hoan setsubi no sugata to tenbou, Railway and Electrical Engineerig of Japan, 2012.

[14] M. Matsumoto, S. Kitamura, D. Watanabe, Zairaisen digital ATC niokeru assurance gijutsu, Railway and Electrical Engineering of Japan, 2001.

[15] M. Fukuda, H. Arai, Ressya no anzenunkou wo sasaeru gijutsu, Railway Technical Research institute of Railway Research Review of Japan, 2010.

[16] M. Fukuda, Kishikata Yukusue, Railway Technical Research institute of Railway Research Review of Japan, 2012.

[17] H. Arai, N. Terada, Singou Ressyaseigyo gijutsu no Hensen to Doukou, Railway Technical Research institute of Railway Research Review of Japan, 2015.

[18] H. Nakamura, Recent Trends of ICT Application to Railway Operation and Signaling Systems: The Innovation of Railway Signaling Systems, Information Processing Society of Japan, 2014.

[19] M. Matsumoto, "Trend of Sensing Technology on Railway Operation," The transactions of the Institute of Electrical Engineers of Japan.E, **127**(11), 461-466, 2007, doi:10.1541/ieejsmas.127.461.

[20] T. Senoo, Ressya kenti souti (SMETgata) you monitor souti, Daido signal corporation quarterly publication of Japan, 2007.

[21] T. Noguchi, M. Suzuki, T. Kobayashi, The development of new measuring device of return current, Technical review of JR East of Japan, 2015.

[22] A. Taguchi, "Improvement of Border Characteristics of Jointless Track Circuit," The transactions of the Institute of Electrical Engineers of Japan.D, **118**(2), 243-252, 1998, doi:10.1541/ieejias.118.243.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 7, No. 4, 15-23 (2022)
www.astesj.com

ASTES Journal
ISSN: 2415-6698

Low-cost Smart Basket Based on ARM System on Chip Architecture: Design and Implementation

Sethakarn Prongnuch[*,1], Suchada Sitjongsataporn[2], Patinya Sang-Aroon[3]

[1]*Department of Robotics Engineering, Faculty of Industrial Technology, Suan Sunandha Rajabhat University, Bangkok 10300, Thailand*

[2]*Department of Electronic Engineering, Mahanakorn Institute of Innovation (MII), Faculty of Engineering and Technology, Mahanakorn University of Technology, Bangkok 10530, Thailand*

[3]*Department of Industrial Design and Packaging, Faculty of Industrial Technology, Suan Sunandha Rajabhat University, Bangkok 10300, Thailand*

ARTICLE INFO

Article history:
Received: 30 May, 2022
Accepted: 04 July, 2022
Online: 18 July, 2022

Keywords:
Low-cost smart basket
Smart basket
ARM
System on Chip (SoC)

ABSTRACT

This paper presents the design and implementation of a low-cost basket based on an ARM system on chip architecture using Raspberry Pi single board computer. The inspiration of this research is how to support the traditional low-income retail store in Thailand driving the local micro-business deal with the economic impacts of survival business from the global retailers. The concept of a smart basket system is to use the open-source software in order to save the budget with the free update system. For the product design, the kansei engineering and form follows function theory are applied. The low-cost basket consists of hardware design based on the system on chip architecture and software design using the proposed smart basket algorithm and user interface. Experimental results show that the proposed smart basket implementation can be convenient for lifestyle shopping experience in the local mini mart. This basket will replace the traditional one, which will help consumers maintain the social distancing and will support the local low-income merchant while running the local business during the COVID-19 pandemic.

1 Introduction

Recently, the modern trade from global retailer changes and affects the traditional low-income retail store in Thailand. Survival strategy for business is how to adapt the traditional business following the digital lifestyle. According to Thailand 4.0 model [1], the economic prosperity is the one of objectives to create a value economy by innovation, technology and creativity supported the well-being style.

The internet of things (IoT) and radio frequency identification (RFID) with the low-power microprocessors and systems on chips are a highly important development used for the communication technology that change people's life in the digital lifestyle [2]. The most important challenges of high-performance embedded system and portable devices are how to manage the power consumption of microprocessors. Moreover, IoT architectures [3] have been upgraded every day to reduce data transmission, latency, power consumption, and bandwidth usage for several applications.

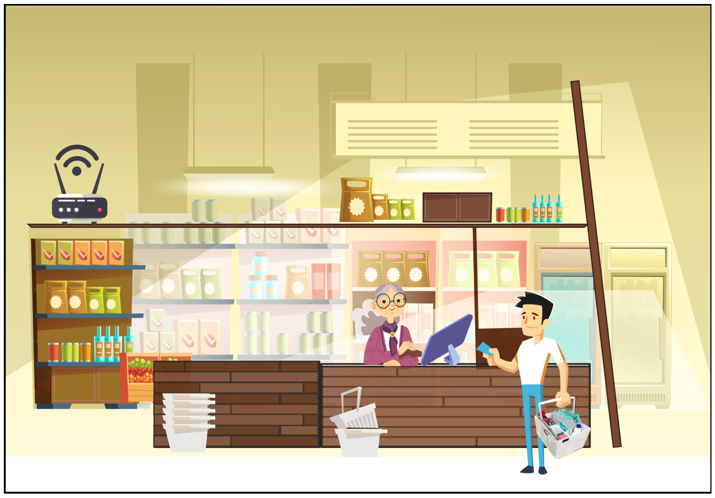


Figure 1: The use of smart basket in traditional retail store

[*]Corresponding Author: Sethakarn Prongnuch, Suan Sunandha Rajabhat University 1 U-Thong nok Rd., Dusit, Bangkok 10300, Thailand, sethakarn.pr@ssru.ac.th

https://dx.doi.org/10.25046/aj070403

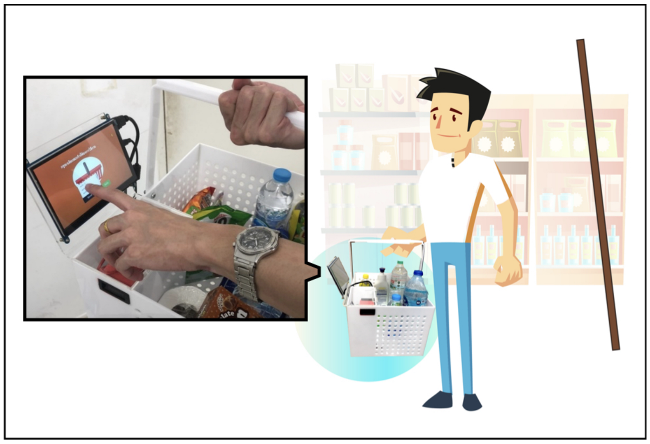


Figure 2: The use of smart basket

According to the embedded system, RISC platform that supports floating-point arithmetic. ARM SoC architecture is embedded in the various industrial control systems, IoT portable devices, and smartphone [4].

In this research, the main objective is to deploy the proposed smart basket for the traditional low-income retail store against the financial crisis and global retailers as shown in Figure 1. The concept of proposed smart basket system is to use the open-source software in order to save cost and easy to update system shown in Figure 2. Design and implementation of a low-cost smart basket should be convenient to improve the lifestyle shopping experience against the COVID-19 at the local supermarkets.

2 Background and Related Work

There are various smart shopping carts and baskets with the RFID, IoT, and embedded system. Nowadays, the effects of coronavirus disease (COVID-19) pandemic living for people around the world has made the changing lifestyle especially shopping at the convenience store and supermarket. In [5], the demonstration experiment of the RFID automatic checkout solution by Panasonic has been proposed. Customers can automatically checkout while walking through the checkout lane with the RFID on basket containing products. In [6], an application on a mobile application to search the shopping shelf in the supermarket has been presented to manage the shopping list.

The utilization of cost-effective smart IoT-based shopping cart has been used in the department store [7]. The development of a smart shopping basket using a barcode reader on a mobile device has been presented in [8]. A smart shopping basket provides a hand-free and hassle-free shopping experience in the supermarket, as detailed in [9]. In [10], an automatic billing on android application for smart shopping has been presented. An automatic WooCommerce application generator framework has been introduced in [11].

There are many researches based on an ARM system such as the spherical magnetic robot using multi-single board ARM computer for controlling and communication [12] and a Mini-UAV for indoor surveillance project in [13]. The advantages of ARM system are affordable to create, low-power consumption, support multiprocessing, and simple circuits.

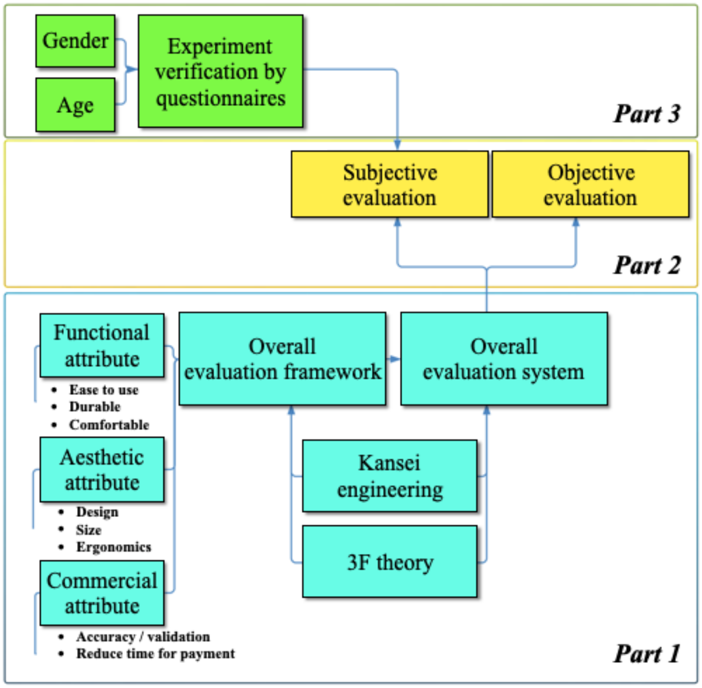


Figure 3: Research framework

For this research, we aim to minimize the production costs of a proposed basket by using the local products in Thailand. The contributions of this paper are summarized as follows: 1) to propose the design and implementation of the low-cost hanging basket based on the kansei engineering and 3F theory for the product design, and 2) to deploy a low-cost basket including with a Raspberry Pi 3 Model B+ SBC, a 1-dimension barcode reader, and a 7-inch TFT LCD display monitor touch screen connected inside the basket by using the online purchase application on the Raspbian operating system for payment.

The rest of this paper is organized as follows. Section 3 presents the proposed low-cost smart basket based on an ARM SoC architecture. Section 4 presents the smart basket algorithm. Section 5 presents the user interface. Section 6 details the experimental results and Section 7 concludes this research.

3 Proposed Low-cost Smart Basket Based on ARM System on Chip Architecture

Proposed low-cost smart basket is used the ARM SoC as a general-purpose processor in a Raspberry Pi SBC architecture to improve the convenience life shopping experience. The design and implementation of a low-cost smart basket are made from plastic, lightweight, durable and suitable for both dry and wet surfaces.

3.1 Proposed Framework

The research framework is divided into three parts as shown in Figure 3. Part 1 is the main contribution of the design process. Kansei engineering and 3F theory are used to construct an overall evaluation system using a hierarchical model including the attribute and evaluation levels.

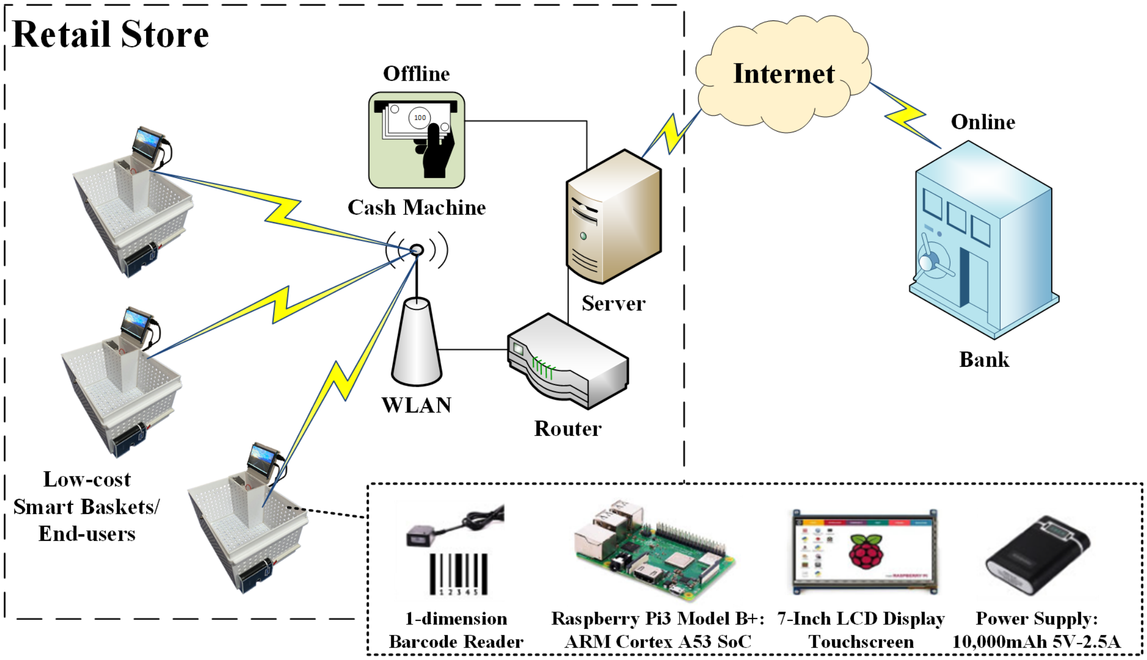


Figure 4: The system architecture of proposed smart basket

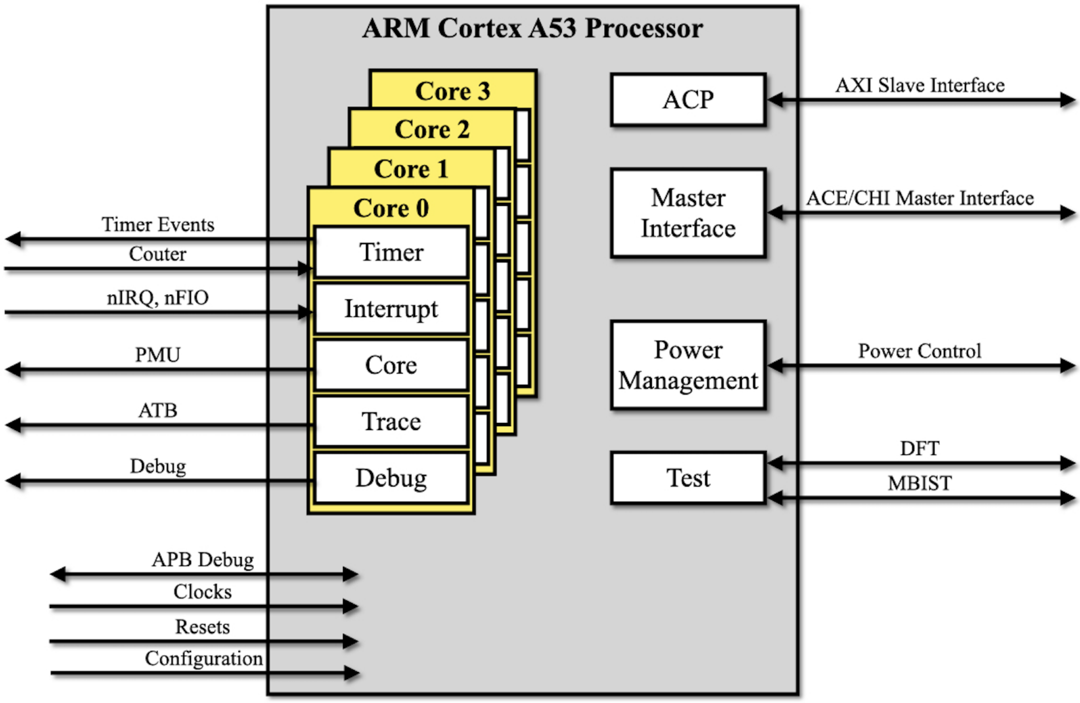


Figure 5: ARM Cortex-A53 processor configuration [15]

Attribute level consists of three levels as the functional attributes: ease to use, durable, comfortable; the aesthetic attribute: design, size, ergonomics; and the commercial attribute: accuracy /validation, reduce time for payment. In part 2, there are two types of evaluation as an objective evaluation and a subjective evaluation. In part 3, the questionnaires are used by the experimental verification for the subjective product evaluation.

3.2 System Architecture

The proposed system architecture of smart basket design as shown in Figure 4 is based on the embedded system and network system. This system architecture can be separated into the retail store and the online banking as follows.

The retail store block divides two parts including as:

1. Network system part includes a WLAN, a router, a server, and the cash machine.

2. Low-cost smart basket part consists of the Raspberry Pi 3 Model B+ SBC with 1.4GHz ARMv8 SoC, 1GB DDR2 low-power memory, a 64GB micro-SD storage, a full size of high definition multimedia interface (HDMI), and 2.4-5GHz IEEE802.11.b/g/n/ac wireless LAN. The DC power supply 10,000mAh 5V-2.5A, a 1-dimension barcode reads, and the 7-inch TFT LCD display monitor 800×480 of resolution with the capacitive touch screen control are installed. The Raspbian is used as the operating system. The component of Raspbian includes a Linux kernel 5.10.17 based on the Debian Linux, which is installed on a micro-SD storage in the Raspberry Pi.

The online banking block with the online purchase application for payment develops by the C/C++, XAMPP as a PHP development environment, HTML, and JavaScript are described in the Section 4. Figure 5 shows an ARM Cortex-A53 processor configuration with four cores and an AXI Coherency Extension (ACE) or the Coherent Hub Interface (CHI) Master Interface.

The 64-bit address version of ARM [14] is used. The Cortex-A53 processor is a mid-range and low-power processor installed the ARMv8 architecture. In this research, the 4-core Cortex-A53 processor is used with an L1 memory system and a single shared L2 cache [15]. There are several low-end IoT devices using ARM as a main processing unit [16] such as OpenMote, LSN50, Memsic Lotus, nRF51 DK, and Arduino MKR1000.

3.3 Design Process and Implementation

Kansei engineering (KE) has been used for the product design following the customer's feeling and needs in the product function [17]. Following the KE concept, we set the adjectives of these influencing factors for evaluation as "weak to strong" on the 5-point scale. Following [18], the product design can be generally divided into three fundamental components as the appearance, functionality, and quality. The form follows function (3F) theory has been associated with the 20th century architecture, engineering, and industrial design related on the function or purpose.

Table 1: Model of proposed smart basket design using the kansei engineering and 3F theory.

Basket management	Level of decision	Basket decision	Design process	Design tools
Smart basket definition	Strategic	Problems to solve with shape and usage	UX/UI	1) Concept of smart basket 2) KE & 3F theory
		Brainstrom	Descriptive statistics	1) Questionnaires 2) User Experience
Basket image	Operational	Formalization	Creation/ usage	Form specification in 3D
Smart basket on embedded system	Operational	Positioning/ Installation	System architecture	C/C++, XAMPP, HTML, PHP
		Inventory & sales report	Stock payment	

Model of the proposed smart basket design using the kansei engineering is shown in Table 1. There are three steps of smart basket management as the smart basket definition, basket image, and smart basket on the embedded system. Level of decision can be divided into two parts as the strategic and operational level.

The basket design process using the kansei engineering and 3F theory are five steps including:

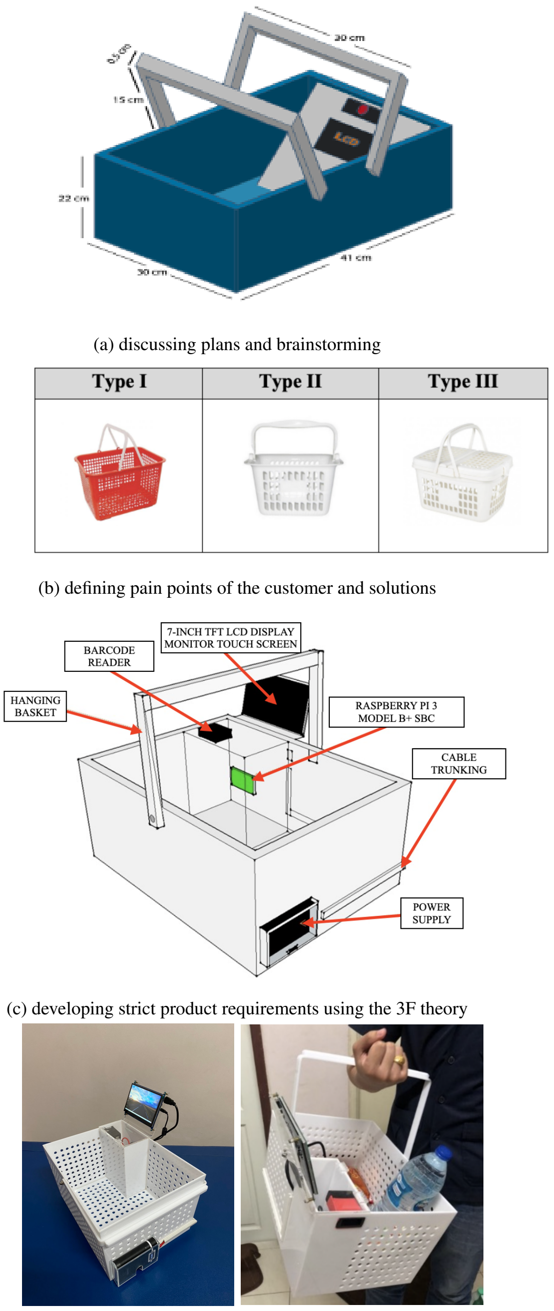
1. Discussing plans and brainstorming for launching a basket: In this step, researchers are surveying about a basket used in the supermarket. Then, they are discussing and selecting a set of baskets. Finally, they make a decision about size of 30×41×22cm shown in Figure 6(a).

2. Defining the pain points of consumers and solutions: In this step, researchers are looking for three different sizes and weights of a basket sold in the hardware stores. Type I, II and III are made from hard plastic material with 30×41×12cm/0.6kg, 17×27×21cm/0.2kg, and 44×33×25cm/0.7kg as shown in Figure 6(b).

3. Developing the strict requirements of basket using 3F theory: In this step, researchers are collecting the information from questionnaires. After that, design and implementation of proposed basket are deployed by 3F theory with all components as shown in Figure 6(c).

4. Production implementation: Prototype white plastic basket is implemented and developed with a size of 30×38×22cm and weight of 2.9kg as shown in Figure 6(d).

5. Testing and modifying: This step is concerning with the user experience to modify and improve for real usage in the future as shown in Figure 6(e).



(a) discussing plans and brainstorming

(b) defining pain points of the customer and solutions

(c) developing strict product requirements using the 3F theory

(d) product implementation (e) testing and modifying

Figure 6: Proposed basket design process using kansei engineering and 3F theory

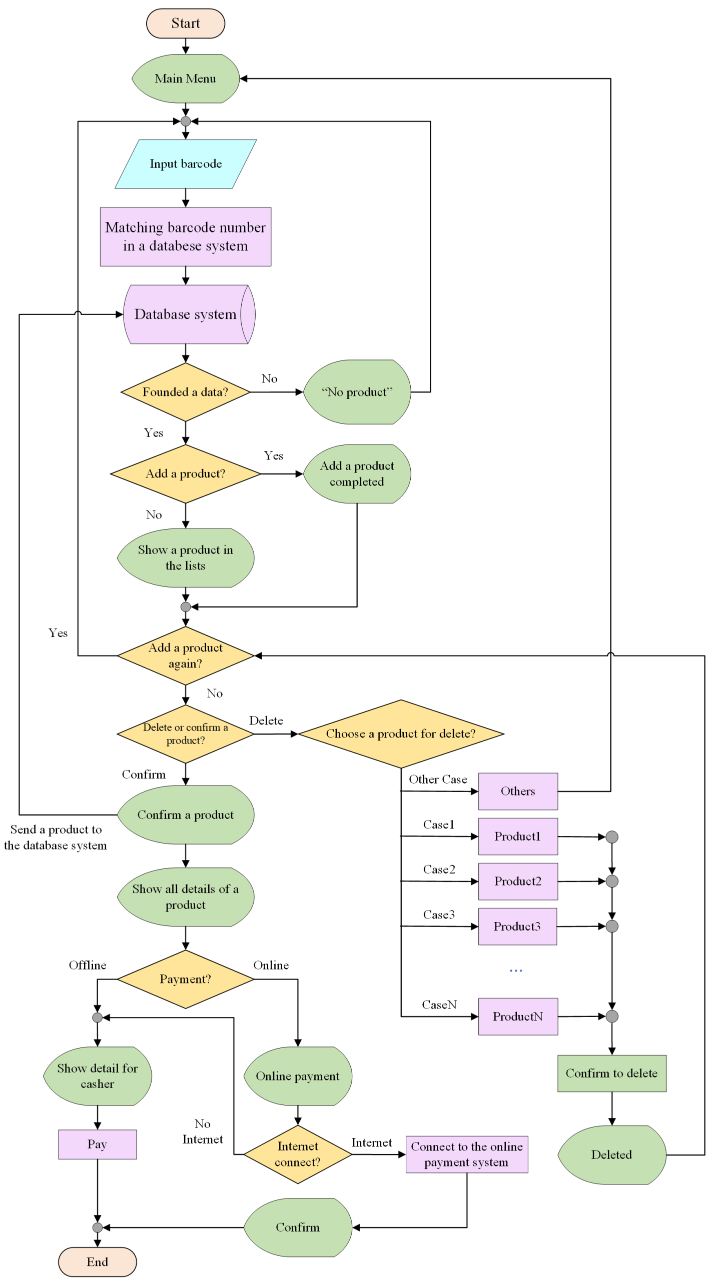


Figure 7: Flowchart of the proposed smart basket

3.4 User Interface and User Experience

User interface (UI) and user experience (UX) [19] are the stage of product interactive development by users. UI is how users can interact with application and tools. UI provides the meaning of input as allowing the users to control the system and output as enabling the system to inform users. UX is an experience and a person's perception related on a system and service that can provide the satisfaction and comfort.

4 Smart Basket Algorithm

In this section, the smart basket algorithm is introduced to manage the stock, database and payment. Flowchart of proposed algorithm is shown in Figure 7.

Basic concept of database system is used as the simple stock management for Thai merchant understanding how to manage and monitor the real-time stock item and sales reports in the small retail store. Smart basket on hand will show the current stock item and reflect the financial plan and forecast, which can prevent the out of stocks happening.

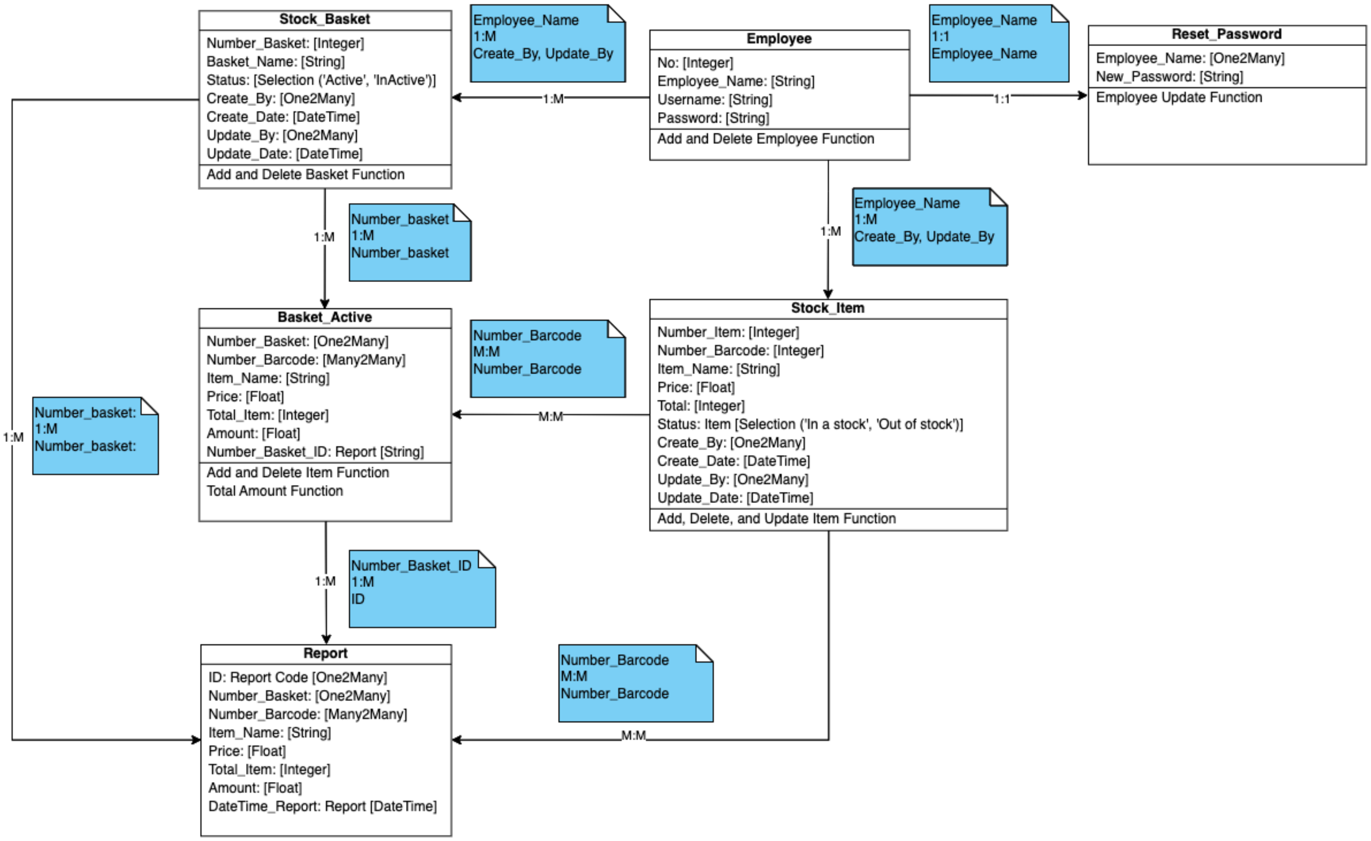
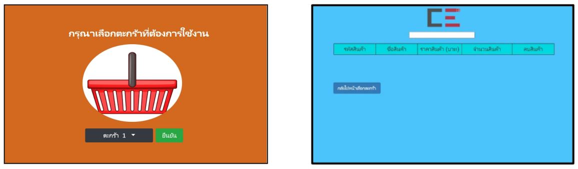


Figure 8: Overall system of the distributed information

The smart basket algorithm can be divided three parts as: add a product, delete a product, and payment.

1. Add a product part: This customer part chooses the product by scanning a barcode. When the Raspberry Pi SBC receives the number of barcode and display the details on a monitor.
2. Delete a product part: At the part of deleting the product, when customers scan the barcode more than one from a barcode reader and require to delete. So, the customers will mark the empty box in front of the product and then confirm to delete.
3. Payment part: When the customers finish the shopping, there are two options of payment as the online and offline. The summary of products in the basket will send to cashier for payment.

Figure 8 shows the unified modeling language (UML) diagram of overall system. The components of UML diagram consist of the stock basket, basket active, report, employee, stock item, and reset password.

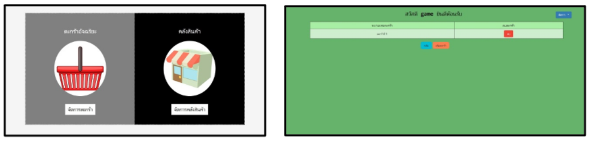


(a) smart basket selection page (b) product details page

Figure 9: User interface for a smart basket design page 1-2

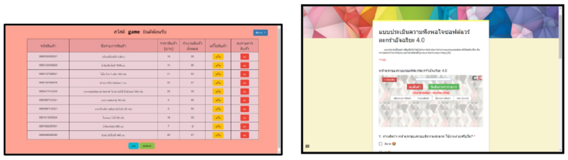
(a) new customer registration page (b) customer login page

Figure 10: User interface for a smart basket design page 3-4



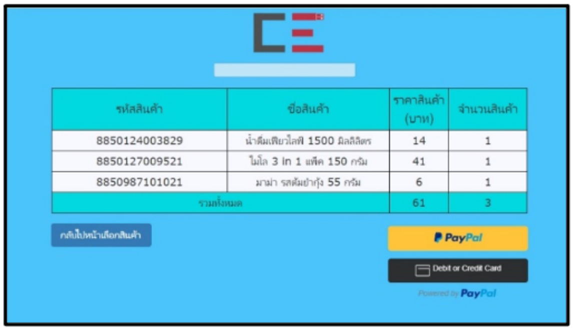
(a) products/stock management (b) product management page

Figure 11: User interface for a smart basket design page 5-6



(a) stock management page (b) customer satisfaction questionnaires page

Figure 12: User interface for a smart basket design page 7-8



รหัสสินค้า	ชื่อสินค้า	ราคาสินค้า (บาท)	จำนวนสินค้า
8850124003829	น้ำดื่มเพียวไลฟ์ 1500 มิลลิลิตร	14	1
8850127009521	ไมโล 3 in 1 แพ็ค 150 กรัม	41	1
8850987101021	มาม่า รสต้มยำกุ้ง 55 กรัม	6	1
รวมทั้งหมด		61	3

Figure 13: User interface for a payment page

5 User Interface of Smart Basket

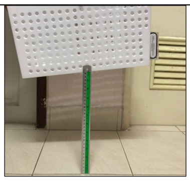
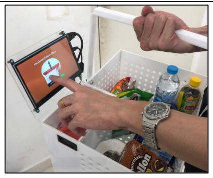
The user interface for a smart basket design for Thai people customers composes of nine pages:1) a smart basket selection page is shown in Figure 9(a), 2) the product details page is shown in 9(b), 3) a new customer registration page is shown in Figure 10(a), 4) the customer login page is shown in Figure 10(b), 5) the products or stock management selection page ss shown in Figure 11(a), 6) the product management page is shown in Figure 11(b), 7) stock management page for tracking and monitoring stock ss shown in Figure 12(a), 8) the customer satisfaction questionnaires page is shown in Figure 12(b), and 9) the payment page is designed for customers to purchase items easily and securely in Figure 13.

6 Experimental Results

Experiments of the proposed low-cost smart basket are composed of the hardware and software tests.

There are three types of hardware test repeated 100 times as weight, drop impact test and battery life. Hardware experimental results are shown in Table 2. It is concluded that the maximum weight capacity of proposed basket hold is around 12 kg. Drop impact tests simulate the drop of proposed basket to guarantee the safety of basket during shopping at the maximum height of 30cm.

Table 2: Hardware experimental results of the proposed smart basket

Weight	Drop impact testing	Battery life
Maximum 12 kg.	Maximum 30cm	Maximum 6 hours
		

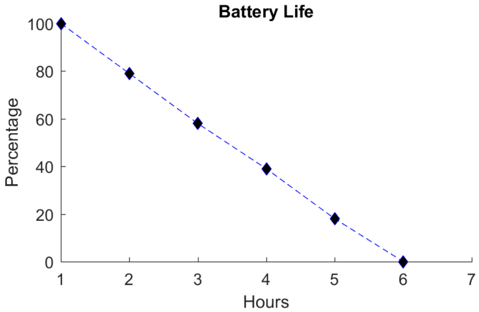
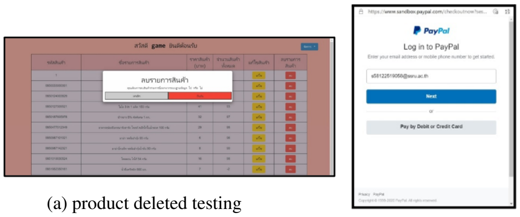


Figure 14: Battery life



(a) product deleted testing

https://www.sandbox.paypal.com/checkoutnow?ses...

Log in to PayPal

Enter your email address or mobile phone number to get started.

s58122519068@ssru.ac.th

PayPal
Next
or
Pay by Debit or Credit Card
Privacy PayPal
Copyright © 1999-2020 PayPal. All rights reserved.

(b) online payment testing

Figure 15: Example of software experiments

For the battery life, a cycle of 10,000mAh battery last is about six hours. Figure 14 shows that the trend of percentage of battery life declines gradually within six hours.

Software test is repeated 100 times concerning with added/deleted products, registration/login and payment. Example of deteted product and online payment tests are demonstrated in Figure 15. The real usage of proposed basket is tested at the local supermarket shown in Figure 16.

Additionally, the results of 100 Thai customers satisfaction questionnaires of the low-cost smart basket that consist of six parts: size, durability, design, application response, appropriate for the supermarket, and comfort are shown in Figure 17(a) to Figure 17(f), respectively. There are five levels of satisfaction including the very satisfied, satisfied, well-done, dissatisfied, and very dissatisfied.

Figure 17 shows the satisfaction from online survey of 100 people with 6 key questions of the use of proposed basket. It can be seen that over 43% of these surveys feel very satisfied for application response to user while shopping in the supermarket. Another 42% gets satisfied for design. In conclusion, it is evident that most customers feel satisfied for ease of application used, while they were shopping.

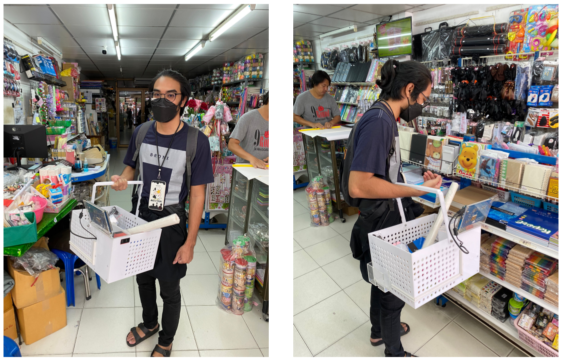
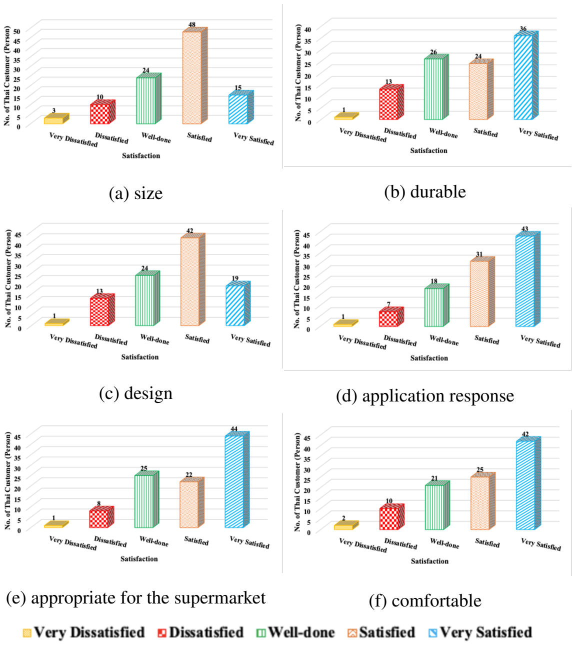


Figure 16: Testing at the local supermarket in Bangkok, Thailand



(a) size

(b) durable

(c) design

(d) application response

(e) appropriate for the supermarket

(f) comfortable

Figure 17: Results of Thai customer satisfaction questionnaires using the low-cost smart basket.

7 Conclusions

In this paper, the design and implementation of a low-cost basket based on an ARM SoC architecture with Raspberry Pi single board computer has been proposed based on the 3F theory of the product design applied for a hanging shopping basket. Hardware design has been applied by ARM system on chip architecture, while the proposed smart basket algorithm has been introduced in the software design to control the user interface and database. Experimental results show that the proposed smart basket implementation can be convenient for the lifestyle shopping experience in the supermarket. This innovation will replace the traditional one, which will help consumers maintain the social distancing during the COVID-19 pandemic.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment The authors would like to thank Watchara Yaima, Phompiriya Amiam, Nattawat Charoenkitja, and Aphirak Thitinaruemit, Department of Computer Engineering, Faculty of Industrial Technology, Suan Sunandha Rajabhat University in Bangkok, Thailand.

References

[1] 304 Industrial Park Co., Ltd., "Industry 4.0: Thailand's Turning Point for the Future of Manufacturing," 2022, [Online]. Available: https://www.304industrialpark.com/th/articles-detail/51/Industry-40-Thailand.[Accessed May. 30, 2022].

[2] S. Mekruksavanich, "Supermarket Shopping System using RFID as the IoT Application," in 2020 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON), 2020, 83-86, doi:10.1109/ECTIDAMTNCON48261.2020.9090714.

[3] R. Krishnamoorthy, K. Krishnan, B. Chokkalingam, S. Padmanaban, Z. Leonowicz, J. B. Holm-Nielsen, M. Mitolo, "Systematic Approach for State-of-the-Art Architectures and System-on-Chip Selection for Heterogeneous IoT Applications," *IEEE Access*, **9**, 25594-25622, 2021, doi:10.1109/ACCESS.2021.3055650.

[4] D. Kusswurm, Modern Arm Assembly Language Programming: Covers Armv8-A 32-bit, 64-bit, and SIMD, Apress Media, 2020.

[5] Panasonic Corporation, "RFID Based Walk-through Checkout Solution for Future Retail," 2018, [Online]. Available: https://news.panasonic.com/global/topics/2018/55288.html. [Accessed May. 30, 2022].

[6] M. Shahroz, M. F. Mushtaq, M. Ahmad, S. Ullah, A. Mehmood, G. S. Choi, "IoT-Based Smart Shopping Cart Using Radio Frequency Identification," *IEEE Access*, **8**, 68426-68438, 2020, doi:10.1109/ACCESS.2020.2986681.

[7] S. Karjol, A. K. Holla, C. B. Abhilash, P. V. Amrutha, Y. V. Manohar, An IOT based smart shopping cart for smart shopping, Cognitive Computing and Information Processing, Springer, 2017.

[8] S. Mekruksavanich, "The Smart Shopping Basket Based on IoT Applications," in 2019 IEEE International Conference on Software Engineering and Service Science (ICSESS), 2019, 714-717, doi:10.1109/ICSESS47205.2019.9040750.

[9] G. Arjun Kumar, Shivashankar, Keshvamurthy, K. Suni Kumar, R. Gatti, M.B. Hegde, "Design And Implementation Of Smart Shopping Basket," in 2020 International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), 2020, 714-717, doi:10.1109/RTEICT49044.2020.9315702.

[10] R. K. Megalingam, S. Vishnu, S. Sekhar, V. Sasikumar, S. Sreekumar, T. R. Nair, "Design and Implementation of an Android Application for Smart Shopping," in 2019 International Conference on Communication and Signal Processing (ICCSP), 2019, 0470-0474, doi:10.1109/ICCSP.2019.8698109.

[11] M. Dehghani, S. Kolahdouz-Rahimi, "An Automatic Generation of Android Application for WooCommerce," in 2019 International Conference on Computer and Knowledge Engineering (ICCKE), 2019, 194-200, doi:10.1109/ICCKE48569.2019.8964732.

[12] S. Prongnuch, S. Sitjongsataporn, "Differential Drive Analysis of Spherical Magnetic Robot Using Multi-Single Board Computer," *International Journal of Intelligent Engineering and Systems*, **14**(4), 264-275, 2021, doi:10.22266/ijies2021.0831.24.

[13] N. Boonyathanmig, S. Gongmanee, P. Kayunyeam, P. Wutticho, S. Prongnuch, "Design and Implementation of Mini-UAV for Indoor Surveillance," in 2021 International Electrical Engineering Congress (iEECON), 2021, 305-308, doi:10.1109/iEECON51072.2021.9440350.

[14] D.A. Patterson, J.L. Hennessy, Computer organization and design, ARM Edition: the hardware software interface, Morgan Kaufmann, 2016.

[15] Arm Limited, "Arm Cortex-A53 MPCore Processor Technical Reference Manual r0p4," 2018, [Online]. Available: https://developer.arm.com/documentation/ddi0500/j/Functional-Description/About-the-Cortex-A53-processor-functions. [Accessed May. 30, 2022].

[16] M.O. Ojo, S. Giordano, G. Procissi, I.N. Seitanidisn, "A Review of Low-End, Middle-End, and High-End Iot Devices," *IEEE Access*, **6**, 70528-70554, 2018, doi:10.1109/ACCESS.2018.2879615.

[17] C. Enrique, G. Venture, N. Yamanobe, "Applying kansei/affective engineering methodologies in the design of social and service robots: A systematic review," *International Journal of Social Robotics*, **13**(5), 1161-1171, 2021, doi:10.1007/s12369-020-00709-x.

[18] Y. Demkiv, "Product design process: 10 steps," 2022, [Online]. Available: https://qubstudio.com/blog/ten-steps-of-the-product-design-process. [Accessed May. 30, 2022].

[19] E. Krisnanik, T. Rahayu, "UI/UX integrated holistic monitoring of PAUD using the TCSD method," *Bulletin of Electrical Engineering and Informatics*, **10**(4), 2273-2284, 2021, doi:10.11591/eei.v10i4.3108.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 7, No. 4, 24-27 (2022)

www.astesj.com

ASTESJ

ISSN: 2415-6698

Using Dynamic Market-Based Control for Real-Time Intelligent Speed Adaptation Road Networks

Jamal Raiyn[*]

Al Qasemi Academic College, Computer Science, Baqa Al Gharbiya, 30100, Israel

ARTICLE INFO

Article history:
Received: 08 February, 2022
Accepted: 16 April, 2022
Online: 18 July, 2022

Keywords:
Data science
Speed adaptation
Market-Based Control
Computational intelligence

ABSTRACT

Traffic road management is becoming more complex due to limited resources and an increasing number of hybrid vehicles. Currently, a number of global classical computing tools are used to manage the traffic road network. This kind of classical computing is resource intensive and expensive, and in the best case, optimizing a traffic network takes a few hours. In traffic road networks, some classes of traffic are much more sensitive to communication delays than other classes. Delays in vehicle- to- vehicle communication cause traffic congestion and sometimes accidents. To honor users' preferences and to improve the road network performance, a market –based control approach is proposed that returns results in quantum space. This paper introduces a market-based control scheme with the goal to manage the traffic flow. Market-Based Control (MBC) is an economic model, which provides new solution to improve the travel data management and to reduce the probability blocking rate in connected autonomous vehicle. The Market- Based Model divided into fixed MBC and dynamic MBC.

1. Introduction

Road accidents have received the most attention from researchers due to its impacts economically [1]-[3]. It is proposed various conventional methods to reduce the number of traffic congestion states in road networks, and reduction of their negative effects and improvement of traffic safety [4], [5].

The route traffic management becomes more complex [6], when the number of vehicles increases in the road networks. The market- based control approach is proposed to describe the traffic flow and to handle traffic congestion road networks by considering the real-time traffic flow. In this case, computational intelligent methods are used, as market-based control model. The Market –based control provided different intelligence features, such autonomy, negotiation learn ability and reasoning. The novelty of market-based control is the dynamic management of traffic flow based on provider and supplier strategy.

The paper is organized as follows: Section 2 presents the traffic control problem. Section 3 describes the proposed concept based on market-based control. Section 4 presents the simulation and results' discussion. Finally, the conclusion summarizes the presented work and points to some future research directions.

2. The Traffic Control Problem

The traffic and road network management problem can be framed in terms of a market-based control model, in which a road network is divided in *m* road sections and allocated *n* activities. The AVs are considered to be consumers and the road sections to be the producers.

The supply of the producer *l* is expressed by r_{lS}.
The demand of the consumer *i* is expressed by r_{iD}.
consumer: $J_i(r_{iD}) \quad \rightarrow max$
producer: $J_l(r_{lS}) \quad \rightarrow max$
s.t. $\quad \sum_{i=1}^{n} r_{iD} = \sum_{l=1}^{m} r_{lD} \geq 0. \qquad (*)$
$r_{iD} \;\geq 0 \;\; f.a.i$
$r_{lS} \;\geq 0 \;\; f.a.l$

The optimization problem (*) is a road network management problem. The objective functions should be maximized at the same time. $R_{j(t,k)}$ is considered the total resources, and *j* is the activity that changes within road sections and within a designated time period. The activity of each road section *j* is presented as $x_{j(t,k)}$.

$$\sum_{i=1}^{m} x_{i,j}(t,k) = \; R_j(t,k), \;\; 1 \leq i \leq n$$

[*]Corresponding Author: Jamal Raiyn, Email: jamal.raiyn@gmail.com

https://dx.doi.org/10.25046/aj070404

Market- based control is employed to manage the traffic and the road network. To maximize utility, a given AV demands road sections with a minimal cost. The operator supplies road sections (resources) at a minimal cost to maximize its profits by offering available, road sections affording zero delay. The factors that influence the road traffic are expressed as cost functions that should be detected and reduced to zero. The cost functions consist of noise, delay and other factors that have a negative effect on road traffic. The cost functions can be collectively expressed as a general formulation:

$$J_k = \sum_{i\in I_C}(C_{ki}q_{ki}) + q_cC_k$$

for $\forall c, \forall k$, *where,* J_k presents the cost units for is the data anomalies in *k*th road sections, I_c denotes the sets of anomalies related to a cluster of road sections, and *C.* C_{ki} denotes the binary status of I_c which breaks down as

$$C_{ki}\left\{\begin{matrix} 0, if\ there\ are\ no\ anomalies \\ 1, if\ there\ are\ anomalies \end{matrix}\right\}$$

and q_{ki} is used to reflect anomalies between the two neighboring road sections and is taken from the anomalies matrix (i.e., $q_{ki} = m_{ij}(t)$). The road section strategy is to keep the supply higher than the minimum supply.

Road section Anomalies	section 1	section 2	section 3	section5
Delays	V	X	X	V
Noise	V	V	X	X
Accidents	V	X	X	X
Road work	X	X	X	X
Interference	X	X	X	X
Environmental factors	V	X	X	X
Human behavior	X	V	X	X

3. Market- Based Control Model

3.1. Classical Market-Based Control

In classical market-based control, there are three main components: Road section, autonomous vehicle, and a cognitive agent. The cognitive agent manages the negotiation between road section and autonomous vehicle. The road bids a resource in position *k* at time *t* with the price $P_k(t)$. The price of the resource $x_k(t)$ is affected by cost function. The cost function is affected by noise and delay. The autonomous vehicle searched resources with minimal cost function.

- Road traffic utility

The goal of road traffic is reducing the traffic congestion in road section *i* at time *t* and detecting the factors that caused congestions.

$$U_i^s(t) = P_i(t)\sum_{j=1}^{n} x_{ij}(t)$$

- AV utility

The goal of the AV is saving time, avoiding congestion, reducing CO_2.

$$U_i^b(t) = -\sum_{j=1}^{n} P_i(t)x_{ij}(t)$$

- Cognitive agent utility

The goal of cognitive agent is to manage a negotiation between AV and road section to assign road section (resources) with minimal cost.

$$U_i(t) = \sum_{j=1}^{n} x_{ij}(t)/N_i$$

Today, the digital devices, algorithm and data presentation are used classical computing. The road sections state is described in binary format, i.e. 0,1.

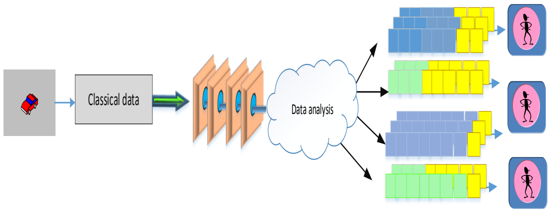


Figure 1: Fixed market-based control

3.2. Dynamic Market-Based Control

By using fixed market based control, all autonomous vehicles received update about the traffic sates at the same time, that performed in *t₀*. For that, when using normal car navigation systems, with the aim to avoid traffic jams are presented for each individual vehicle; however, because all vehicles are received nearly identical routes, the vehicles concentrate along the same route, causing traffic congestion. In dynamic market-based control, the road traffic is calculated in distributed mode, in different time slots, and according to AV demands. In distributed mode, the system estimates the traffic based on individual AV demands. Dynamic market based control model can present differing best routes for each of many vehicles that have the same starting points and destinations by distributing their routes as much as possible. Dynamic market-based control can prevent traffic congestion before it occurs instead of merely mitigating it.

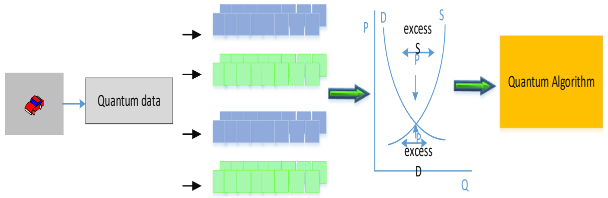


Figure 2: Dynamic market based control

4. MBC oriented computational Intelligence

The operations of the system model consist of four main phases: traffic road sensing; creating neural networks based on multi- layered deep learning technology, detecting anomalies and allocating resources as illustrated in Figure 3. The cognitive GOA obtains the road network, divides the roads into sections, and assigns a neuron to each road section to manage the traffic. The neuron aims to detect data anomalies and the factors that cause them. Then the road sections are classified by the effects of the anomalies, and on each road section is assigned a score that indicates its quality. The highest-scoring road sections offer balanced service, where the demand (for available road) is equal to the supply. If the demand is greater than the supply, it means, there is traffic congestion in road section. The traffic flow in the road section is dynamic which is described by using binary bits (0 and 1) along with the option of the bit being both 0 and 1 at the same time (superposition), this creates the third state. Eventually 3 states are 0,1 and 0–1. Quantum bit or Qubit is the basic unit of quantum information, that can be in a state of both 0 and 1 at the same time. The 3rd state creates more processing power. How states are processing in 3-bit systems described below.

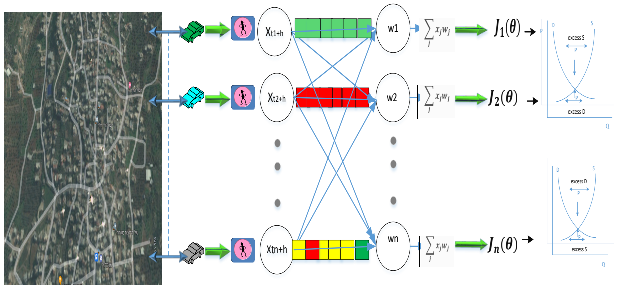


Figure 3: System model

5. Discussion and Evaluation

Travel data is collected by several tools, such as the magnetic loop detectors [7] and mobile services [8]-[14]. Traffic congestion appears when too many vehicles attempt to use a common transportation infrastructure with limited capacity [15, 16]. For a successful forecast of traffic flow, it ought to update the local travel data. Furthermore, it is important that the forecast model takes into consideration the anomalies data that caused in real-time [17]-[20]. Today, the digital devices, algorithm and data presentation are used classical computing which are described in binary format. As result, the road section state is described in binary format also as illustrates Figure 4. A fixed market-based control model is used in road traffic management, which evaluates a road network at time *t=0* for all road sections, despite the fact that traffic varies from section to section and changes from moment to moment. To overcome this disadvantage, a dynamic market- based control model is introduced, which assesses road sections at differing intervals of time, based on the condition of the traffic in each section as illustrates figure 5. The road section is described on quantum computing. Quantum computing uses binary bits (0 and 1) along with the option of the bit being both 0 and 1 at the same time (superposition), this creates the third state. Eventually 3 states are 0,1 and 0–1. Quantum bit or Qubit is the basic unit of quantum information, that can be in a state of both 0 and 1 at the same time. The 3rd state creates more processing power. How states are processing in 3-bit systems described below. Fixed market-based control can be implemented in Matlab for evaluation and simulation results. However, for implementing an agent goal oriented is needed dynamic environment. The challenge is creating virtual environment that support parallel processing of data in qubit form. The simulation platform will be described in Python to implement input data that is represented in qubit form.

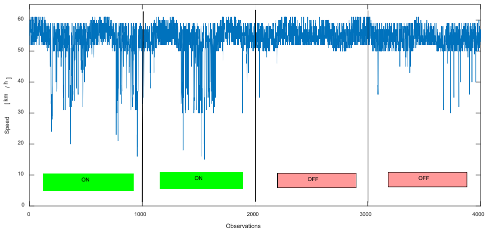


Figure 4: Speed calculation in binary format

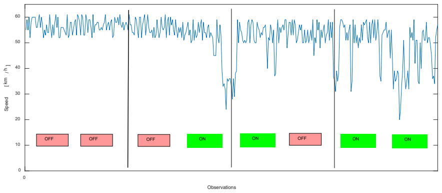


Figure 5: Speed calculation in qubit format

- Data collection

In classic computing, the data is presented in 0,1 which is based on a given threshold. However, in dynamic market based control, the data changes continuous depended on strategy of supplier and demander. In this case, the data is presented in two digits of binary form. One digit for the current state of the road section at time t_0 and one digit for the state of road section after t_0. In fact, the data presentation of dynamic market-based control is the form of quantum data. e.g. it needed qubit form to deal with those data.

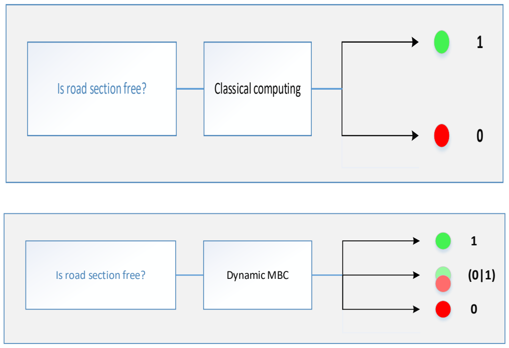


Figure 6: Binary vs. qubite form

- Algorithm based quantum computing

The dynamic market-based control which is proposed to process quantum data is based on advanced deep learning approach. Each road section is assigned a neuron. The task of the neuron is to manage the data in road section. The road section is divided in two parts. The first part is assigned 1 or 0 e.g. free or busy. The second part is described by (0|1) or (1|0) e.g. the traffic flow is in transition from 0 to 1 or 1 to 0 as illustrated in next figure. The road section is assigned a score for current state at t_0 and for transition at t_1.

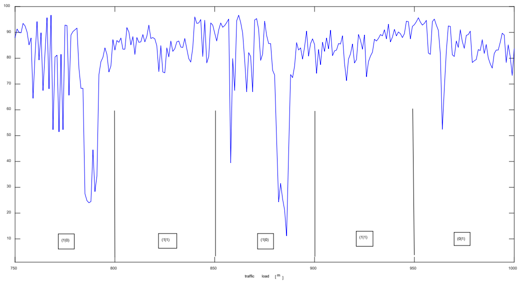


Figure 7: Speed description in qubit

6. Conclusion

Market- Based Model approach provided high reliability and the ability of a very efficient resource allocation even in very complex environments. In this paper, the travel data management is discussed. The travel data has been collected by mobile services. Due to urban coverage lack in cellular system are anomalies data and incomplete data caused congestions and accidents. Thus, the proposed system may play significant role on people life that had direct and positive impact at their life. The evaluation results of the proposed system show that dynamic market based control scheme can detect the any anomaly case early. Furthermore, the forecast results are too closest to the real traffic flow. Future work will consider anomalies data management caused in heterogeneous road networks.

References

[1] R. Elvik. "How much do road accidents cost the national economy?," Accident Analysis and Prevention 32,849-851,2000. http://dx.doi.org/10.1016/S0001-4575(00)00015-4.

[2] A. Alkhatib, T. Sawalha. "Techniques for Road Traffic Optimization: An Overview", Journal of Computer Science and Engineering, v. 11(4). 311-320, 2020. DOI: 10.21817/indjcse/2020/v11i4/201104063

[3] X. Zheng, M. Liu. "An overview of accident forecasting methodologies," Journal of Loss Prevention in the Process Industries 22 (4). 484-491, 2009. http://dx.doi.org/10.1016/j.jlp.2009.03.005

[4] W. Hong, K. Choi, E. Lee, S. Im, M., Heo. "Analysis of GNSS Performance Index Using Feature Points of Sky-View Image," IEEE Transactions on Intelligent Transport Systems, 15(2), 889-895, 2014. **DOI:** 10.1109/TITS.2013.2282631.

[5] J. Gong, Z. Yu, N. Chen. "An analysis of drivers' route choice behavior in urban road networks based on GPS data". In: Proc. of the Int. Conf. on Transportation Engineering ICTE, American Society of Civil Engineers, 515–520, 2007.

[6] L. Berzina, A. Faghri, M.T. Shourijeh, M Li. "Developing of a Post-Processing Automation Procedure for the GPS-Based Travel Time Data Collection Technique," Journal of Transportation Technologies, 4(1), 63-71, 2014. **DOI:** 10.4236/jtts.2014.41006

[7] T.M. Borzacchielo, "The use of data from mobile phone networks for transportation applications," TRB 2010 Annual Meeting, 2010.

[8] M. K. Habib, K. Ilhem, J. Casillas, A. Abraham, M. Alimi,"Adapt-Traf: An adaptive multiagent road traffic management system based on hybrid ant-hierarchical fuzzy model", Transportation Research Part C 42 ,147–167, 2014. DOI: 10.1016/j.trc.2014.03.003

[9] K.K. Santhosh, D.P. Dogra, P.P. Roy. "Anomaly Detection in Road Traffic Using Visual Surveillance: A Survey," ACM Computing Surveys, 53(6). 1-26, 2020. https://doi.org/10.1145/3417989.

[10] Y. Lv, S. Tang. "Real-time Highway Traffic Accident Prediction Based on the K-Nearest Neighbor Method." International Conference on Measuring Technology and Mechatronics Automation, Volume 3, 547-550, 2010.

[11] Z. Xiaoqiang, L. Ruimin, Y. Xinxin. "Incident Duration Model on Urban Freeways Based on Classification and Regression Tree." 2nd International Conference on Intelligent Computation Technology and Automation, TRB 2010 Annual Meeting, 2, 526-528, 2010. DOI:10.1109/ICECE.2010.934

[12] J. Raiyn. "Using Cognitive Radio Scheme for Big Data Traffic Management in Cellular Systems," International Journal of Information Technology and Management, 16 (3). 301-313, 2017. https://doi.org/10.1504/IJITM.2017.084985.

[13] T. Yucek, H. Arslan. "A Survey of Spectrum Sensing Algorithms for Cognitive Radio Applications," IEEE COMMUNICATIONS SURVEYS & TUTORIALS, VOL. 11, NO. 1, FIRST QUARTER 2009.pp. 116-129, 2009. http://dx.doi.org/10.1109/SURV.2009.090109.

[14] X. Zheng, M. Liu. "An Overview of Accident Forecasting Methodologies". Journal of Loss Prevention in the Process Industries, 22, 484-491, 2009. http://dx.doi.org/10.1016/j.jlp.2009.03.005.

[15] J. Andrada-Felix, F. Fernandez-Rodriguez. "Improving Moving Average Trading Rules with Boosting and Statistical Learning Methods." Journal of Forecasting, 27(5), 433-449, 2008. DOI: 10.1002/for.1068

[16] X. Zhang, J.A. Rice. "Short-term travel time prediction". Transport. Res. Part C: Emer. Technol. 11 (3–4),187–210, 2003. doi:10.1016/S0968-090X(03)00026-3.

[17] J. Raiyn. "Speed Adaptation in Urban Road Network Management", 17(2). 111-121. 2016. DOI 10.1515/ttj-2016-0010.

[18] R. R. Andrawis, F.A Atiya. "A New Bayesian Formulation for Holt's Exponential Smoothing", Journal of Forecasting, 28(3), .218-234, 2009. https://doi.org/10.1002/for.1094.

[19] X. Ma, Z. Dai, Z. He, J. Ma, Y. Wang, Y. Wang. "Learning Traffic as Images: A Deep Convolutional Neural Network for Large-Scale Transportation Network Speed Prediction, Sensors, 17(v1), 1-16, 2017. https://doi.org/10.48550/arXiv.1701.04245

[20] B. Chen, H.H. Cheng. "A Review of the applications of agent technology in traffic and transportation systems." IEEE Trans. Intell. Transport. Syst. 11 (2), 485–497, 2010. DOI. 10.1109/TITS.2010.2048313.

Advances in Science, Technology and Engineering Systems Journal
Vol. 7 No. 4, 28-33 (2022)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

Assessment of Electromagnetic-Based Sensing Modalities for Red Palm Weevil Detection in Palm Trees

Mohammed M. Bait-Suwailam*,[1,2], Nassr Al-Nassri[1], and Fahd Al-Khanbashi[1]

[1]Department of ECE, Sultan Qaboos University, Muscat, P.C. 123, Oman

[2]Remote Sensing and GIS Research Center, Sultan Qaboos University, Muscat, P.C. 123, Oman

ARTICLE INFO

Article history:
Received: 02 June, 2022
Accepted: 12 July, 2022
Online: 23 July, 2022

Keywords:
Date palm
Microwaves
Red palm weevil

ABSTRACT

In this paper, we investigate the utilization of three effective detection methods to identify potential threats of insects in date palm trees. The detection techniques presented here are the application of infrared radiation, microwave antennas and metamaterials based sensors. Experimental trials using IR radiation took place in a local farm. Moreover, the second sensing system is based on microwave antennas that are designed and numerically simulated at the 2.45 GHz-band. Lastly, the third detection method focuses on the design and development of low-power microwave sensor based on metamaterials concept. Based on the processed and analyzed results, the aforementioned sensing techniques are able to predict existence of red palm weevils within date palms.

1 Introduction

Food security and production are considered very essential infrastructures that have great impact on the way we consume food and dietary products towards human well-being and health. In principle, food systems include all stages of food production, including harvesting, processing, distribution and storage. It is understood that the quality of food products can be compromised or even degraded when farmers/supply chains are not taken safety measures and food quality standards into consideration. For instance, the use of an improper storage spaces for food or the use of unhealthy measures for food harvesting and production [1]-[3].

One of food quality degradation can be attributed to the surrounding environment. In real-life, growth and safety of crops is directly related to the environment. For instance, date palm trees are attacked heavily by many insects.

Oman relies heavily on the production of dates, which helps quite well in the development of the country's economy. The overall area of dates trees in Oman is over 30,000 hectares [4]-[6]. Currently, Oman has carried out numerous initiatives and projects over the Sultanate area in order to enhance and maximize the yield from such an important agricultural crop in Oman. Unfortunately, numerous palm pests deeply affect the health and safety of such trees, among which red palm weevil is a highly risk factor [7]-[8].

The life cycle of RPW is summarized in the diagram shown in Fig. 1, where RPW adults could last between 14 to 20 days. It is not an easy task to visualize by human eyes occurrence of an early stage infections within date palms, which is due to the hidden larvae.

Potential signs of infection, include very slow production rate of date palms, appearance of yellow sticky paste in palms (see Fig. 2), and appearance of apertures on the trunk [9]. It can be observed that infections in palm trees extend up to almost 2 meters high from soil level.

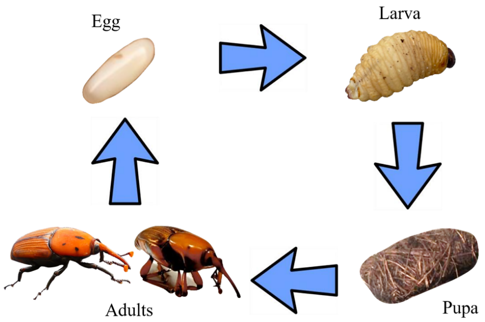


Figure 1: A conceptual chart showing red palm weevil growth cycle.

[*] Corresponding Author: Mohammed M. Bait-Suwailam (email: msuwailem@squ.edu.om)

https://dx.doi.org/10.25046/aj070405

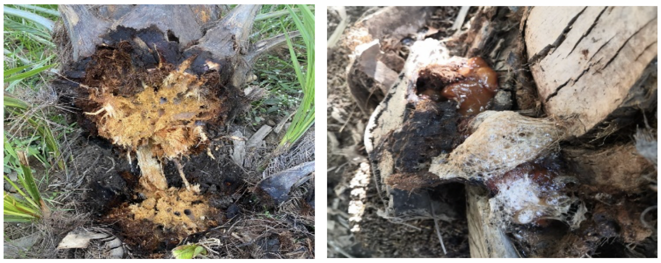


Figure 2: An image from a local date palm farm, where waste and watery liquid can be seen as a clear symptom of RPW existence in palm trees.

Microwave-based sensing and detection technique is mature and well-known for its non-destructive and hygienic testing modality for food sensing, evaluation and analysis, with more emphasis on the medical and industrial sectors [10], [11]. In principle, microwave sensing technology takes advantage of electromagnetic radiation within microwaves band, ranging from 300 MHz - 30 GHz. There have been extensive research studies focusing on the development of microwave sensing and imaging systems for a large number of applications, including but not limited to healthcare and biomedical treatments, industrial, scientific and security needs.

There are a number of techniques in the literature aiming to mitigation effects of infestation in date palms, with the objective of providing an early detection mechanism for possible infestations. The detection methods include but not limited to the use of trained dogs [12], sound-based sensors for monitoring palm pests activities [13], utilization of microwaves and high power systems treatment [14], [15], deployment of remote sensing and geographic information systems [16]–[18] and development of transmission lines based resonators [19]. The research work in [13] presented an experimental trials using acoustic signals to capture RPWs' activities through pizoelectric sensing device, where sound signals were recorded in-situ and then digitally processed. The application of microwave heating treatment was introduced in many studies, for instance the research work in [14], [15] and references therein, with the objective of treating the infected trees using high-power microwaves. Although earlier detection techniques based on acoustics, fiber optics, or even high-power microwave illumination are effective, they suffer from the complexity in the design and setup needed in monitoring a large number of RPW infestation or the need of deploying high-power radiators that are harmful and costly, thus limiting their practical use by farmers. The utilization of low-power sensing alternatives is expected to be more attractive and cost effective, which this research work aims to explore a number of alternatives based on low-power microwave sensing modalities.

In [20], a comparison between two sensing modalities was presented, namely: microwave-based sensory antennas and infrared-based sensing. However, the work presented here develops a more comprehensive study covering a number of attractive sensing modalities for the detection of RPW within date palm trees with more emphasis on the application of microwaves using antennas and highly-sensitive sensors based on metamaterials concept. In the remaining parts of the article, additional numerical and experimental results are presented and discussed.

2 Potential sensing and detection techniques for RPW detection

This section covers a comprehensive study of potential yet effective measures for sensing RPW in date palms, namely: 1) deployment of infrared-based sensing systems, 2) integrating microwave sensors (antennas) with a processing unit, and 3) deployment of microwave-based metamaterial sensors.

2.1 RPW sensing using Infrared-based imagery data

An infrared imaging is a very attractive technique, which utilizes radiation to identify distinctive features of temperature distribution. Nowadays, IR-based sensors are embedded in many electronic devices. This makes it easy to deploy IR in many applications, including detection of undesirable insects in food products.

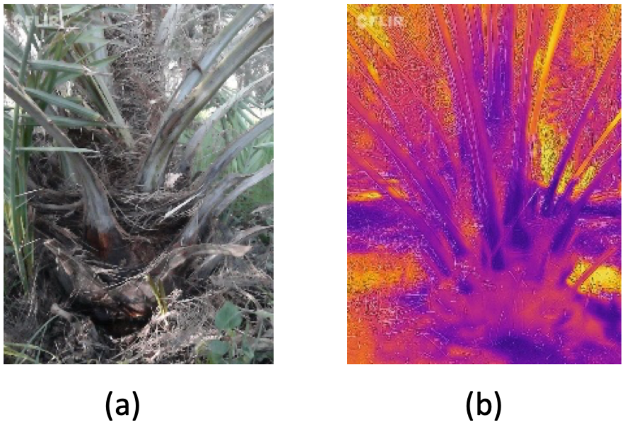


Figure 3: (a) Photo of a normal date tree and (b) the thermal heat distribution of the same palm tree by the IR sensor.

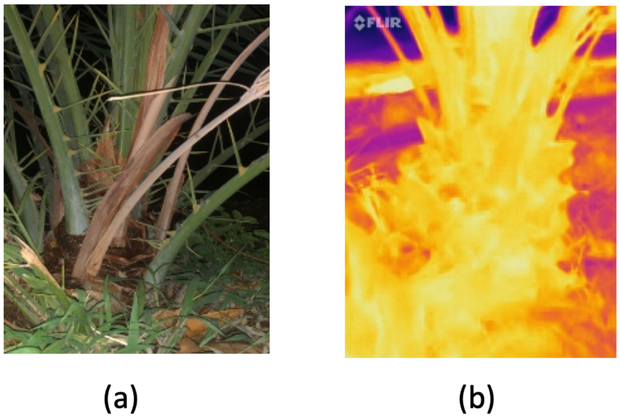


Figure 4: (a) Photo of RPW infestation case at night time and (b) the thermal heat distribution of the same palm tree by the IR sensor.

Since IR radiation has longer wavelength than visible band, thus it makes it difficult for humans to visualize IR thermal distribution. We can then utilize infrared sensors to inspect thermal distribution in palm trunk to detect any activities from the red palm weevil. IR images from healthy palm trunk were also recorded and processed for comparison purposes. Fig. 3 shows the temperature distribution within the trunk. As shown, the thermal heat distribution from the trunk was not significant enough, from which we could conclude that the palm is healthy. The heat distribution for the case of unhealthy trunk can be seen in Fig. 4 during night time, where more heat is expected from the trunk body itself.

Finally, the larva of the RPW was extracted. Interestingly, it was gluey. Additionally, it was observed that sufficient high radiated thermal energy was generated by the RPW and identified as yellow color from the sensor, as in Fig. 5(b).

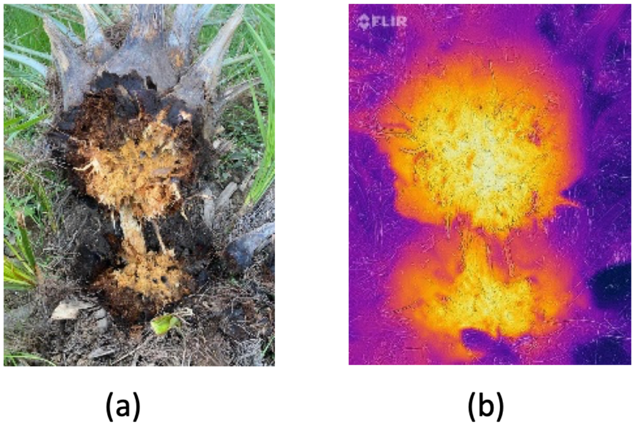


Figure 5: (a) Photo illustrating waste from infected palm trunk and (b) the thermal heat distribution of the same trunk by the IR sensor.

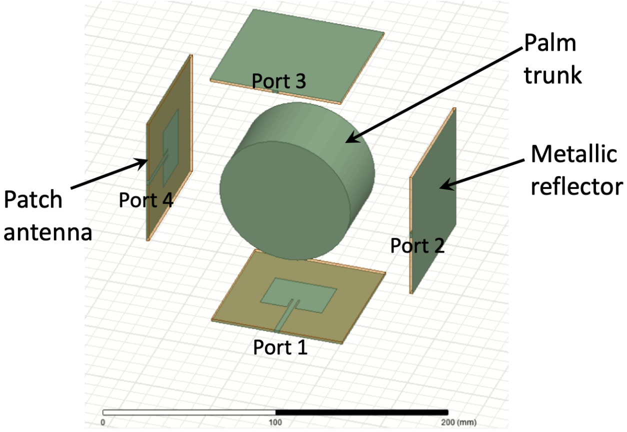


Figure 6: The developed model of the trunk surrounded with four microstrip antennas.

2.2 RPW sensing using microwave antennas

We investigate next the application of microwave-based sensors for red palm weevils detection. We have numerically modeled and simulated a small-scaled model comprising an unhealthy date palm and within its body a modeled pest. For ease of microwave sensing development, microstrip patch antennas are used here as the elementary sensors in order to predict any abnormalities within the coupled energy between the antenna elements. The spatial distance separating the radiating elements is maintained at 80mm. The modeled trunk has a height of 44.7mm and a diameter of 80mm (its dielectric constant, $\epsilon_r = 31.5$ and loss factor of 11.5). Since the solution to such 3D numerical model is mainly dependent on finite-element method of HFSS, storing the solution in each iterative process until convergence is reached would require huge amount of memory. To minimize the demands to such memory and disk space in the computer, a small-scale numerical model was constructed. Furthermore, the dimension of the RPW insect was modeled as a cylindrical object, with height = 2mm and radius = 3mm, with electric permittivity of 9.3 and bulk conductivity of 0.38 S/m. A single palm pest was placed in two different scenarios: within trunk's center and edge side of the palm, as illustrated in Figs. 7 (a)-(b). Note that all aforementioned constitutive parameters and associated losses were embedded into the models [15].

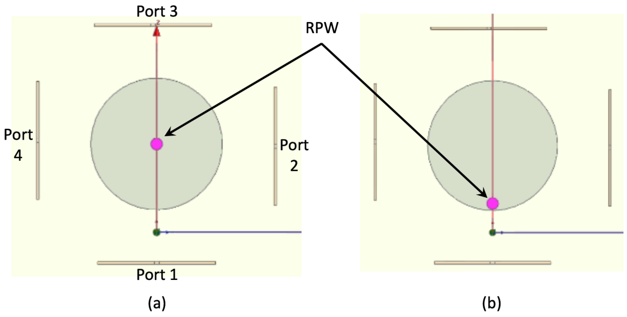


Figure 7: Top view of the trunk with a homogenized model mimicking RPW placed at two different scenarios, located at: (a) center, and (b) near to inner surface of the trunk.

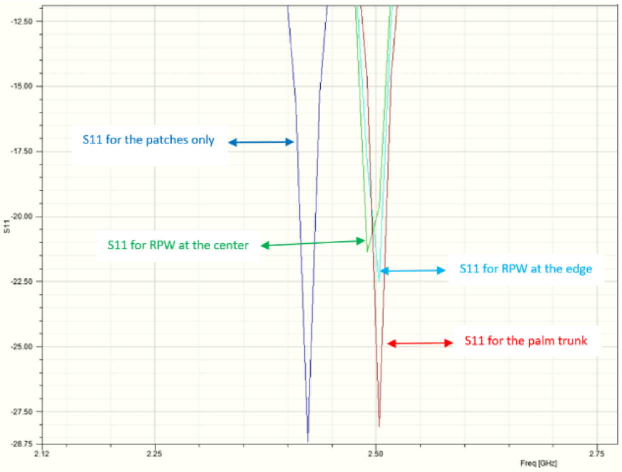


Figure 8: Simulated $|S11|$ for the healthy and infected palm trunk.

For comparison, a finite size structure consisting of the four patch antennas was considered. The case of the four patch antennas alone, i.e. no palm trunk, is expected to provide good impedance matching and low coupling coefficients between the individual antenna elements. The case of a finite-sized healthy trunk (modeled here as cylindrical lossy object) was numerically simulated, with the four antenna elements surrounding the trunk, as illustrated in Fig. 7. Good matching for the microwave antenna elements is still maintained as shown in Fig. 8. Furthermore, an appreciable shift in the transmission coefficient between ports 1 and 3 ($|S13|$) was achieved, which is referred to the presence of RPW inside the palm trunk. It is worth mentioning here that no major frequency shift in the peak of $|S13|$ was observed when comparing the cases of RPW at center and at the edge, as shown in Fig. 9.

The transmitted energy strength from port 1 to 4 was also considered. The sudden peak movement of $|S14|$ for the unhealthy palm is expected, as depicted in Fig. 10. Furthermore, we note that transmission coefficient, $|S13|$, is stronger than that from ports 1 and 4, which is due to the strong mutual interaction from port 1 to 3 in comparison to the case of coupled energy from ports 1 and 4.

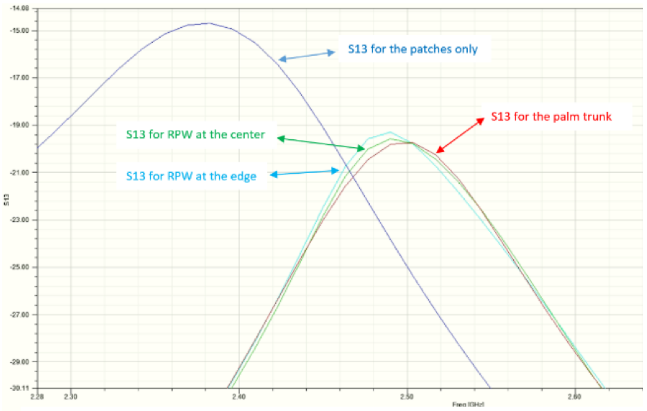


Figure 9: Simulated $|S13|$ for the healthy and infected palm trunk.

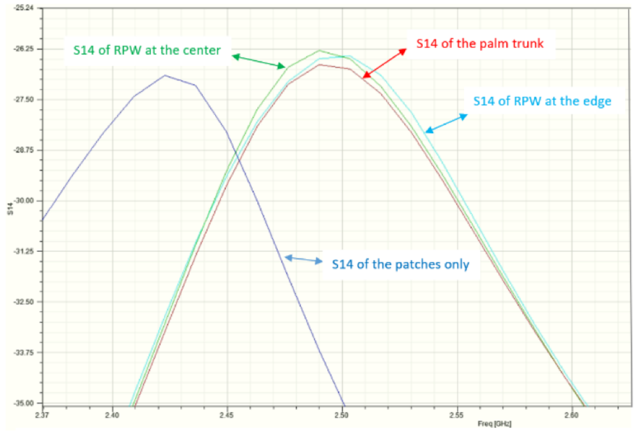


Figure 10: Simulated $|S14|$ for the healthy and infected palm trunk.

2.3 RPW sensing using microwave TL-based metamaterial sensors

In this section, we present a numerical study of an interesting sensing modality using microwave transmission line-based metamaterial sensors. Such sensors are engineered resonant particles or unit cells that have found lots of applications from engineering and science. Amongst the popular structures is the artificial magnetic conductor, also known as split-ring resonator (SRR). The dual of the SRR is the complementary split-ring resonator (CSRR), i.e., artificial electric conductor [21]. A schematic representation of the physical structure of the two inclusions can be seen in Fig. 11, where square rings are depicted for convenience.

In this study, a single unit cell is presented for ease of demonstration and compactness of the sensor. However, it is also permissible that more concentric or cascaded resonant units as a 1 dimensional array of rings could be explored. Details of the dimensions of the SRR/CSRR inclusions include length of the ring represented by L, a is the width of the SRR ring (or slot for CSRR), g is the cut gap and S_p is the spacing between the two rings. From the physical operational point of view, the SRR resonates well once excited through a normal external magnetic field component, while the CSRR resonates very well through an external normal electric field. Interestingly, the dual-operation mechanism of SRR/CSRR is based on the Babinet's principle, as presented in [21].

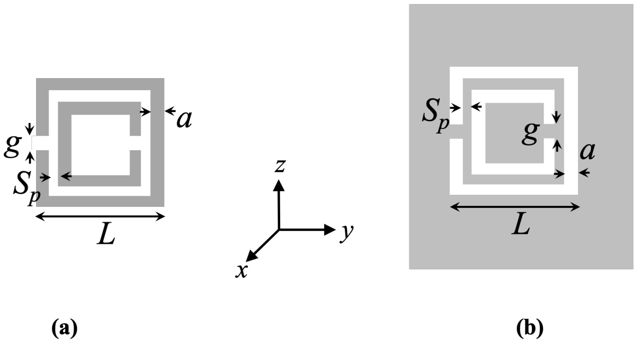


Figure 11: (a) the artificial magnetic conductor unit inclusion, and (b) its counterpart, the artificial electric conductor unit inclusion.

Fig. 12 depicts the developed microwave TL-based metamaterial structure. Two ports are deployed here in order to characterize the performance of the sensor through the computed transmission coefficient between the two ports. Once the sensing element is brought in close proximity to the date palm trunk, the resonance of the sensing element is expected to shift, and even further change is expected whenever RPW insects are within the inside of the trunk. The optimized dimensions of the developed CSRR microwave sensor at 2.45 GHz-band are available in Table. 2, along with the dimensions of the modeled 3D palm trunk.

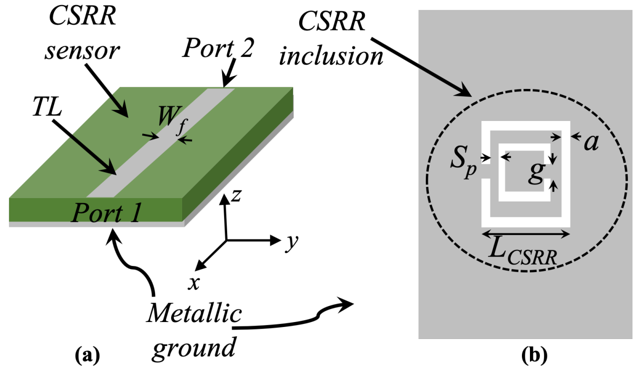


Figure 12: The developed TL-based metamaterial two-port sensor: (a) lateral view, and (b) bottom view of the metallic ground showing the adopted CSRR unit inclusion. The grey area here represents metallization.

Table 1: A comparison between several existing techniques suitable for RPW infestation in palm trees.

Techniques	sample required?	special set	Trained labor required?	Cost	Suitability
Sniffing dogs	required	No	Yes	Low	small (uncontrolled) farm
Acoustic	required	Yes	Yes	Moderate	small (controlled) farm
IR imaging	No	Yes	Yes	Moderate	large scale environment
microwave heating	No	Yes	Yes	High	High power radiation exposure
remote sensing	Yes	Yes	Yes	Moderate	small and large scale environment
microwave detection	Yes	Yes	Yes	Moderate	small (uncontrolled) farm

Table 2: The optimized parameters of the microwave detection structure.

Model parameters	Dimension(in mm)
Feedline width, W_f	2.52
Sensor thickness	0.8
Substrate length	40
Substrate width	30
Palm trunk diameter	50
Palm trunk height	40
Stand-off distance, S_z	2
L_{CSRR}	10.5
a	0.45
S_p	1.575
g	0.5

Fig. 13 shows a 3D view of the developed TL-metamaterial sensor that is placed on top of a date palm sample. For convenience, a stand-off distance, S_d, of 2mm was considered. Coupled energy via the sensor's ports is then computed and used as a metric to quantify the performance of the sensor in detecting RPW in date trees. Comparison was also made for the case of healthy palm trunk. Note that the resonance frequency of the reference sensor was 2.45 GHz. According to full-wave simulations, the close proximity of the sensor to a dry palm (representing dead one) resulted in a shift to the resonance frequency by 2.53% to lower frequencies due to the loss nature of the palm, while further increased shift to resonance frequency by 5.84% was incurred for the case of unhealthy (infected) palm tree.

Table. 1 provides a a comparison between several techniques that are available for the detection of infested date trees. It can be seen that each method has its own advantage and limitation, depending on various metrics.

3 Conclusion

In this paper, we presented a comparison study concerning the effectiveness of three detection techniques for red palm weevils in date palm trees. In the first detection method, detection of thermal variations within the date palm trunk was permissible using IR sensor. Experimental studies were performed in a small farm, containing a number of healthy and unhealthy palm trees. The experiments were carried out at different timings. Based on IR imagery data, quite noticeable heat was generated from the infested palm trees. A numerical model for the detection of RPW using four microstrip antenna elements working at 2.45 GHz was developed. Throughout the presented detection method using microwave antenna system and depending on the adopted number of ports, any sudden abnormalities in the response of either S11 or S21 magnitude and/or phase can be taken as an indicator for existence of RPWs in palm trunk.

Lastly, a 3D numerical model was developed for the case of deploying metamaterials-based microwave sensors. Based on the findings from this study, the two microwave-based detection models are considered as low-cost and effective tools for sensing the existence of RPWs.

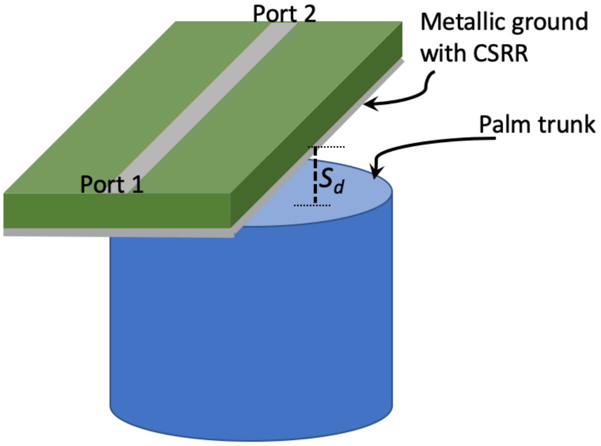


Figure 13: A 3D schematic showing the TL-based CSRR sensor in close proximity to a small-scale date palm trunk.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment The authors would like to acknowledge the support of Sultan Qaboos University, Oman, and Omantel company for providing the financial resources.

References

[1] T. Bosona, G. Gebresenbet"Food traceability as an integral part of logistics management in food and agricultural supply chain," Food Control, **33**, 32–48, 2013, doi: 10.1016/j.foodcont.2013.02.004.

[2] K. G. Grunert, "Food quality and safety: consumer perception and demand," European Review of Agricultural Economics, **32**(3), 369–391, 2005, doi: 10.1093/eurrag/jbi011.

[3] J. Zhang, L. Liu, W. Mu, L. M. Moga, X. Zhang, "Development of temperature-managed traceability system for frozen and chilled food during storage and transportation," Journal of Food, Agriculture and Environment, **7**(3), 28–31, 2009.

[4] A.S. Al-Marshudi, "Oman traditional date palms: production and improvement of date palms in Oman," Tropiculture, **20**(4), 203–209, 2002.

[5] R. Al-Yahyai, M. Khan, Date Palm Status and Perspective in Oman, in: Al-Khayri, J., Jain, S., Johnson, D. (eds) Date Palm Genetic Resources and Utilization, 1st Edition, Springer, doi: 10.1007/978-94-017-9707-8-6.

[6] R. Al-Yahyai, "Improvement of date palm production in the Sultanate of Oman," Acta Hort, **736**, 337–343, 2007, doi: 10.17660/ActaHortic.2007.736.32.

[7] K. Azam, S. Razvi, I. Al-Mahmuli, "Survey of red palm weevil, (Rhynchophorus Ferrugineus Oliver) infestation in date palm in Oman," Iraqi date Palms Network, http://www.iraqi-datepalms.net [accessed 4th October 2021].

[8] K. Al-Kindi, P. Kwan, N. Andrew, M. Welch, "Impacts of human-related practices on Ommatissus lybicus infestations of date palm in Oman," PLoS ONE, **12**(2), 1-17, 2017, doi: 10.17605/OSF.IO/HXPYP.

[9] D. Kontodimas,V. Soroker,C. Pontikakos,P. Suma,L. Beaudoin-Ollivier,F. Karamaouna,P. Riolo,"Visual identification and characterization of Rhynchophorus Ferrugineus and Paysandisia archon infestation," in: Handbook of Major Palm Pests Biol. Management, 187-208, 2016, John Wiley and Sons, doi: 10.1002/9781119057468.ch9.

[10] E. Fear, S. Hagness, P. Meaney, M. Okoniewski, M. Stuchly, "Enhancing breast tumor detection with near-field imaging," IEEE Microwave Magazine, **3**(1), 48-56, 2002., doi: 10.1109/6668.990683.

[11] E. Fear, P. Meaney, M. Stuchly, "Microwaves for breast cancer detection?" IEEE Potentials, **22**(1), 12-18, 2003, doi: 10.1109/MP.2003.1180933.

[12] S. Salem, "Accuracy of trained dogs for early detection of red palm weevil, Rhynchophorus ferrugineus Oliv. infestations in date palm plantations," Swift Journal of Agric. Res., **1**(1), 1-4, 2015.

[13] A. Hertzroni, V. Soroker, Y. Cohen, "Toward practical acoustic red palm weevil detection," Computers and Electronics in Agriculture, **124**, 100-106, 2016, doi: 10.1016/j.compag.2016.03.018.

[14] S. Nelson, "Review and assessment of radio-frequency and microwave energy for stored-grain insect control," Transactions of the ASAE, **39**(4), 1475–1484, 1996, doi: 10.13031/2013.27641.

[15] R. Massa, G. Panariello, D. Pinchera, F. Schettino, E. Caprio, R. Griffo, M. Migliore "Experimental and numerical evaluations on palm microwave heating for Red Palm Weevil pest control," Scientific Reports, **7**, 45299, 2017, doi: 10.1038/srep45299.

[16] F. Marzukhi, M. Said, A. Ahmad, "Coconut Tree Stress Detection as an Indicator of Red Palm Weevil (RPW) Attack Using Sentinel Data," International Journal of Built Environment and Sustainability, **7**(3), 1-9, 2020, doi: 10.11113/ijbes.v7.n3.459.

[17] H. Kurdi, A. Al-Aldawsari, I. Al-Turaiki, A. Aldawood, "Early detection of red palm weevil rhynchophorus ferrugineus (olivier), infestation using data mining," Plants, **10**(1), 95, 2021, doi: 10.3390/plants10010095.

[18] D. Kagan, G. F.Alpert, M. Fire, "Automatic large scale detection of red palm weevil infestation using street view images," ISPRS Journal of Photogrammetry and Remote Sensing, **182**, 122-133, December 2021, doi: 10.1016/j.isprsjprs.2021.10.004.

[19] M.M. Bait-Suwailam, "Numerical Assessment of Red Palm Weevil Detection Mechanism in Palm Trees Using CSRR Microwave Sensors," Progress in Electromagnetic Research Letters, **100**, 63-71, 2021, doi: 10.2528/PIERL21080303.

[20] F. Al Khanbashi, Nassr Al Nassri, M. Bait-Suwailam, "A Comparative Study of Electromagnetic-Based Sensing Modalities for Red Palm Weevil Detection in Palm Trees," in: 3rd IEEE Middle East and North Africa COMMunications Conference (MENACOMM), 69-73, 2021, doi: 10.1109/MENACOMM50742.2021.9678240.

[21] F. Falcone, T. Lopetegi, M. A. G. Laso, J. D. Baena, J. Bonache, M. Beruete, R. Marqués, F. Martín, M. Sorolla, "Babinet Principle Applied to the Design of Metasurfaces and Metamaterials," Physical Review Letters , **93**(19), 197401, November 2004, doi: 10.1103/PhysRevLett.93.197401.

Advances in Science, Technology and Engineering Systems Journal Vol. 7, No. 4, 34-41 (2022)

www.astesj.com

ASTESJ

ISSN: 2415-6698

The use of Integrated Geophysical Methods to Assess the Petroleum Reservoir in Doba Basin, Chad

Diad Ahmad Diad [1], Domra Kana Janvier[2], Abdelhakim Boukar[1,*], Valentin Oyoa[2]

[1]Department of Physics, Teachers Training Higher School of N'Djamena, University of N'Djamena, Chad, P.O. Box 460, N'Djamena, Chad

[2]Department of Mines, Petroleum and Water Resources Exploration, Faculty of Mines and Petroleum Industries, University of Maroua, P.O. Box 08, Kaele, Cameroon

ARTICLE INFO

Article history:
Received: 13 December, 2021
Accepted: 20 February, 2022
Online: 23 July, 2022

Keywords:
Static reservoir
Petrophisical parameters
3D seismic
Well log
Central African rift system
Doba Basin
Chad
Geology
Earth science
Africa

ABSTRACT

Hydrocarbon exploration and production has been successful in the central region of Doba basin, Chad, north-central Africa. In order to optimize the hydrocarbon production in this area, the combination of seismic and well log datas have been processed and analyzed to better characterize, image and capture the reservoirs. The 3D seismic and well log datas were used to obtain the horizon grids, fault polygons, and petrophysical parameters. The results show continuous and divergent horizons which are associated to differential subsidence and thickening of series on the inclined bedrock. The reservoir is an anticlinal structure where the hydrocarbons are trapped, in particular with reservoir levels interposed between the two stratigraphic sequences. Four mayors fault that cross the reservoir have been identified. The calculation of the average percentage of the encountered facies enabled to highlight the high percentage of sands compared to clays and clayey sands. Porosities are uniform in clays and sands in the two out of three wells, and higher in coarse sands. The permeabilities are average in sands and clays, but decrease in the fine sands. The 3-D static reservoir model integrated with structural and petrophysical parameters gives a better understanding of spatial distribution of the discrete and continuous reservoir properties. This work contributes to a future prediction of the reservoir performance, characteristics and production behavior in Doba basin.

1. Introduction

According to the national report of Chad, the petroleum production started in 2003 and reached its peak in 2004 (8.7 Mt). However, it has decreased steadily since 2006 (5.7 Mt in 2011), despite the commissioning of the new fields. This decrease is due to the technical difficulties and upwelling, which caused production to stop on several fields. This decline in production is mainly caused by the depletion of the fields. To remedy this deficit, the Chadian State has promoted several other blocks. The prospecting of these blocks begins with a better geological and petrophysical knowledge of the neighboring fields or in the exploitation of the basin. To contribute to this government-run project, this study proposes to determine a static reservoir model and petrophysical properties of cretaceous basin of Doba using logging and 3D seismic methods. This basin is part of the West African and Central African Rift system. The Lower Cretaceous formations mainly contain lacustrine sediments, the sandstone and mudstone interdicts are major reservoir strata, and above all, a large set of thick schists developed at the top of the formation, the Upper Cretaceous is dominated by the abundant riverine sandstone, with alluvial plain mudstone at the top. The Cenozoic formations contain coarse-grained clastic fluvial sediments.

The use of novel methods of geological and geophysical interpretation gives a realistic result in the oil and gas industry. Consequently, the use of the integrated approach of geology, geophysics, petrophysics, geostatistics and reservoir engineering for detailed characterization of reservoirs and their properties is a crucial approach, as also demonstrated in existing studies [1]-[3].

Hydrocarbon characterization can be achieved by the integration of seismic, well logs and geological data commonly used independently in hydrocarbon exploration and exploitation studies [4]-[9]. However, the detailed analysis of reservoir depends on how well data integration is performed [10]-[13].

[*]Corresponding Author: Abdelhakim Boukar, boukar.abdelhakim@gmail.com

https://dx.doi.org/10.25046/aj070406

Moreover, the use of seismic attributes such as dip, azimuth, amplitude, envelope, frequency and variance can enhance seismic interpretation [14]-[19] for a proper reservoir characterization. Furthermore, accurate 3D modeling of the reservoir can facilitate estimating of hydrocarbon reserves and determining the efficient way to recover as much of the hydrocarbon as possible for economic resource exploration [20]-[23].

2. Study area and geological setting

The Doba Basin is a sedimentary basin located in southern Chad, central Africa. It has a relatively high outcrop in the north and east, a lower outcrop in the south, a deeper outcrop in the north-west and forms a half bowl in the center, the Figure 1. The coordinates of the study area are 8° 31 'N and 16° 47' E. The study area forms a part of the West African and Central African Rift System, stretching across the central part of Africa, from Nigeria to Kenya. It is stored in the shear zone between these two countries.

The Doba Basin is a late Mesozoic sedimentary basin. This basin contains up to 10 km of the non-marine sediments that record the tectonic and climatic evolution of the region from the Lower Cretaceous to the present day. Those include stretching of the African plate during the tectonic plate movements, the orogeny and the dislocation of Gondwana, formed depressions in the center of the African continent which are the part of the West African and Central African rift system. The Doba Trench is the most depressed topographic part of the region. The ultrasound measurements undertaken in order to assess the depth of the Cretaceous sediments under the Continental Terminal detected the basement at 3500 m under the Paleo Chadian sands and sandstones, the Cretaceous marls and the Continental Interlayer. In [24], the author evaluated the respective thicknesses of these layers near Doba as follows: - Continental Terminal: from 0 to 700 m; - Cretaceous marls: from 700 to 1500 m; - Continental Intercalaire: from 1500 to 3500 m.

The presence of Bébo siliceous sandstones that could be linked to the Continental Intercalaire would attest to the existence of a Continental deposit above a highly developed ante-Karro surface. It is on the same surface, perhaps already deformed, that the Benue transgression would take place from west to east to the depressed area that started developing in the Doba region. The end of the Cretaceous transgression was accompanied by folds and dislocations as evidenced by faulted synclines of Léré-Figuil and dips of the Lamé series. The Bénoué basin and the Doba pit evolved separately from these deformations. In the West, the evacuation of detrital formations from Tertiary is oriented towards Niger. In the east, on the contrary, it accumulates in the Doba and Sarh pits subject to subsidence.

The Doba Basin has a basement composed of Precambrian metamorphic rocks covered by at most 7,500 m of continental deposits of thinner formations of the Lower Cretaceous (Mangara, Kedeni, Doba and Lower Kome), and thick formations of the Upper Cretaceous (Upper Kome and Miandoum). The Lower Cretaceous formations mainly contain lake sediments, sandstone and mudstone interlits are the major reservoir strata, and above all a large set of thick schists is developed at the top of the formation. The Upper Cretaceous is dominated by the abundant river sandstones, with alluvial plain mudstones at the top. The Cenozoic formations contain coarse-grained clastic fluvial sediments. The Doba pit is made up of the following layers: • Reservoir-Lower Cretaceous rock of Lower Kome, Kedeni and Mangara Formation and Upper Cretaceous of the Miandoum Formation of various arkosic sandstones. • Source-Lower Cretaceous rock of the Doba schists; Lacustrine schists of the first rift with a TOC content of 1 to 4% (organic matter of type I and III) • Cover rock, thick shales developed at the top of the formation, acting as good regional cover rocks.

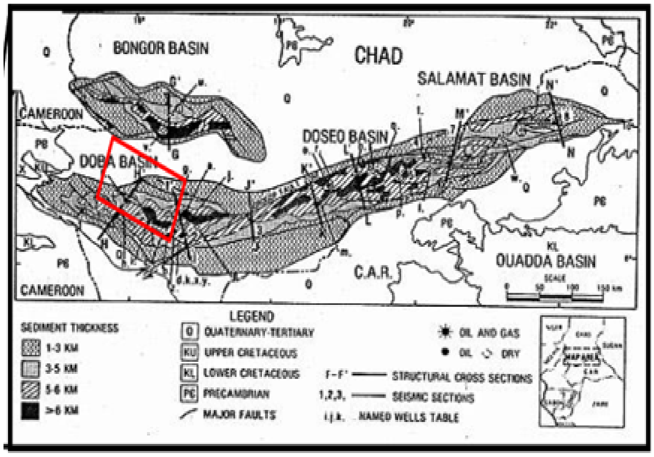


Figure 1: Map of studied location (red box), Doba Basin of Southern Chad which is a part of Central African Rift System, CAS [23].

3. Materials and methods

3.1. Materials and data

The implementation of this work was made possible due to the use of available documentation from the previous existing reports on fieldwork and also from existing reports of several oil companies. The computer interpretation and modeling tools included the following software: Golden Surfer, Microsoft Excel, Publisher, Google Earth, Google Map, and Pétrel from Schlumberger. The data used for this study include results of seismic reflection and well logging of the Doba basin in Chad. The logging data used include the following types: Gamma ray, which measures the natural radioactivity of a formation and is expressed in API (American Petroleum Institute); Neutron porosity which measures the porosity of a formation in percentage (%); Permeability which measures the permeability of a formation in milli darcy (mD); The Net to Gross which gives the ratio between the quantified fasciae and the global fasciae (NTG).

3.2. Methodology

The Petrel software was used as a 3D geocellular modeling package to represent the reservoir geology, structure, stratigraphic envelope, reservoir sublayers and faults in 3Dviewshowing structural and properties models. The reservoir volume was divided into a 3D mesh of cells, a typical geocellular model having hundreds of thousands to millions of cells in it. For each cell, a litho facies and rock properties such as porosity was assigned. The seismic data set comprises of both inline and cross line sections. Each well consist of the following well logs: gamma ray, spontaneous potential, calliper, density, neutron, sonic, and resistivity logs.

We extracted the faults using the guided manual picking method, proposed by Simpson and Howard (Simspon and Howard, 1996). The workflow methodology was based on the existing guide in a manual: • The interpreter points a polygonal line of a few points on a vertical section of the cube; • The ends of the line are automatically extended down and up the section; • The sub-vertical line obtained is then automatically readjusted laterally on the section, so that it follows the zones where the attribute marks the fault as accurate as possible; • The fault is now pointed by a line on a vertical section. This line is copied and translated into the adjacent vertical cross-section. • The new line is in its turn readjusted to the attribute of the new section. This approach has the advantage of allowing the interpreter to intervene at any time to rectify a misplaced point. The extraction is therefore well controlled, which makes it possible to obtain a good result, since this method is mostly manual and human-controlled. The multi-well geological correlation was performed in order to relate geophysics to geology. In particular, we performed cutting of the stratigraphic horizons on the well section, for the corrections of the structural model, the parameterization of the speed model, which allowed to transform modelling from the time domain to the depth domain. It was therefore necessary to make a geological correlation of various boreholes carried out. We linked the four wells based on their calculated average porosities on the one hand, the Net-to-Gross ratio, and the river fasciae identified by logging results on the other. The Figure 2 illustrates the workflow.

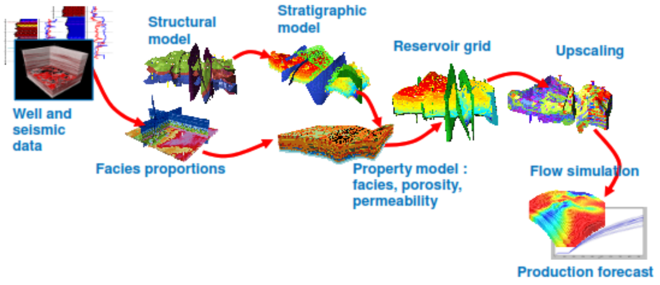


Figure 2 flowchart of methodology used

4. Results

4.1. Result of sedimentary interface termination

Upon applying the attribute combination cosine of phase and relative acoustic impedance, we observed the interface termination. This surface was compared to the erosion truncation that we extracted from the seismic cube, as presented in Figure 3.

After analysis, it is possible to see that the structure observed in nature is the same as in our seismic image. Details were not clearly visible on a normal profile without the use of these attributes. This result is a first step towards the extracting of geological objects aimed to derive information that an interpreter is looking for.

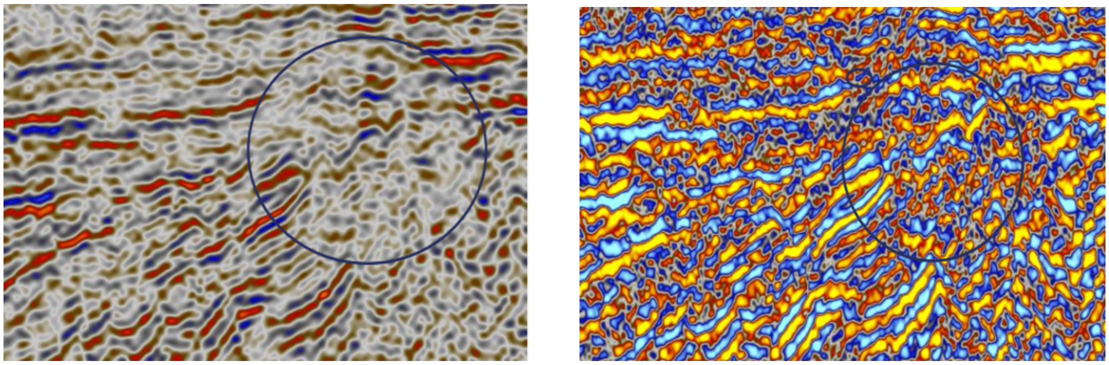


Figure 3: Extraction of an interface termination before picking

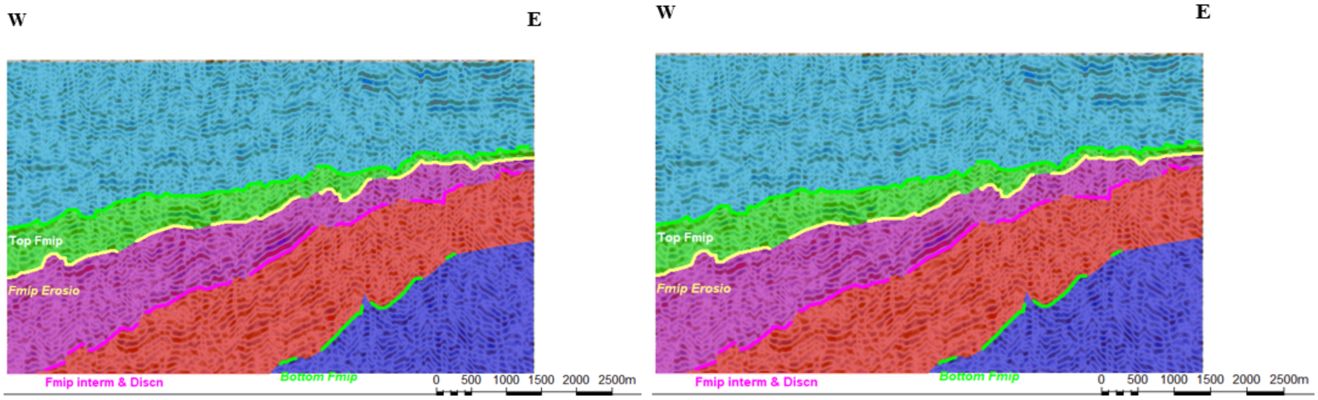


Figure 4: Blocks of cut-out horizons

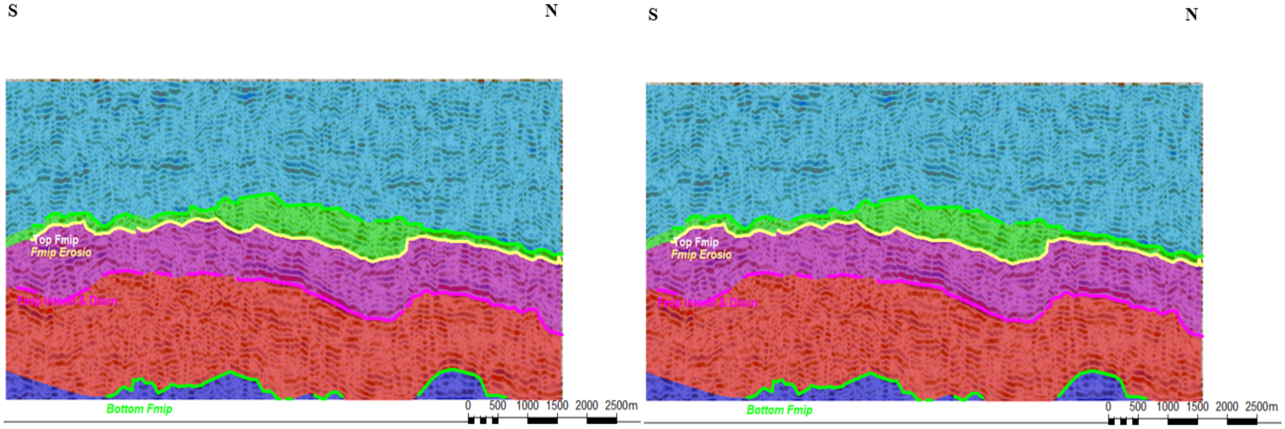


Figure 5: Reconstructed blocks of horizons

4.2. Result of horizon

Based on the performed graphical plotting, we are initiating the projection into space of the horizons being tracked in order to better highlight the stratigraphic sequences of the sedimentary units. From a geological point of view, we can distinguish the two visible sequences. Figure 4 shows the continuous configurations of the divergent types. They refer to a differential subsidence and a thickening of series on an inclined substratum. The continuous configurations of the subparallel types are implying subsidence with a constant rate of sedimentation.

Figure 5 is a visualization of the division made by reconstructing different blocks with a range of colors, each representing a horizon block. We can see here the succession of layers deposition with the migration of off lap breaks which are more distinct in some cases, while less visible in others, see Figure 5.

4.3. Fault extracting results

After applying the ant tracking and meta attribute, we obtained the image to detect and track flaws. However, to avoid errors during manual extraction, we applied Meta attribute after the Ant tracking, in order to improve this detection. The faults appear much clearer and are easily distinguished from the horizons. Figure 6, 7, 8 and 9 demonstrate the obtained results.

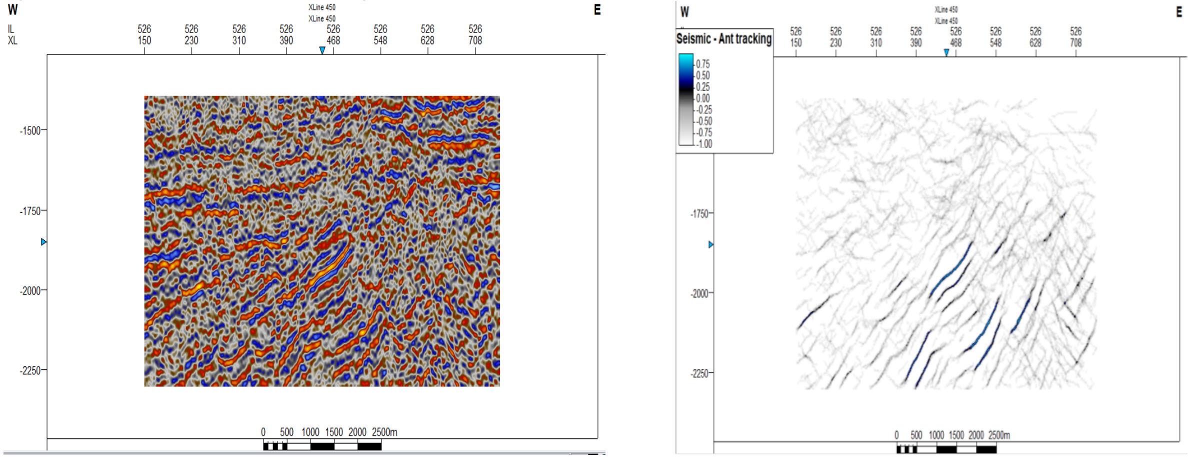


Figure 6: Fault extraction ant tracking on inline

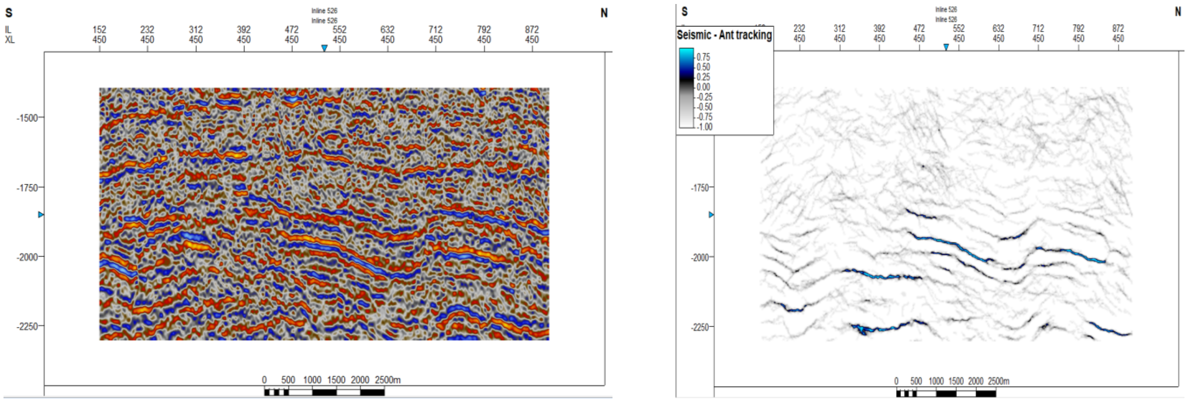


Figure 7: Fault extraction ant tracking on crossline

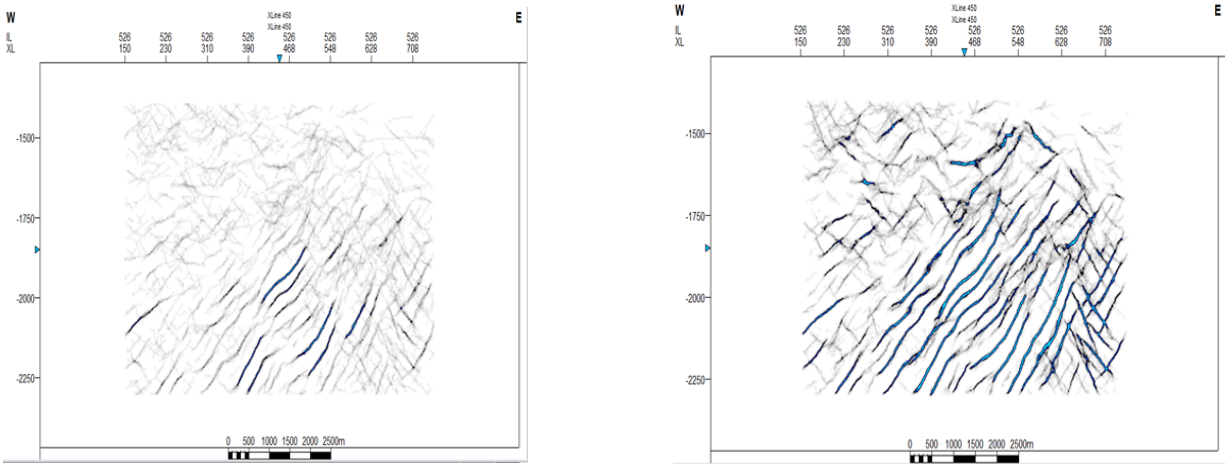


Figure 8: Result of apply the meta attribute on ant tracking (inline)

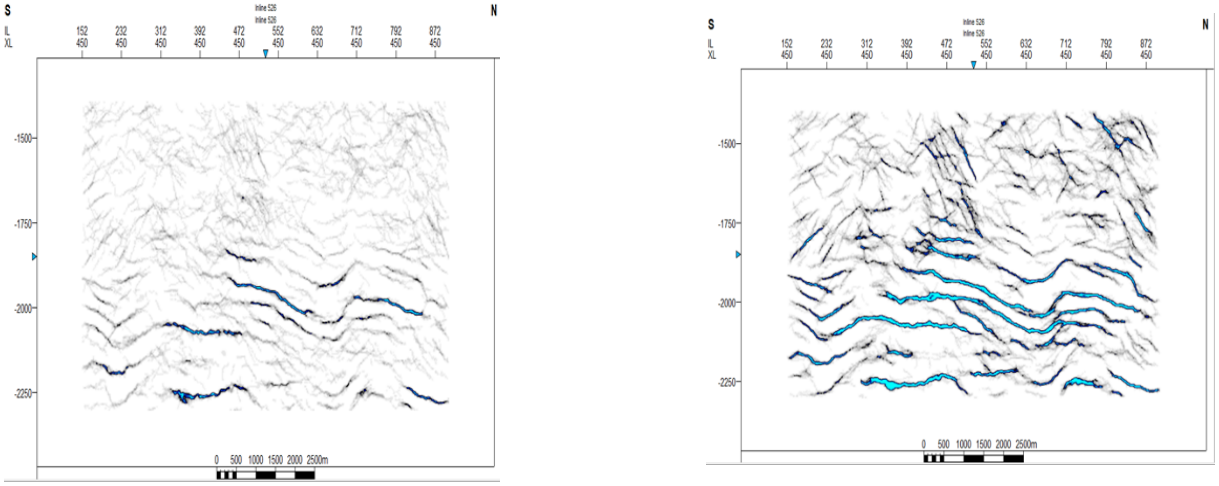


Figure 9: Result of apply the Meta attribute on ant tracking (crossline)

The three-dimensional model of the four major faults is shown in the Figure 10.

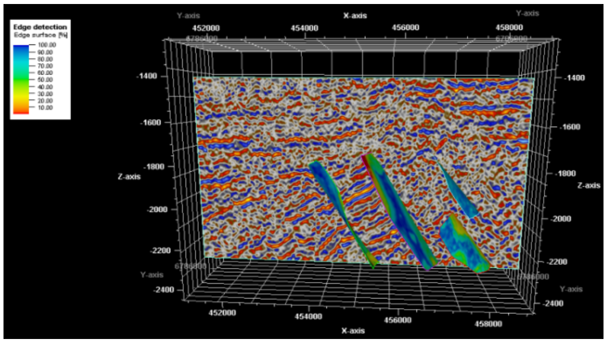
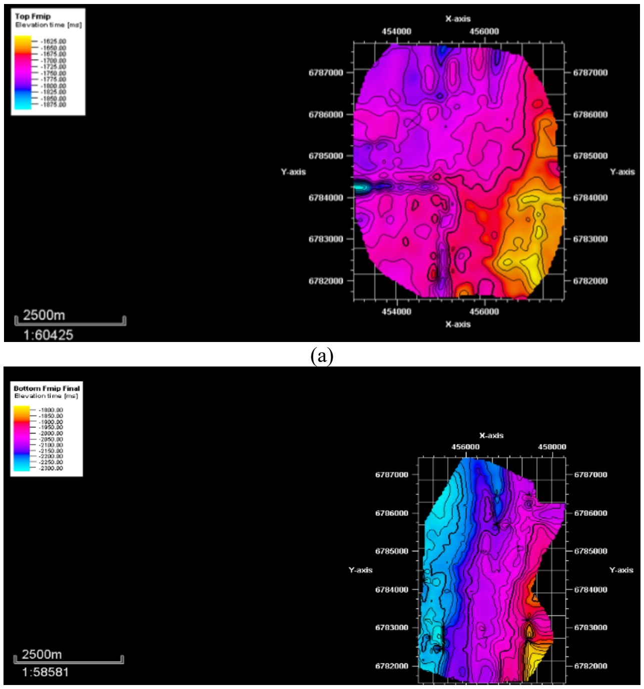


Figure 10: 3D modeling of major faults (inline)

4.4. Result of Top and Bottom of Horizons

The maps of top and bottom horizon show a variation of color on a time scale in milli second (Figure 11). The low values are associated to the first arrivals and could be interpreted as the sub-surface areas, while the high values inform on the depth layers. The pace of the curves show high depressions that refer to a more pronounced folding (turbulence zones) especially on the top horizon, where a complex mixing of the layers occurred.



(a)

(b)

Figure 11: Top and bottom horizon(a top and b bottom)

The structural model obtained after picking and tracking of geological setting is given in Figure 12. The four major faults have been centralized inside the model. This model can be sued to plot static model.

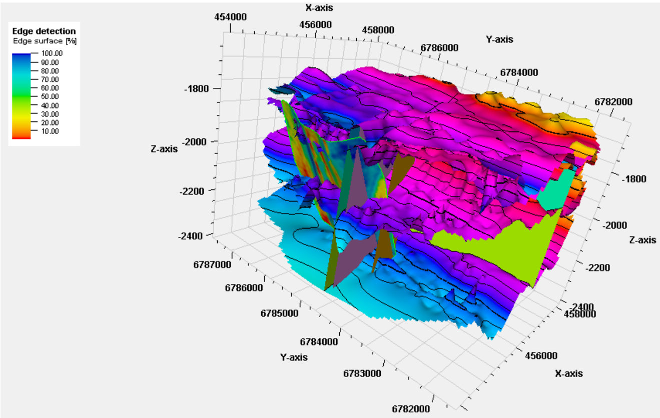


Figure 12: Structural model

4.5. Petrophysical model

From the perspective of relating seismic to geology, it is important to correlate the wells to highlight the geology of the basin and to make the link between the different sedimentary units. Here we presented the three wells drilled in the study area, respectively the A16; B2; B9.For this purpose, our correlation is made according to the identified fascies, the average porosity, and the average Net to Gross ratio.

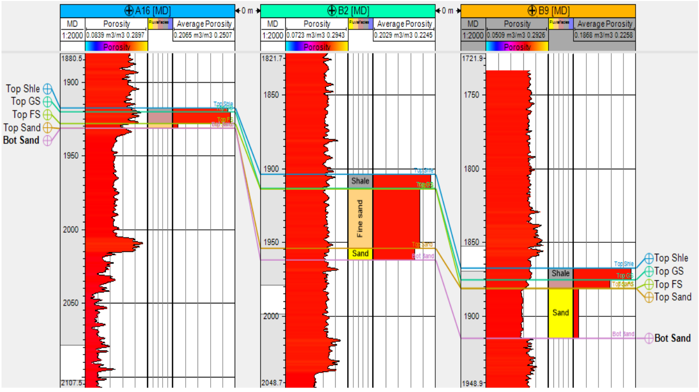


Figure 13: Average fascias per zone in the wells

The Figure 14 below illustrates the average porosity calculated in each well based on the overall porosity records. We can therefore notice high percentage in the coarse sands, decreasing slightly in clays and then becoming medium in sand. However, by applying the calculation of the average porosity in these wells as a function of general porosity, we noticed that coarse sands retain their greater volume ratio compared to middle- or fine-grained sand. However, the proportion of clays is even greater than that of coarse sand.

Here we applied our correlation according to the Net to Gross, which emerges the ratio of a pure fascies on all the fasciates present in order to characterize the quality of the reservoir in question. According to the analysis, we can see a large variation to the right, for clay and coarse sand, which is totally opposite to that of sand which has a large deflection to the left. However, when we calculate the average Net to Gross (Figure 15), the sand which had a large deflection to the left totally drops, but clay, fine sand and coarse sand have a high percentage, with a slightly large peak for coarse sand.

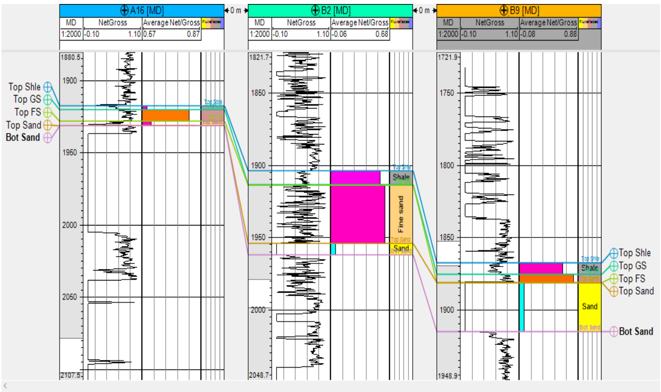


Figure 14: average porosity in each well

Compared to clay, we can therefore conclude that the percentage of the pure coarse sand and pure fine sand are in excess of all the fasciae present, and that implies a sandstone reservoir. The following Figure 15 shows the results.

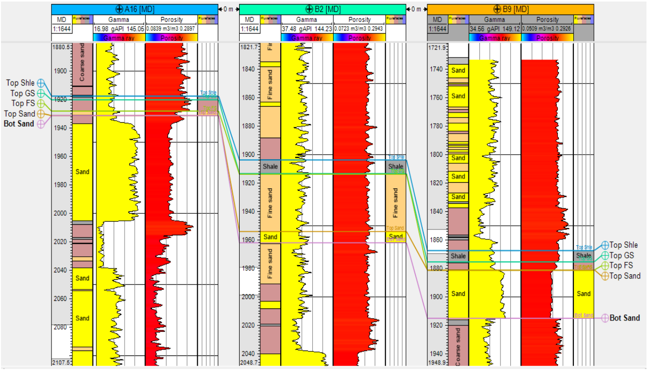


Figure 15: Net to cross average in each well

Following previous calculations, we can move on to the stratigraphic model which consists inplotting the structural model with the petro-physical properties and the fascies. We therefore plot the different zones created on our three wells with the determined fascies and the properties calculated above. We first present the rock volume based on a color scale which gives values of the gross volume. Each color corresponds to any volume, and subsequently the reservoir grid. We present them in the following Figure 16.

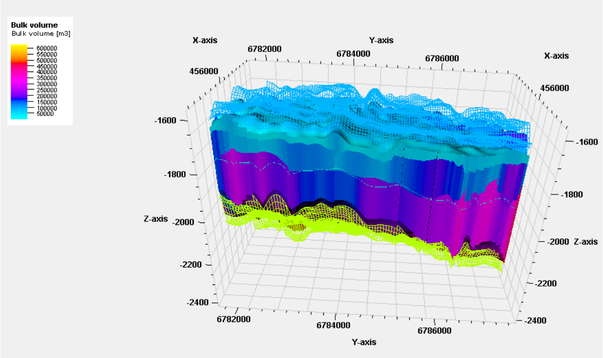


Figure 16: Result of rock volume

Upong modelling fascies and petro-physical properties, the next step included plotting grid. This model is not completed because the physical petro properties and fascies have exceeded the extent of the grid.

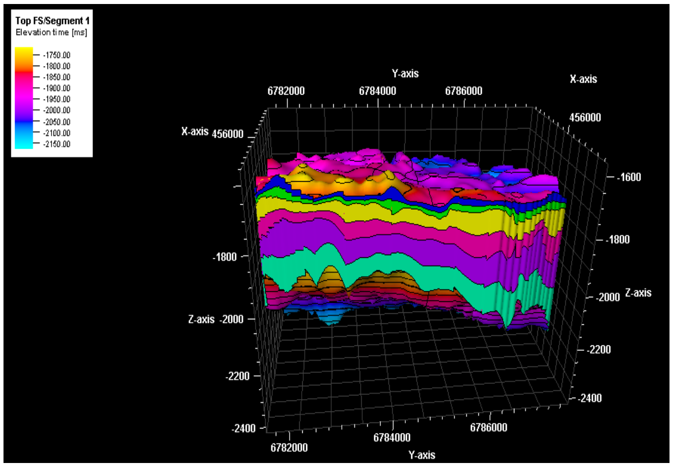


Figure 17: Result of static reservoir model

5. Discussion

The major research question of this study included interpretation of the seismic data of one of the blocks in the Doba basin in Chad. Through this interpretation we observed that the reservoir in question dates from the Cretaceous period and contains clay and sandstone rocks inside. It is affected by tectonic and stratigraphic accidents which allowed us to delimit the reservoir. In fact, it is an anticlinal structure where hydrocarbons are trapped into the two to three reservoir levels interposed between the first stratigraphic sequence and the second, which is subdivided in the two divergent types of the continuous configurations. Thus, this refers to a differential subsidence and a thickening of the series on an inclined substratum.

The continuous configurations of the sub parallel types, imply the subsidence with a constant rate of sedimentation. The calculation of the average percentage of the encountered fasciae enabled to highlight the high percentage of sands compared to clays and clayey sands. The porosities are uniform in normal clays and sands in the two out of three wells, and higher in coarse sands, respectively. The permeability is average in medium-grained sands and clays, but decreases in the fine-grained sands. We also noted that the waterproof roof was not extended over the entire block due to the tight mesh. Therefore, we extended a few more in lines to maximize the reservoir grid while controlling the extension of the porous fasciae in certain areas which did not present trapping structures.

Based on the performed analysis we interpret this area to be a sandstone reservoir, which is an anticline dating from the Cretaceous period. It was developed during the formation of the West African and Central African Rift system, and now extend through the central part of Africa, from Nigeria to Kenya through Chad. Thus, we achieved the major objective of this study, which aimed at a clear reinterpretation of the block. The limitations of the study include the following data shortage. Since our well data did not include especific information regarding resistivity and density, we were unable to determine saturation and from there estimate the reserves in place. It is therefore recommended for future studies to extend the research by including resistivity and density of porous materials for more detailed geological analysis.

6. Conclusion

Seismic and well logs data have been used for static modeling and petrophisical characterization of the reservoirs in Doba field, Chad. The reservoir is an anticlinal structure where the hydrocarbons are trapped, in particular with reservoir levels interposed between the two stratigraphic sequences. In this study we have identified the four major faults that crossed the reservoir. The calculation of the average percentage of the facies demonstrated high percentage of sands compared to clays and clayey sands. The porosities are uniform in clays and sands in the two out of three wells, and higher in coarse sands. The permeability is average in sands and clays, but decreases in fine sands. The 3-D static reservoir model integrated with structural and petrophysical parameters has been plotted and visualised on a series of the presented graphs. This paper contributes to the future exploration of geological resources of Chad, where the presented results could be used for correlation in the field of Doba.

References

[1] Y.Z. Ma,, Uncertainty analysis in reservoir characterization and management: how much should we know about what we don't know? In: Ma, Y.Z., La Pointe, P.R. (Eds.), Uncertainty Analysis and Reservoir Modeling: AAPG Memoir, **96**, 1–15, 2011.

[2] O. O. Osinowo,, Ayorinde, J.O., Nwankwo, C.P., Ekeng, O.M., Taiwo, O.B., Reservoir description and characterization of Eni field offshore Niger Delta, southern Nigeria. J. Petrol. Explor. Prod. Technol. **8** (2), 381–397., https://doi.org/10.1007/s13202-017-0402-7, 2018.

[3] X.Y. Yu, Y.Z. Ma,, D. Psaila, Pointe, P.L., Gomez, E., Li, S., Reservoir character- ization and modeling: a look back to see the way forward. AAPG Mem. **96**, 289–309, 2011.

[4] A.W. Mode, Anyiam, A.O., Reservoir characterization: implications from petro- physical data of the "Paradise-Field", Niger Delta, Nigeria. Pac. J. Sci. Technol. **8** (2), 194–202, 2007.

[5] A.P., Aizebeokhai, I. Olayinka, Structural and stratigraphic mapping of Emi field, offshore Niger Delta. J. Geol. Min. Res. **3** (2), 25–38, 2011.

[6] A.O. Adelu, Sanuade, O.A., Oboh, E.G., Offeh, E.O., Adewale, T., Mumuni, O.S., Oladapo, I.M., Omolaiye, E.G., Hydrocarbon field evaluation: case study of 'Tadelu' field shallow offshore Western Niger Delta, Nigeria. Arabian J. Geosci. **9** (2), 116. https://doi.org/10.1007/s12517-015-2028-8, 2016.

[7] O.A. Sanuade, A.O. Akanji, Olaojo, A.A., Oyeyemi, K.D., Seismic interpretation and petrophysical evaluation of SH field, Niger Delta. J. Petrol. Explor. Prod. Technol. **8** (1), 51–60, 2017b.

[8] Sanuade, O.A., Kaka, S.I., Sequence strati graphic analysis of the Otu Field, onshore Niger Delta using 3D seismic data and borehole logs. Geol. Q. **61**(1), 106–123. https://doi.org/10.7306/gq.1318, 2017.

[9] A.O. Akanji, O.A. Sanuade, Kaka, S.I., Balogun, I.D., Integration of 3D seismic and well log data for the exploration of kini field, offshore Niger delta. Petrol. Coal **60** (4), 752–761, 2018.

[10] S. Chopra, R.J. Michelena, Introduction to Reservoir Characterization. Special edition of The Leading Edge, 35–37, 2011.

[11] T.A. Adagunodo, L.A. Sunmonu, M.A. Adabanija, Reservoir characterization and seal integrity of Jemir field in Niger Delta, Nigeria. J. Afr. Earth Sci. 129, 779–791, 2017.

[12] O. O. Osinowo, J.O. Ayorinde, C.P. Nwankwo, O.M. Ekeng, O.B. Taiwo, Reservoir description and characterization of Eni field offshore Niger Delta, southern Nigeria. J. Petrol. Explor. Prod. Technol. **8** (2), 381–397. https://doi.org/10.1007/s13202-017-0402-7, 2018.

[13] Y.C. Ajisafe, B.D. Ako, 3-D seismic attributes for reservoir characterization of "Y" field Niger Delta, Nigeria. IOSR J. Appl. Geol. Geophys. **1** (2), 23–31, 2013.

[14] E.A. Ayolabi, A.O. Adigun, The use of seismic attributes to enhance structural interpretation of Z-field, onshore Niger Delta. Earth Sci. Res. **2** (2), 223–238, 2013.

[15] E. Jegede, B.D. Ako, Adetokunbo, P., Edigbue, P., Abe, S.J., Seismic stratigraphy and attribute analysis of an offshore field, Niger Delta, Nigeria.

Arabian J. Geosci. **8** (9), 7537–7549. https://doi.org/10.1007/s12517-014-1665-7, 2014.

[16] K.D. Oyeyemi, A.P. Aizebeokhai, Seismic attributes analysis for reservoir char- acterization; o shore Niger Delta. Petrol. Coal 57 (6), 619–628, 2015.

[17] P. Adetokunbo, A.A. Al-Shuhail, S. Al-Dossary, 3D seismic edge detection using magic squares and cubes. Interpretation **4** (3), T271–T280, 2016.

[18] O.A. Sanuade, S.I. Kaka, Sequence strati graphic analysis of the Otu Field, onshore Niger Delta using 3D seismic data and borehole logs. Geol. Q. 61 (1), 106–123. https://doi.org/10.7306/gq.1318, 2017.

[19] L. Adeoti, N. Onyekachi, O. Olatinsu, J. Fatoba, M. Bello, Static reservoir modeling using well log and 3-D seismic data in a KN eld, o shore Niger Delta, Nigeria. Int. J. Geosci. **05** (01), 93–106, 2014.

[20] O. Duvbiama, J. Ikomi, 3D Static Modelling of an O shore Field in the Niger-Delta. Nigeria Annual International Conference and Exhibition Held in Lagos, Nigeria. 31 July – 2 August 2017 SPE-189161-MS, 2017.

[21] G.O. Emujakporue, Petrophysical properties distribution modelling of an onshore field, Niger Delta, Nigeria. Curr. Res. Geosci. **7** (1), 14–24, 2017.

[22] Bouteyre G., J. Cabot, J. Dresch; Observations sur les formations du continental terminal et du Quaternaire dans le bassin du Logone (Tchad). Bulletin de la Société Géologique de France 1964;; **S7-VI** (1): 23–27, doi: https://doi.org/10.2113/gssgfbull.S7-VI.1.23

[23] R. E. HOWARD, Landmark Graphics Corporation. Method for attribute tracking in seismic data. Brevet American 5, 056 – 066, 1991

[24] M. (n.d) Aremu, Chad Prepares to Be an Oil Producers. Oil and Gas Online. Retrieved on September, 29th 2015 from http://www.oilandgasonline.com/doc/chad-prepares-to-be-an-oil- producer-0001

Advances in Science, Technology and Engineering Systems Journal
Vol. 7, No. 4, 42-48 (2022)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

Computer Vision Radar for Autonomous Driving Using Histogram Method

Hassan Facoiti*,[1], Ahmed Boumezzough[1], Said Safi[2]

[1]*Applied Physics and New Technologies Team, Sultan Moulay Slimane University, Beni Mellal, Morocco*

[2]*Laboratory of Innovation in Mathematics, Application and Information Technologies, Sultan Moulay Slimane University, Beni Mellal, Morocco*

ARTICLE INFO

Article history:
Received: 24 May, 2022
Accepted: 12 July, 2022
Online: 27 July, 2022

Keywords:
Radar
Autonomous driving
ADAS
Computer vision
Line lane detection
Histogram

ABSTRACT

Mobility is a fundamental human desire. All societies aspire to safe and efficient mobility at low ecological and economic costs. ADAS systems (Advanced Driver Assistance Systems) are safety systems designed to eliminate human error in driving vehicles of all types. ADAS systems such as Radars use advanced technologies to assist the driver while driving and thus improve their performance. Radar uses a combination of sensor technologies to perceive the world around the vehicle and then provide information to the driver or take safety action when necessary. Conventional radars based on the emission of electromagnetic and ultrasonic waves have been consumed in the face of the challenges of the constraints of modern autonomous driving, and have not been generalized on all roads. For this reason, we studied the design and construction of a computer vision radar to reproduce human behavior, with a road line lane detection approach based on the histogram of the grayscale image that gives good estimates in real-time, and make a comparison of this method with other computer vision methods performed in the literature: Hough, RANSAC, and Radon.

1 Introduction

Advanced driver assistance systems (ADAS) are multiplying with the emergence of new technologies and the fall in the price of sensors and computers. The ADAS systems not only act on the vehicle in the event of an emergency but also make it easier to drive the vehicle by delegating certain tasks to the vehicle. ADAS are electronic systems that have access to the restitution, traction, braking, and steering components of the vehicle, thus allowing drivers to benefit from assistance and/or to temporarily delegate driving to an automatic co-pilot under certain conditions of traffic [1].

In the literature on driver assistance systems, there are two types: informative systems and active systems. Informative ADAS: Anti-collision system, LDW (Lane Danger Warning), BSD (Blind Spot Detection system), Park assist, and DMS (Driver Monitoring System). Active ADAS: EBA (Emergency Brake Assist), AEB (Automatic Emergency Braking), VSL (Variable Speed Limiter), VCR (Variable Cruise Regulator), ACC (Adaptive Cruise Control), LKA (Lane Keeping Assist), and LPA (Lane Positioning Assist) [2].

The Lane Danger Warning (LDW) warns the driver in the event of involuntary lane crossing. The device uses for this an infrared sensor or a camera pointed toward the ground in front of the car. The sensor locates the white lines to the left and right of the vehicle and calculates the relative position of the vehicle in relation to these lines. As soon as the car (bites) the line, an alert is triggered. For this, we were interested in developing this device to create a computer vision radar to take the raw images and then processed, them so that an automatic driving system could be built [3].

Images are important information vectors that represent a considerable amount of data. Their analysis makes it possible to make the decision in terms of space management and locate the areas of interest and carry out processing in real-time [4].

Computer vision tools enable rapid and automatic extraction of qualitative and semantic information, performing relevant information extraction and intelligent interpretation operations on images [5]. For this purpose, the new radars for industrial use or onboard modern cars migrate to the use of computer vision technology thanks to its reliability and cheaper than other types of radars, since they are based on cameras.

In this paper, we present a histogram method processing image data provided by the camera to detect line lanes and compared it with different methods (HOUGH, RANSAC, and RADON) [6], [7]. This paper is organized into 4 sections: In section 1, we present the general processes to detect line lanes. In section 2 we illustrate how

*Corresponding Author: Hassan FACOITI, Hassan.facoiti@gmail.com

https://dx.doi.org/10.25046/aj070407

to find lanes from the track with the histogram method. In section 3, we present the detection of line lanes by the three methods indicated and their obtained results. In section 4, we compare data from experimental results to the presented methods on videos obtained under real conditions.

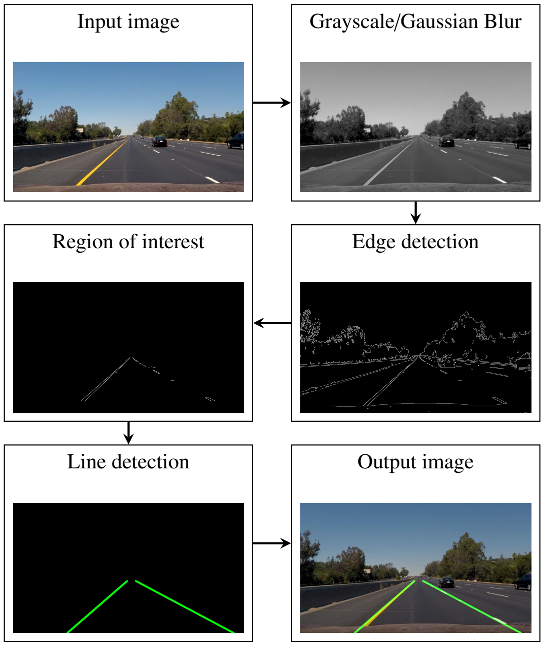


Figure 1: Diagram of the roadway detection & estimation method

2 General approach to the extraction of road marking

In the Figure 1 we illustrate the step-by-step the image processing methodology of the roadway detection and estimation method that achieves the objectives proposed in this paper. This processing is programmed by hybrid languages C++, python-OpenCV, and Matlab.

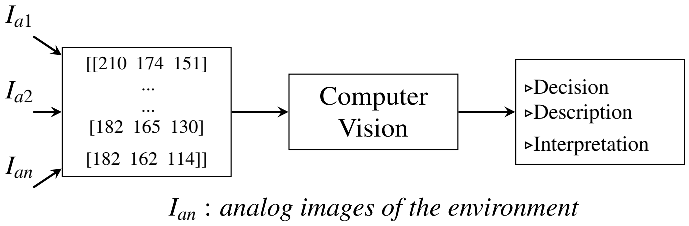


Figure 2: Block diagram of a computer vision system

2.1 Computer vision

The development of computing machines makes it possible to construct a vision system for visual perception capable of naming the objects that surround it, in order to provide the necessary knowledge without ambiguity.

The computer vision system Figure 2 is based on the processing of a sequential sequence of digital images (in the form of pixels) taken by a camera (raw information). Each pixel of the digital image represent an information giving an indication of the amount of light and color coming from the surrounding space [8].

2.2 Frame per second

The general challenge of vision applications is costly in computational time due to the complexity of the algorithms developed. For this reason, the first step to reduce the calculation time is to reduce the processed data depending on time, for this, we have adapted our image capture program to take one image (frame) per second without affecting the data.

2.3 Perspective transformation (Bird-Eye-View)

The Bird-Eye-View transformation technique is to generate a top view perspective of an image, as illustrated in Figure 3. This technique can be classified in digital image processing as a geometric image modification. The Bird-Eye-View transformation can be divided into three stages. First, represented the image in a system of offset coordinates, second rotate the image, third project the image on a two-dimensional plane. The basic diagram of the transformation is given in Figure 4.

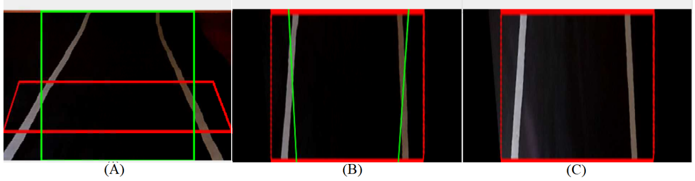


Figure 3: (A) Image originale, (B) Bird-Eye-View image, (C) Output image

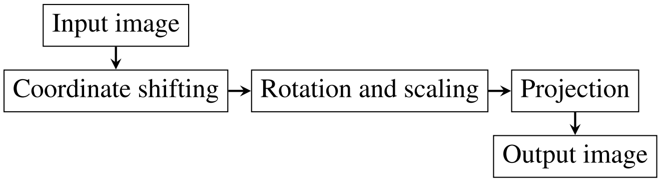


Figure 4: Bird-Eye-View block diagram

2.4 Edge detection

Edge detection is based on Canny Filter, it is a computational approach to detection of marking contours, which represent a fast approach but sensitive to noise because it is accentuated by derivation [9], [10].

After performing a Gaussian smoothing on the image, in order to remove the impurities, then we apply the Canny filter, which calculates the norm of the intensity gradient and the angle of the normal to the gradient (direction of the contours) for each pixel of the smoothed image.

The formulas to apply for each pixel are described as follows:

$$G_r(x,y) = \sqrt{G_x^2(x,y) + G_y^2(x,y)} \quad (1)$$

$$\theta = \pm \arctan(\frac{G_x}{G_y}) \quad (2)$$

$G_r(x,y)$ represent the function of the intensities of pixels, in all directions, x and y.
G_x : represents the gradients in x.
G_y : represents the gradients in y.
G_x et G_y are two convolution masks defined previously, one of dimension 3 x 1 and the other 1 x 3:

$$G_x = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \quad G_y = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad (3)$$

Low gradient	Strong gradient
[5 5 14 14]	[0 0 255 255]
[5 5 14 14]	[0 0 255 255]
[5 5 14 14]	[0 0 255 255]
(a)	(b)

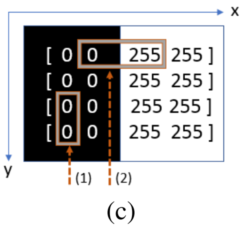


Figure 5: (a) Low gradient. – (b) Strong gradient. – (c) Corresponding image of strong gradient (b). – (1) No change, derivative = 0. – (2) large change, large derivative (gradient).

Canny's method is based on the derivation variation of $G_r(x,y)$ to determine the image contours, it draws the edge with a big change of intensity (large gradient) and is displayed in the form of white pixel outlines Figures 5 and 6.

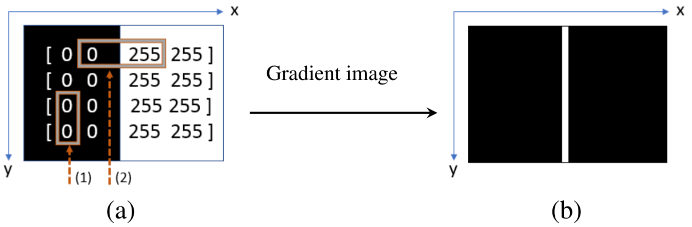


Figure 6: (a) Image. – (1) No change, derivative = 0. – (2) large change, large derivative (gradient). – (b) Image contours in white pixels.

The completely black areas Figure 6(b) correspond to a small variation of intensity between the adjacent pixels. While the white line represents a region of the image where the intensity varies considerably and exceeds the threshold. The threshold is to refine the filtering of the weak contours and to keep only the significant contours, by using two thresholds: High (S_H) and Low (S_L). If the value of a contour is higher than the highest threshold, it is preserved. And if the threshold is lower than the low threshold, the corresponding pixel becomes black.

At this location, two transformations are made of the image obtained after perspective transformation Figure 7(A) to detect the edges, the first is a threshold operation Figure 7(B) and the second is the Canny method Figure7(C). At the end, the two resulting images are confused to guarantee the detection of the edges of the image Figure 7.

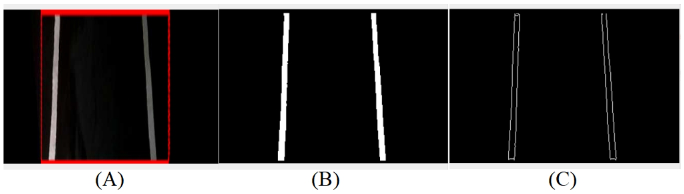


Figure 7: (A) perspective Image, (B) Grayscale image, (C) Canny image

2.5 Region of interest

Automatic detection of the region of interest on an image often has multiple areas is a difficult problem to solve. The proposed method Figure 8 is to determine the region of interest is based on the creation of a digital mask of the image, this mask consists of two types of pixels: 0 represent the black and 255 represent the white, a triangle (or rectangle) pixels of values 255 identical to the region of interest that we are trying to extract and the rest of the pixels have the value 0 Figure 8(b). The goal is to do the and-logic operation between the bit of each pixel homologous to the image and this mask, thus masking the complex image to show only the region of interest corresponding to the mask Figure 8(c).

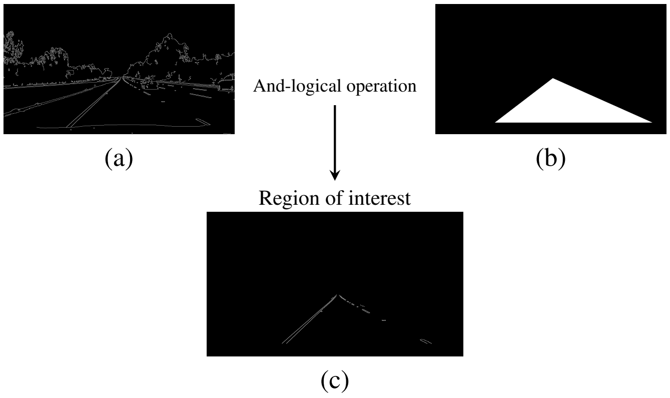


Figure 8: (a) Canny image. – (b) Image mask. – (c) Region of interest.

As the pretreatment steps mentioned and detailed in [8], a region of interest (the red rectangle) is created as illustrated in Figure 9.

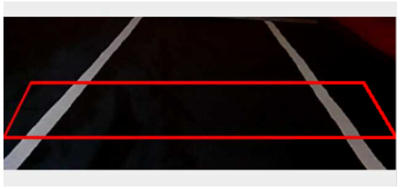


Figure 9: Region of interest

3 Find Lanes from Track using the histogram method

3.1 *Histogram method*

The histogram method makes it possible to find the exact position of the road line lane from the grayscale image compared to the position of the radar system. This method is based to create a region of interest forms a rectangle in the final image Figures 10(1) and 11(C), divide into different strips.

Algorithm 1: Histogram algorithm

```
Input : FrameFinal;
initialization : DynamicAreas.size();
for int i=0; i < FrameFinal.size().widht ;i++ do
    RegionOfInterestLane=FrameFinal(rectengle.size());
    divide(255, RegionOfInterestLane,
      RegionOfInterestLane);
    push all the intensity valueus in DynamicAreas;
    histogrameLane.push_back(int)(
      sum(RegionOfInterestLane)[0] );
end
LeftLaneParametre= max_element(histogrameLane.begin(),
  histogrameLane.lenght()/2);
LeftLanePosition=
  distance(histogrameLane.begin(),LeftLaneParametre);
RightLaneParametre=
  max_element(histogrameLane.lenght()/2,
  histogrameLane.end);
RightLanePosition=
  distance(histogrameLane.begin(),RightLaneParametre);
Output : Final_image, LineLanePosition;
```

In one step, there will be total pixels and each pixel has an intensity of 0 if is black or 255 if is white, we can calculate multiply each step of pixels by its intensity and get the resultant intensity and store this intensity in dynamic areas Figure 10(2), then replace each intensity equal to 255 by X Figure 10(3)(4) and save the coordinates value to find the line lane position. Finally, draw these lines in green and their average on the final image Figure 11(C).

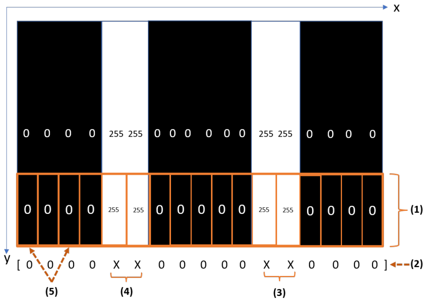


Figure 10: (1) Region of interest, (2) Dynamic areas, (3) Right lane position, (4) Left lane position , (5) Pixels strips

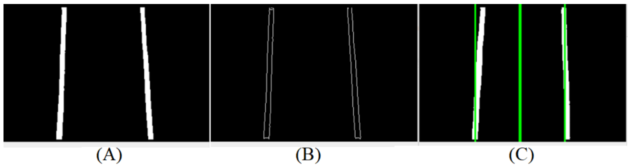


Figure 11: (A) Grayscale image, (B) Canny image, (C) Final image

3.2 *Calibrage*

We take the average of the lines detected as a reference, it is the instruction of our automatic closed-loop system which must be followed by the car, it is the principle of autonomous driving. Each time the car deviates from this average (the deviation is indicated by a blue line), the system asks to perform the calibration Figures 12, 13 and 14.

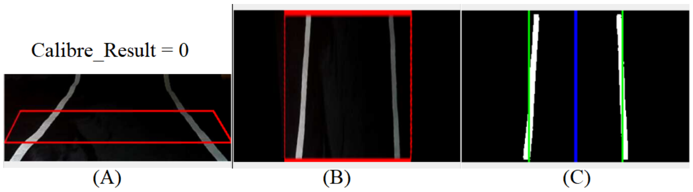


Figure 12: The right position: (A) Original image, (B) Bird eye viw image, (C) Result Image

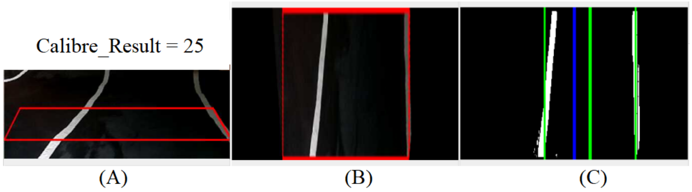


Figure 13: Deviation to the left: (A) Original image, (B) Bird eye viw image, (C) Result Image

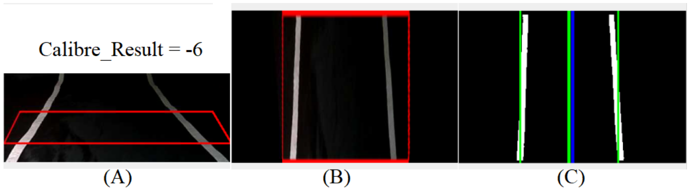
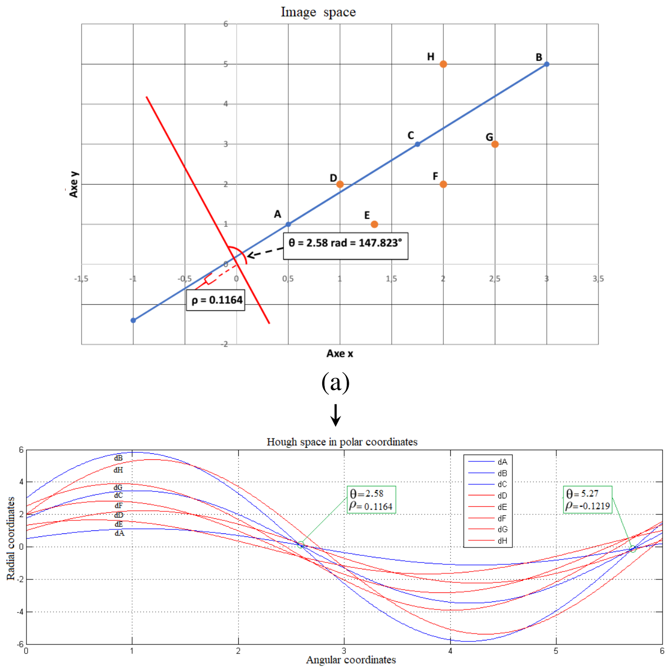


Figure 14: deviation to the right: (A) Original image, (B) Bird eye viw image, (C) Result Image

4 Computer vision methods find lanes from the track

4.1 *HOUGH Transform*

The Hough Transform is an efficient form in recognition tool, the practical application to detect in a camera image the presence of parametric curves from a set of characteristic points, essentially uses the spatial information of the characteristic points, that is, their position in the image. The generalized Hough transform can detect other forms [7].



(a)

(b)

Figure 15: Principle of Polar System. – (a) Distribution of noisy points. – (b) Hough Transform

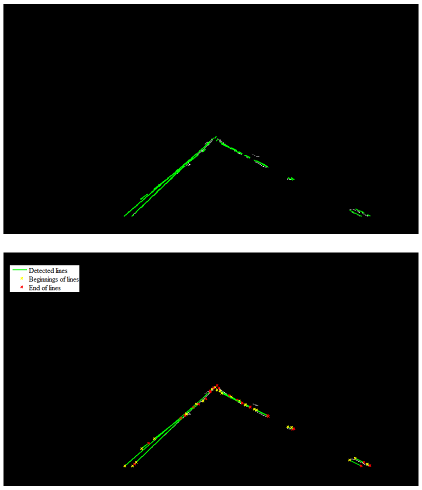


Figure 16: HT Detected lines lanes

The Hough transform algorithm based on the polar system uses an accumulator matrix that represents the space (ρ,θ) [11]–[14], of dimensions (L,C) where L is the number of possible values of ρ and C the number of values of θ. In a distribution of noisy points (x_i, y_j) of the Figure 15(a), each point becomes a sinusoid of the equation "(4)" in the parameter space (ρ,θ). The Figure 15(b) shows, at the end of the accumulation, the sinusoids corresponding to the points of the same line intersect at the point (ρ,θ) setting this line.

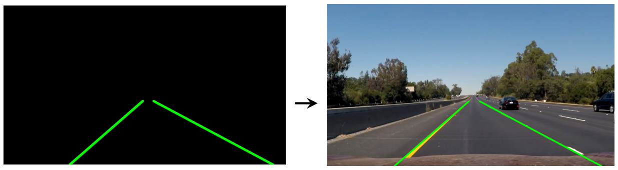
$$x\cos(\theta) + y\sin(\theta) = \rho \tag{4}$$

The Figure 16 shows the execution of this algorithm on the image of the regions of interest Figure8(c), so that the accumulator correspond to the votes, which obtained the highest values of the parameters correspond to the lines of the contours of the road way in the treated image.

4.2 Optimization

The detection of roadside contours by the Hough transform is robust as well as the information of the positions of the lines of these contours, this information must be optimized, so that to give a single line in each limit of the road lane Figure 17(a), the objective is to invest in the engineering of Advanced Driver Assistance Systems (ADAS) to build an embedded control system of the vehicle position. The optimization algorithm is summarized in "Algorithm. 2".

Finally, the experimental results as they appear in the output image Figure 17(b) show that the detection algorithm followed in this paper improves visibility of the roadway and reduces noise.



(a)

(b)

Figure 17: (a) Optimization of line detection. – (b) Output image

Algorithm 2: Optimization algorithm

```
Slope_interception_left = [ ];
Slope_interception_Right = [ ];
for line in linesDET do
    % linesDET: the lines detected by the transform;
    % each linesDET detected is a 2D array containing the
      coordinates in [[x1, y1, x2, y2]];
    Convert 2D coordinates of linesDET to 1D;
    % 2D [[x1, y1, x2, y2]] ⇒ 1D [x1, y1, x2, y2];
    Calculate the slope $a_i$ and the interception $b_i$ of each
      linesDET;
    if slope $a_i < 0$ then
        % note that the y axis is reversed. Values increase in
          descending;
        Slope_interception_left = [$a_i$,$b_i$];
    else
        Slope_interception_Right = [$a_i$,$b_i$];
    end
end
Left_line_parameter = average.(Slope_interception_left);
Right_line_parameter = average.(Slope_interception_Right);
```

4.3 Probabilistic voting method RANSAC

The RANSAC (Random Sample Consensus) method is a probabilistic voting method based on the use of minimal data to accurately estimate the parameters of a model, even if the data is noisy. This method has been proposed to reduce the calculation time of conventional voting methods such as the Hough Transform. The algorithm of this method is based on a number of iterations, at each turn, it randomly selects a subset of data (two random points), then finds the model for the selected data, then tests all the data by model and determine the relevant points according to the threshold. In the end, if the new model is better than the best model (based on the number of relevant points), then the new model becomes a best model. From these processes, the estimate preserves the parameters of the searched lines. Let ”p” be the probability of obtaining a good sample, ”s” the minimum number of points in a sample to estimate the parameters of the model and ”r” the probability to have a valid point in all selected contour points. The minimum number ”m” of random draws necessary to have a probability of the correct parameters is as follows:

$$m = \frac{log(1-p)}{log(1-(1-r)^s)} \quad (5)$$

Finally, in Figure 18 the vehicle circulation lane is determined by exploiting the RANSAC algorithm after 12 iterations, the execution of this algorithm on the same image returns different results because it is based on the random selection of the data, and all the results obtained are valid. The choice of 12 iterations amounts to minimize the calculation time in parallel to obtain a good estimate. After optimizing the results with the optimization algorithm ”Algorithm. 2”, we obtain a single line in each path limit which is clear in Figure 18(c).

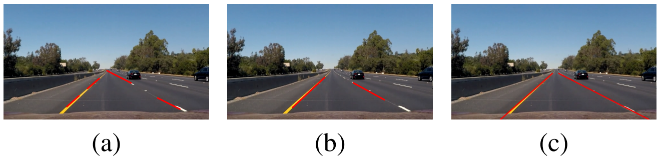

(a)
(b)
(c)

Figure 18: (a) and (b):Track markings detected by our RANSAC algorithm after 12 iterations. (c):Grouping of line segments.

4.4 Radon Transform

The Radon transform is a mathematical technique developed by Johann Radon, it is defined on a space of lines L in R^2 according to two arguments (ρ, θ). The application of this transform on a two-dimensional function I(x,y) (image) is based on several projections of the image under parallel beams under different given angles, in order to calculate the line integrals of this beam in directions specified. The resulting image R(ρ, θ) of the projection is sum of intensities of the pixels in each direction, can be translated by [15]–[17]:

$$R(\rho, \theta) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} I(x,y)\delta(x\cos(\theta) + y\sin(\theta) - \rho)\,dxdy \quad (6)$$

The robustness of the Radon transform lies in its ability to detect lines in noisy images. Like the Hough transform, it can also transform lines into easy-to-solve pics that correspond the parameters of the lines in the outlines of the image. This property of the Radon transform allows us to identify and detect the contour lines of the roadway in the image of the region of interest. After the optimization, we find very important and valid results Figures 19 and 20.

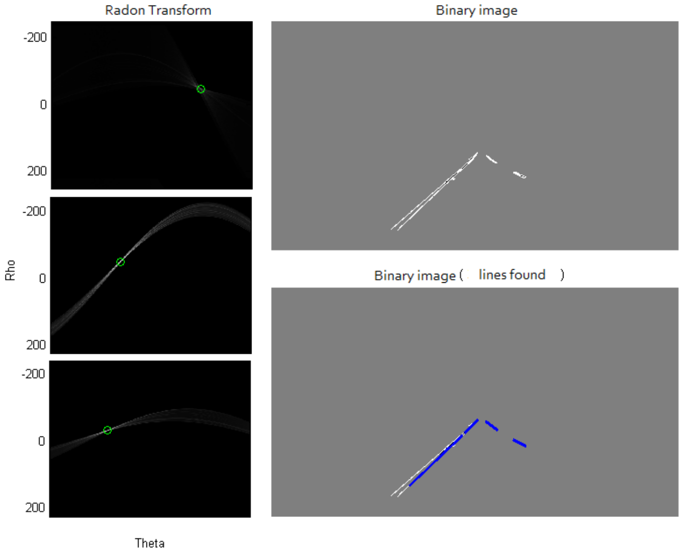


Figure 19: Radon algorithm lane detection

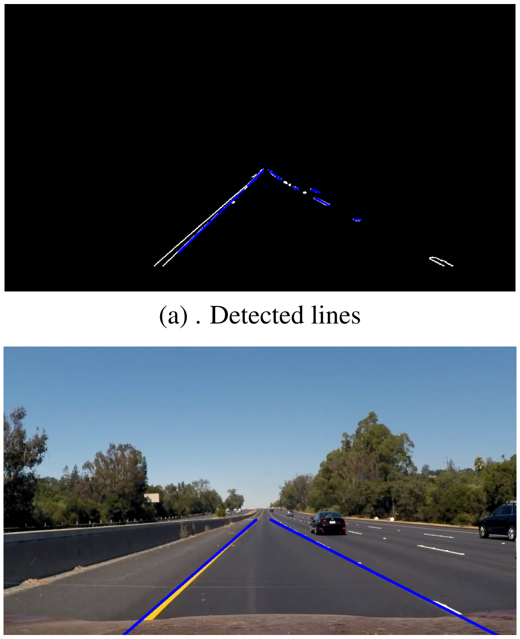
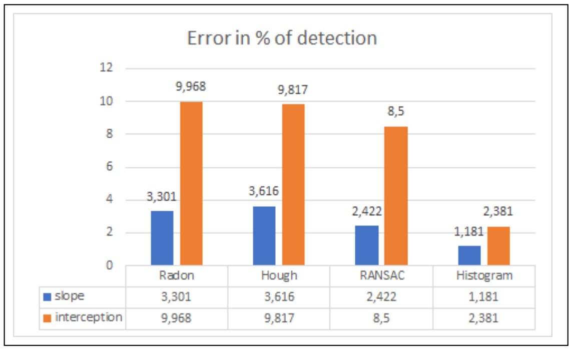

(a) . Detected lines
(b) . Grouping of lines

Figure 20: Radon algorithm results

5 Experimental data

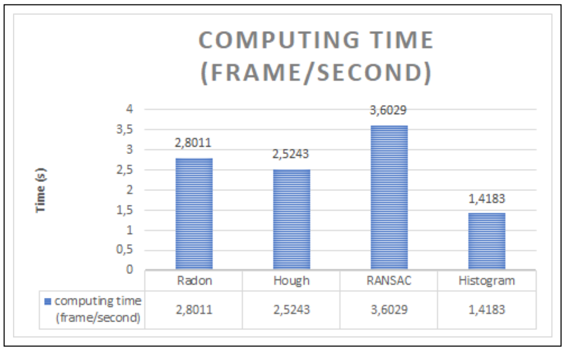
We have based on the error (%) (false alarms) of the estimate of the slope and the interception of the detected lines lanes to prove the robustness of the algorithms presented and compare them, by using several tests on sequences of images (videos) of resolution (1280 x 720) in different conditions. Each image sequence has an almost constant width of the line lanes and also has real parameters (slope, interception). The false alarm rate determined after a series of experiments varies acceptably among the four methods compared to the real values Figure 21.



	Radon	Hough	RANSAC	Histogram
■ slope	3,301	3,616	2,422	1,181
■ interception	9,968	9,817	8,5	2,381

Figure 21: Detection error percentages for each of the four methods

Note, however, that the method based on the histogram presents a minimum error and minimum computing time compared Figure 22 to RANSAC, HOUGH, and RADON methods, due to the data processing being done at the rectangle of the region of interest after generating a perspective view from the above an image.



	Radon	Hough	RANSAC	Histogram
■ computing time (frame/second)	2,8011	2,5243	3,6029	1,4183

Figure 22: Computing time (frame/second)

6 Conclusion

The comparative study has shown that the proposed vision algorithms (Histogram, Hough, RANSAC, and RADON) are robust to detected contour points with a predefined model describing the curvature of the road, and are quite efficient and have a very efficient accuracy rate despite the fact that noise is present in the images of the road. But the Histogram method presents more efficiency compared to other methods.

The use of these methods is necessary to treat the real environment, but in terms of execution time is very important, which poses a problem of synchronization with the real time of a sequence of images.

In order to complete our study, we want to continue to improve vision algorithms able to be generalized to all roads and detect obstacles and remain robust to different conditions and compatible with real-time processing.

Conflict of Interest The authors declare no conflict of interest.

References

[1] K. Bengler, K. Dietmayer, B. Farber, M. Maurer, C. Stiller and H. Winner, ”Three Decades of Driver Assistance Systems: Review and Future Perspectives,” in IEEE Intelligent Transportation Systems Magazine,**6**(4), 6-22, winter 2014, doi: 10.1109/MITS.2014.2336271.

[2] M. Nolwenn. ”From driving assistance systems to connected autonomous vehicles”. 2019. Doctoral thesis in Automation, Production, Signal and Image, Cognitive Engineering. Bordeaux.

[3] R.N. Mahajan and Rashmi A. M. Patil, ”Lane departure warning system”. International Journal of Engineering and Technical Research, **3**(1), 120-123, 2015.

[4] C.J. Jacobus, D. Haanpaa.:All weather autonomously driven vehicles. U.S. Patent **9**, 989-967. Jun 2018.

[5] M. Olivier, Y. Poncet. ”Computer vision techniques applied to aerial images of floodplains”.2001.

[6] M.A. Rahman, M. F. I. Amin and M. Hamada, ”Edge Detection Technique by Histogram Processing with Canny Edge Detector,” 2020 3rd IEEE International Conference on Knowledge Innovation and Invention (ICKII), 128-131, 2020, doi: 10.1109/ICKII50300.2020.9318922.

[7] M. Henri.:A panorama of the Hough transformation. Signal processing. **2**(4), pp 305-317, 1985.

[8] H. Facoiti, A. Boumezzough and S. Safi. ”Comparative study between computer vision methods for the estimation and detection of the roadway”, 2021 7th International Conference on Optimization and Applications (ICOA), 1-6, 2021. doi: 10.1109/ICOA51614.2021.9442633

[9] B. BESBES, Christele LECOMTE, Peggy SUBIRATS.:New lane line detection algorithm. XXIIe colloque GRETSI, Signal and image processing, September 2009.

[10] C. John, A computational approach to edge detection. Readings in computer vision. Morgan Kaufmann,184-203, 1987.

[11] R.O. Duda, E. Hart, Use of the Hough transformation to detect lines and curves in pictures. No. SRI-TN-36, Communication of the ACM, 15, 11-15, janvier 1972.

[12] H. Facoiti, S.Safi, A.Boumezzough.:Analyse, caractérisation et conception d’un système radar anticollision embarqué dans les véhicules. sciencesconf.org:icsat-2020:326086, 13-18, International Colloquium on Signal, Automatic control and Telecommunications, Caen, France, juin 2020.

[13] S. Haykin.: Neural Networks and Learning Machines. 3rd Edition, New York 2008.

[14] C. John.:A computational approach to edge detection. Readings in computer vision. 184-203, 1987.

[15] J.X. Wu, S.V. Kucheryavskiy, Linda G.Jensen, Thomas Rades, Anette Müllertz, Jukka Rantanen.:Image Analytical Approach for Needle-Shaped Crystal Counting and Length Estimation. Crystal Growth and Design, Bind 15, s.4876-4885 2015.

[16] H. Wang, Q. Chen.:Real-time lane detection in various conditions and night cases. Intelligent Transportation Systems Conference, 1226-1231, Canada, September 2006.

[17] K. Osman, J. Ghommam, M. Saad.:Vision Based Lane Reference Detection and Tracking Control of an Automated Guided Vehicle. 25th Mediterranean Conference on Control and Automation (MED), Valletta, Malta, July 2017.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 7, No. 4, 49-58 (2022)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

ARAIG and *Minecraft*: A Modified Simulation Tool

Cassandra Frances Laffan[*,1], Robert Viktor Kozin[1], James Elliott Coleshill[1], Alexander Ferworn[1], Michael Stanfield[2], Brodie Stanfield[2]

[1] *Computational Public Safety Lab, Department of Computer Science, Toronto Metropolitan University, Toronto, M5B 1Z4, Canada*

[2] *Inventing Future Technologies Inc. (IFTech), Whitby, L1N 4W2, Canada*

ARTICLE INFO

Article history:
Received: 10 May, 2022
Accepted: 19 July, 2022
Online: 27 July, 2022

Keywords:
Haptics
Octree
Pathfinding
*A**
Search and Rescue
Digital Games

ABSTRACT

Various interruptions to the daily lives of researchers have necessitated the usage of simulations in projects which may not have initially relied on anything other than physical inquiry and experiments. The programs and algorithms introduced in this paper, which is an extended version of research initially published in ARAIG And Minecraft: A COVID-19 Workaround, create an optimized search space and egress path to the initial starting point of a user's route using a modification ("mod") of the digital game Minecraft. We initially utilize two approaches for creating a search space with which to find edges in the resulting graph of the user's movement: a naive approach with the time complexity of $O(n^2)$ and an octree approach, with the time complexity of $O(nlogn)$. We introduce a basic A algorithm to search through the resulting graph for the most efficient egress path. We then integrate our mod with the visualization tool for the "As Real As It Gets" (ARAIG) haptic suit, which provides a visual representation of the physical feedback the user would receive if he were to wear it. We finish this paper by asking a group of four users to test this program and their feedback is collected.*

1 Introduction

As most people, readers and the general populace alike, are aware, the COVID-19 pandemic has hampered the plans of many researchers [1]; moreover, ongoing supply chain interruptions have compounded this issue. From research using specialized equipment to field testing, the past two and a half years have been full of ingenious "workarounds" to both workplace restrictions and material shortages. We feel our current project is no exception: in a time where our research, as introduced in our previous work [2], should be applied to tangible hardware and tested in physical environments, access to our lab and materials is limited. The need for corporeal results is clear and thus, we propose and implement a workaround in this paper. This work is an extended version of previously published conference proceedings: *ARAIG and Minecraft: A COVID-19 Workaround* which may be found here [3].

As touched upon in [2], the field of search and rescue is going to experience various changes over the coming decades. Canada, home to both this project and its researchers, is one of many countries currently feeling the metaphorical and literal heat of climate change's effects [4]. Natural disasters, such as the forest fires plaguing the Canadian prairie provinces [5, 6], will become more prevalent across the country and globally before climate action takes effect [4]. Thus, now is the time to act in terms of curbing future disasters while empowering emergency workers to safely and efficiently respond to high impact, low (though increasing) frequency events. While we cannot influence government actions, domestically or globally, we can certainly do our best to assist in preparing our first responders for future disasters.

One of the most dangerous situations a firefighter may face in disaster environments, particularly in enclosed spaces such as homes and buildings, is potentially growing disoriented and thus, lost [7]. In fact, there is ongoing research on this exact issue [7], as this is still an unsolved problem. The motivation for this research, in conjunction with the increasing incidence of natural disasters, is this issue: how can we utilize modern technology in a lightweight fashion to assist firefighters in navigating out of these low visibility environments? Due to the restrictions the pandemic has brought about, as well as ongoing supply chain issues, the initial answer to this question comes in the form of simulation.

Minecraft is a digital game which focuses on the exploration, and building, of randomly generated environments. The graphics, game mechanics and game world are simple: the player is placed, without warning or preamble, into the *Minecraft* environ-

[*] Corresponding Author: Cassandra Frances Laffan, George Vari Engineering and Computing Centre, 245 Church Street, +1 (647) 983-4070 & Cassandra.Laffan@ryerson.ca

https://dx.doi.org/10.25046/aj070408

ment, which is composed of “blocks”, not dissimilar to voxels. The game is centred around navigating, manipulating and accumulating blocks while exploring a procedurally generated world. Since the world of *Minecraft* is so simple, and there are numerous frameworks which support modding the game (these programs will be referred to as “mods”), we have found it suits the above requirements for a simulation medium given our problem domain. Moreover, the game has a well established and active “modding” community [8, 9].

2 Related Work

There has been considerable research into simulations over the past few years, particularly for reasons outlined in this paper’s introduction. It is of note, however, that the usage of digital games as an avenue for simulating experiments predates the pandemic and supply shortages. “Project Malmo”, a project published in 2016 by Microsoft, is an earlier example of this. The authors write an Application Programming Interface (API) and abstraction layer for *Minecraft*. With it, they train an artificial general intelligence (AGI), to complete various tasks [10, 11]. The inclusion of [10] in our *Related Work* is due to their motivations for using *Minecraft* as their training medium:

- *The environment is rich and complex, with diverse, interacting and richly structured objects [10, 11].* The platform must offer a worldspace with which a user, or otherwise autonomous agent, can interact. This worldspace must be varied and robust.
- *The environment is dynamic and open [10, 11].* The platform needs to offer unique settings which allow us to mimic real-world environments.
- *Other agents impact performance [10, 11].* Other agents, such as AI or other humans, should be able to impact the simulation.
- *Openness [10, 11].* The platform should be cross-platform and portable.

Our ongoing research does not necessitate training AI in *Minecraft*. However, the above guidelines summarize why we believe the game is an optimal platform on which we can test our algorithms and simulate our experiments. The third point in the above list, touching upon how other agents impact performance, is discussed again in various sections of this paper.

Research into simulation, space division and construction of point clouds, graphs and pathways is limited, especially in the realm of search and rescue. Thus, we explore a more generalized approach to 3D pathfinding which has a lower time complexity than more conventional approaches to navigating Euclidean space. In that regard, the authors in [12] explore pathfinding using 3D voxel space in the digital game *Warframe*.

In [12], they propose utilizing an octree, comprised of voxels or “octants”, to split the 3D world of *Warframe* into a low time complexity, searchable space which allows for more efficient pathfinding. Every octant is the centre point of a corresponding voxel. The program explores all 26 nearby voxels to find the next best space to move to. The authors determine the next available voxels via the following constraints: are there obstacles in the closest voxel? If so, the voxel is left out of the potential path as the agent cannot occupy the same space as another object. Is the voxel out in the “open”, away from cover? It is undesirable for an agent to be out in the open, as it leaves it vulnerable to enemy attacks.

In [13], the authors implement an approach to processing point clouds in Euclidean space, much like what we are attempting to accomplish in our own research; in this case, they wish to navigate through buildings and other structures. The researchers observe that sorting and naively navigating through unordered point clouds can have needlessly high time and space complexities. They propose using an octree to circumvent these issues, allowing new buildings to be mapped internally with a lower demand for computation time. It should be noted that the point clouds these researchers are using are already constructed before data processing; our research differs in that, not only are we constructing the point clouds as a user navigates through the worldspace, but points are not pre-processed.

The authors in [13] use these point clouds to determine the location of obstacles and throughways, such as furniture or doors, respectively. They concern themselves more with identifying and labelling specific objects and terrains, such as stairs, whereas we are more concerned with efficient navigation. More precisely, our focus is on whether or not the firefighter’s elevation has changed, as well as the most efficient egress path through a given point cloud and resulting graph.

Further on in this paper, we implement an A* algorithm for searching the resulting graph with appropriate edges drawn between nodes. Two papers which are referenced in this publication for the implementation of A* are: *Algorithms and Theory of Computation Handbook* [14] and *Artificial Intelligence: A Modern Approach* [15]. Both publications act as guidelines for the implementation of not only A*, but other algorithms which are discussed in the *Future Work* section. The latter also provides guidance for both the time and space complexities of A*, which are necessary to explore given the problem domain.

3 Methodologies

The next sections are divided and ordered in a way mimicking the timeline of this project. First, methods for modifying *Minecraft* are outlined, as this is a necessary step for producing meaningful datasets and egress settings for our experiments. Next, we explore dividing the worldspace of *Minecraft* and ensuing datasets into a searchable graph with edges (at times referred to in this work as “adjacencies”). Splitting the worldspace and creating a graph in a timely manner emulates the urgency necessary in the real world environment for firefighters. Pathfinding through the graph follows the graph creation, as is the logical progression of events; again, efficiency in time complexity is discussed since urgency is one of the most important factors in creating the egress path for a first responder. Navigating through the *Minecraft* worldspace follows, alongside using the ARAIG visualization tool. Creating directions and output for the suit, while important, needs to be done with the intention of making them intuitive for first time users. Thus, a short survey is conducted with a small group of users to get initial feedback on the simulated system.

4 Modding *Minecraft*

Minecraft follows a server-client network model. This means there are two avenues for modding the game: server-side or client-side [16]. The server-side handles logic, game state and updates from clients. The client-side handles rendering the game state and sending updates. This separation of concerns is important. If we wish to illustrate this concept in *Minecraft*, a creature, its location and where it moves is handled by the server. The information about said creature is sent to the client where it is rendered. The client can also send updates, for example, when the player wishes to perform an action such as jumping or hitting a block. In this respect, as it can only modify what it has access to, the client cannot control where the aforementioned creature is and the server cannot control how the aforementioned creature appears.

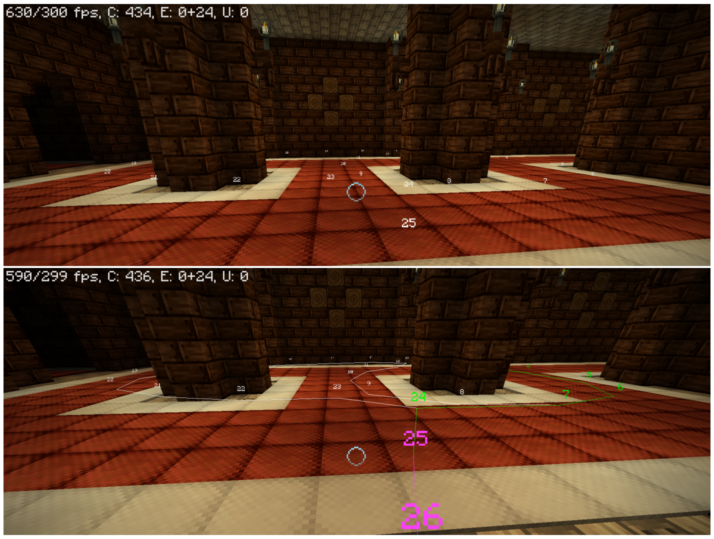
Server-side modding tends to be relatively straightforward, since the network interface is well understood and documented. There are numerous "community sourced" implementations and application programming interfaces (APIs) which extend the *Mincecraft* server. The API we use for this mod is Spigot [17], which provides a way to run code on events, such as player movement. It also enables programs to react to said events, namely cancelling the movement, recording coordinates or even running actions, such as smiting the player with a bolt of lightning.

Client-side modding is often more complex than its server-side counterpart. To extend the game, client-side mods have to "hook" directly into the "vanilla" client using a variety of complex methods. An example is runtime "bytecode" manipulation, where the Java runtime environment is used to modify compiled code while it is running [9]. *Minecraft* client code is often complex in nature as it deals largely with in-game rendering. In addition, the code is obfuscated, making it difficult to read and understand.

There are two big projects that make modding *Minecraft* easier: Forge [18] and Fabric [8]. Forge is more established and has a larger scope of supported modding functionalities. However, the consequence of this is that the framework takes longer to update and is less light-weight. Fabric, in contrast, is newer and lighter, using modern techniques and providing more low level control. In this project, we opt to use Fabric, as we prefer to keep the mod lightweight, allowing for a more agile approach to mod development.

Originally, we attempted to create a server-side mod. Our goal was determining if location data and its visualization is viable in *Minecraft*. Despite our success in proving its viability, the limitations of server-side modding in regards to visualization quickly made themselves known. As previously mentioned, the server has limited control as to what the client "sees". While the server can spawn creatures or create particles for visualization, we require more customization. Given these circumstances and our requirements, the client-side model is best suited for our mod. This allows us to use the same code to render our custom visualizations that the client uses to render the game.

Our *Minecraft* mod visualizes a graph data structure by drawing the edges as lines and the nodes as numbers in the *Minecraft* worldspace. This gives the user the ability to visualize the recorded coordinate points and the connections created between them. Coordinate points closer to the real world are now easier to collect; there is no need for random numbers or predetermined coordinates.


630/300 fps, C: 434, E: 0+24, U: 0

590/299 fps, C: 436, E: 0+24, U: 0

Figure 1: An example of how the *Minecraft* mod visualization tool renders the player's pathway before and after the pathfinding algorithm is run.

As stated above, *Minecraft* is a Java based game. However, in order to seamlessly interface with the ARAIG simulation software, the pathfinding program is in C. Thus, the *Minecraft* mod and the pathfinding algorithms must be split into separate programs, with communication being done over a network socket using a simple protocol. This protocol can be described as a request-response byte encoded message protocol. The first byte is the message type and the rest is the message body in the request. In contrast, the reply is comprised of only a message body. For example, to get the current location from the mod, the pathfinding program sends a $GET\_LOCATION$ byte identifier. The server then responds with 4 big-endian byte encoded floating point numbers containing to the x, y, z, and yaw components respectfully.

The control flow of the mod is as follows:

1. **Capture location data:** The mod has two commands, */start* and */stop*, the user can enter in the "chat bar" to control the recording of coordinate points. This emulates a GPS device by providing the coordinates of a wearer to an external machine. The location data is stored as a list of coordinate points. While the user is recording his path, a numbered node is placed in the worldspace and recorded as a set of coordinates.

2. **Transfer captured location data:** The pathfinding program then requests all of the recorded coordinate points from the mod, allowing processing to begin.

3. **Transfer live location data:** Additionally, the pathfinding program queries the user's live location data from the client. This allows for live tracking of the user along the path so that egress instructions are updated accordingly in real time.

4. **Send draw updates:** Finally, in response to the captured and live location data, the pathfinding program sends draw

commands to the client. The commands include: drawing or deleting a line between points, changing the colour of a given line, and changing the colour of the points.

While our initial belief of having the data collection and visualization as separate from the pathfinding program may result in unnecessary complexity, the resulting mod is, in fact, the opposite. Using two separate programs for data processing and data visualization allows the code to be more organized and decoupled. One of the benefits for this is the pathfinding program could be restarted and debugged without needing to restart the *Minecraft* client. Since the two are decoupled, both can be developed at different paces.

5 Creating the Graph

Two approaches are outlined here for graph creation, which is necessary for finding an egress path given points in 3D space: a naive approach which has a time complexity of $O(n^2)$ and an octree approach which runs in $O(nlogn)$ time.

5.1 Naive Implementation

Algorithm 1 is the initial approach for creating the shortest egress pathway possible given the user's nodes.

Algorithm 1: *add_node*(*graph*, *node*) Adds a node to the graph. Then, adds it to the adjacency arrays of any existing nodes within a radius *R*.

```
Input: A graph with all previous nodes already inserted
       and connected: $graph$, a newly created node to
       insert into the graph: $node$
foreach $existing\_node \in graph.nodes$ do
    if $distance(existing\_node, node) \leq R$ or
     $existing\_node == graph.nodes[-1]$ then
        /* Appends the node to the adjacency
           array of a given node in the graph
           */
        $existing\_node$.adjacencies ← $node$
end
$graph$.nodes ← $node$
```

When a new node is created, the *add_node* function is called to add it to the graph. The node is compared to every existing node already in the graph; if the node is within *R* radius of the node it is being checked against, it is added to the second node's adjacency list. This approach is simple and intuitive, yet inefficient. As mentioned previously, its time complexity is $O(n^2)$. The algorithm works as a proof of concept but it quickly becomes evident that with sufficiently large *n*, a more efficient approach is necessary. See Figure 2 for a visualization of this function.

5.2 Octree Implementation

An octree object has three member variables associated with it. First, the *bounds* variable, which contains the boundaries for the 3-Dimensional (3D) box that the octree resides in; these are stored as a set of coordinates. Next, it contains a *children* array. This array contains either zero or eight octants. These octants are the result of splitting the *bounds* object into eight equally sized boxes.

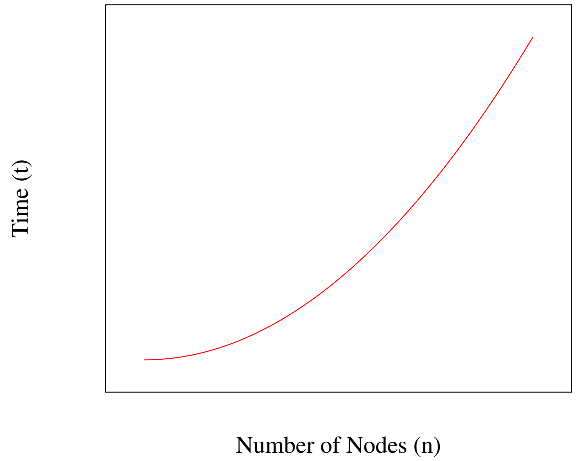


Figure 2: Plotting number of nodes n against time t where $t = n^2$.

Algorithm 2: *make_octree*(*nodes*, *bounds*) Creates an octree given a set of points in 3D space.

```
Input: An array of nodes: $nodes$, the current boundaries for
       the octree: $bounds$
Output: An octree containing either eight octree "children"
        (octants) or $N \geq$ leaf nodes
initialize $octree$
initialize $boundaries$
if $!nodes$ then
    $octree$.children ← NULL
else
    if $length(nodes) \leq N$ then
        $octree$.nodes ← $nodes$
    else
        /* Splitting the space into octants   */
        $boundaries$ ← split($bounds$)
        initialize $octree$.children
        for $i = 0; i \leq 8; i++$ do
            initialize $next\_nodes$
            foreach $node \in nodes$ do
                if $node \in boundaries[i]$ then
                    $next\_nodes$ ← $node$
            end
            $octree.children[i]$ ←
             $make\_octree(next\_nodes, boundaries[i])$
        end
    end
end
return $octree$
```

Finally, an octree object has a *nodes* array; it remains empty unless the octree object is a leaf in the greater data structure. It should be noted that, unlike the naive approach above, this algorithm is run after the subject has ended the path recording. The pseudo-code for the octree creation is illustrated in Algorithm 2.

The initial octree object is created with its bounding box characterized by the minimum and maximum (x, y, z) coordinates of the whole graph. The function *make_octree* is then called on an array of every node and the aforementioned boundaries. Then, one of three things must happen: first, if the array of nodes is empty, the octree's *children* array is set to *NULL* to indicate that it is empty. Second, if the array of nodes exists, and contains less than or equal to the amount of allowed nodes in a given octant space, the octree's *nodes* array is populated and the octree object becomes a "leaf".

Finally, if neither of the previous two conditions are true, the function recursively calls itself. A new set of octants is created by splitting the boundaries into eight identically sized cubes. Next, for each new octant, each node is compared to its bounds and, if it is contained within the octant's boundaries, it is placed in a new array. Once the array is populated with the appropriate nodes, the *make_octree* function is recursively called on it and its respective boundaries. The resulting octree is added to the *children* array for the current octree.

The construction of this octree creates an efficient, though spatially complex, search space which allows for the appropriate edges in the graph to be created. The pseudo-code for edge creation, as a series of octree searches, is in Algorithm 3.

Algorithm 3: $find\_adjacencies(octree, node)$ Populates a node's adjacency array with nodes within a radius R.

```
Input: An octree object: $octree$, the current node that we
       wish to find the adjacencies for: $node$
if $octree.nodes$ then
    foreach $oct\_node \in octree.nodes$ do
        if $node \notin oct\_node.adjacencies$ and
         $distance\_between(oct\_node, node) < R$ then
            $node.adjacencies \leftarrow oct\_node$
    end
else
    for $i = 0; i < 8; i++$ do
        if $node \in octree.children[i].bounds$ then
            $find\_adjacencies(octree.children[i], node)$
    end
end
```

As Algorithm 3 demonstrates, $find\_adjacencies$ takes a fully formed octree and a node as arguments. It is called on every node in the graph once. The function checks if the octree object is a leaf, and, if it is, the selected node is compared to each node in the octree's *nodes* array. If any nodes are within the given distance R, they are added to the adjacency array for the node. Conversely, if the octree object is not a leaf in the octree, the bounds for each member of its *children* array are checked against the coordinates of the node. If the node is within the given child's bounding box, the $find\_adjacencies$ function is called on the child and the node. This is done recursively until all of the appropriate octree branches for a given node are explored.

It should be noted that the *average* case runtime for creating the octree and recreating its path is $O(nlogn)$, where n is the total number of nodes in the system. The worst-case runtime for this octree approach is the same as the above *naive implementation*, which is $O(n^2)$. The worst case for this algorithm occurs if and when a node is compared to every leaf in the octree and thus, every other node in the system.

The bottleneck for efficiency in this case is not the creation of the octree, which is in fact linear time, nor is it any given single search of the octree, which is also $O(logn)$. Rather, the bottleneck is that the search must be performed for each node in the graph. Thus, the time complexity for this algorithm is $T = n + nlogn$, which is the sum of both the runtime of the octree construction and the creation of each node's adjacency list. Idiomatically, the runtime is $O(nlogn)$. See Figure 3 for a visualization of this function.

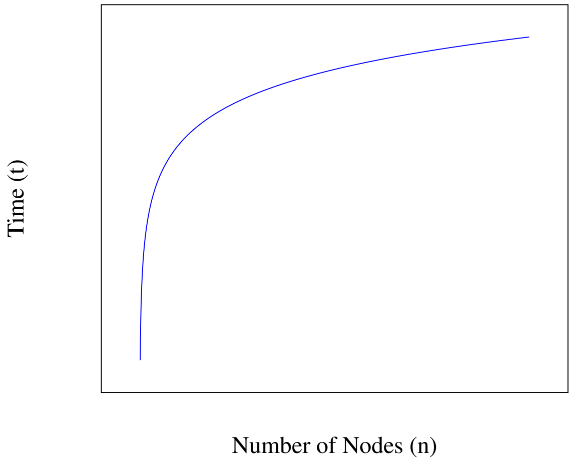


Figure 3: Plotting number of nodes n against time t where $t = nlog(n)$.

6 Egress Path Creation

This section assumes one of the previous two graph creation algorithms was run and there now exists a graph with appropriate nodes and their respective adjacencies. Algorithm 4 creates a path back to the first node of the graph for the user. See Figure 4 for an example overview of a reconstructed path.

Algorithm 4: $create\_pathway(graph)$ Creates a stack containing the nodes comprising the most efficient path back to the beginning of the graph.

```
Input: A graph with appropriate adjacencies already
       created: $graph$
initialize $stack$
initialize $node \leftarrow graph.nodes[0]$
while $node \neq graph.nodes[-1]$ do
    $stack.push(node)$
    $node \leftarrow node.adjacencies[-1]$
end
/* Gives the directions to the user in order
   */
initialize $next\_node$
while $stack$ do
    $next\_node \leftarrow stack.pop$
end
```

6.1 Naive Search

Algorithm 4 navigates through the graph by "jumping" to the last neighbour in the current node's edges. Both of the previous algorithms place the latest recorded neighbour at the end of the adjacency array. The naive algorithm accomplishes this by appending new nodes to the graph as they are created and updates the adjacency arrays accordingly by comparing each existing node to the new node. The octree algorithm emulates this behaviour by comparing a node's ordering ID to the neighbour's ID. If the node's ID is less than the incoming neighbour's ID, the node is appended to the adjacency array. Thus, a stack is the only thing necessary when recreating the shortest path with which the user can navigate back to the beginning of the graph. The time complexity for this approach is linear, $O(m)$, where m is the number of nodes in the egress path; the worst case runtime is $O(n)$, where n is the number of nodes in the graph.

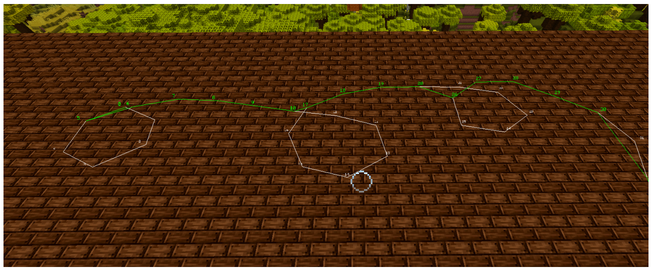


Figure 4: An aerial view of a reconstructed path.

Algorithm 5: A* Search

```
Input: A graph with all adjacencies drawn: graph
Output: A path containing the nodes from the user to the
        closest goal node: path
initialize priority_queue
priority_queue.enqueue(start_node, 0)
while !priority_queue.is_empty do
    u ← priority_queue.deqeueue
    if u == graph.end then
        final_node ← u
        break
    else
        foreach adjacency ∈ u.adjacencies do
            if adjacency.g + distance(adjacency, u) <
             adjacency.g then
                adjacency.previous ← u
                adjacency.g ←
                 u.g + distance(adjacency, u))
                f ← adjacency.g + adjacency.h
                priority_queue.enqueue(adjacency, f)
            if adjacency.visited then
                adjacency.reExpansions + +
            else
                adjacency.visited ← true
            end
        end
    end
end
return path(start_node, final_node)
```

6.2 A* Search

The issue with the previous implementation is that it does not take into account a future feature for this system: multiple users. Even worse, there are some instances of path creation in which jumping to the highest neighbouring node will not actually create the most efficient pathway back for a single user. An edge case which would cause this undesirable behaviour, for example, is if the user continuously navigates in "loops" or interconnected circles. Algorithm 5 is a basic implementation of A*, which replaces the naive approach above. While the naive approach is linear in its time complexity, the average time complexity for A* is $O(b^d)$, where b is the branching factor and and d is the depth of the solution [15].

The heuristic function utilized by Algorithm 5 is the Euclidean distance value from any given node to the final node. This allows for prioritization of nodes, leading the user in a straighter line towards the goal. This heuristic, given only one goal node, is admissible [14], meaning it does not overestimate the cost of a given node to the goal node [14].

7 Giving Directions in *Minecraft*

This section assumes the egress path has already been constructed. As the *Modding Minecraft* section states, the game was created and released in the early 2010 [19]. The programmers behind the original Java-based game made some unusual implementation decisions, particularly in regards to its coordinate system. In essence, it is a right-handed system with unconventional axes. The three issues which we circumvent later in this paper, in order to provide accurate navigation, are:

1. The y-axis is the measure of how high or low the player is relative to the ground, as opposed to the conventional z-axis for this task;
2. The angles between points in the XZ-plane are given clockwise, instead of the conventional counter-clockwise that is generally utilized for trigonometry;
3. The positive Z axis, which is South, is 0 radians in *Minecraft*. Consequently, North, which is negative Z, is π radians.

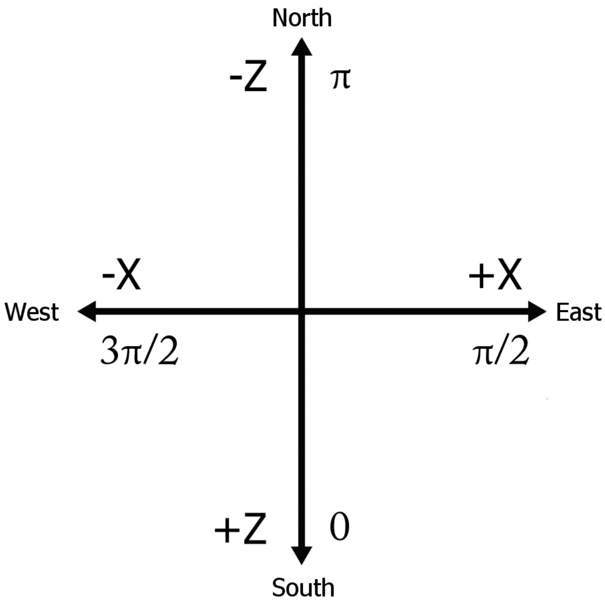


Figure 5: A diagram of the *Minecraft* coordinate system.

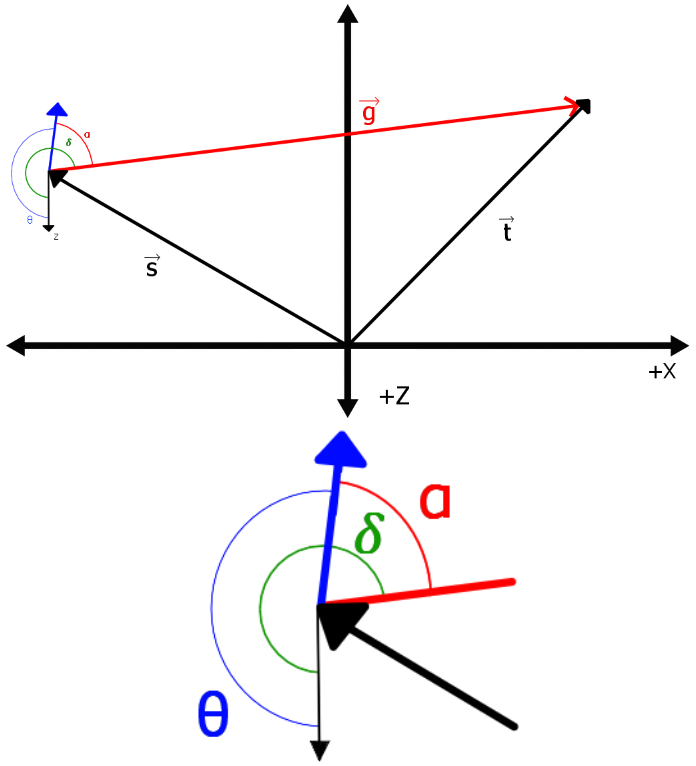


Figure 6: An example of the resulting system of vectors when a player's yaw is facing a different direction than the next desired node. θ and the blue arrow are the player's yaw. $\vec{t}$ is the node vector, $\vec{s}$ is the player's position vector.

Once the shortest path has been created for the player, we need to give him directions to the next immediate node in that path. To do this, we calculate the angle of the next node relative to the player's yaw. We shall call this angle α. First, we take the vector $\vec{t}$, which is the vector pointing to the desired node, and the player's position vector, $\vec{s}$, and calculate $\vec{g}$, which we will refer to as our *direction vector*:

$$\vec{g} = \vec{t} - \vec{s}$$

Once we have our direction vector, we calculate β, which is the angle of $\vec{g}$ relative to the native coordinates to *Minecraft*. For this, we use 2-argument arctangent:

$$\beta = atan2(-g_z, g_x) + \pi$$

Once β is calculated, we have to adjust the angle so that it is relative to the flipped coordinate system that we are now working in. We will call this angle δ:

$$\delta = 3\pi/2 - \beta$$

The final step to find the next desired direction is to calculate the difference between the resulting δ and θ, where θ is the player's yaw. This is α:

$$\alpha = \delta - \theta$$

For the sake of convenience, in order to ensure that the resulting angle is easy to use when giving directions from the suit in the next node, we may add 2π to normalize the value of α. Please refer to Figure 6 for a visualization of an example in this system. Thus, the algorithm used to calculate the next desired directional instruction, relative to the player, is outlined in Algorithm 6.

Algorithm 6: Returns the angle of the node relative to the player's position and yaw.

```
Input: A player's coordinates and yaw: player, the desired
       node: node
$g_x \leftarrow node.x - player.x$
$g_z \leftarrow node.z - player.z$
$\beta \leftarrow atan(-g_z, g_x) + \pi$
$\delta \leftarrow (3\pi)/2 - \beta$
$\alpha \leftarrow \delta - player.yaw$
if $\alpha < 0$ then
    $\alpha \leftarrow \alpha + 2\pi$
return $\alpha$
```

A public repository including these algorithms may be found in [20]. The functionalities which are specific to the ARAIG suit, as discussed below, have been removed for the sake of licensing.

8 Integration with ARAIG

The "As Real As It Gets" (ARAIG) suit, as outlined in IFTech's specifications in [21], has numerous vibratory and stimulus sensors. In order to provide a distinct set of visual cues on the simulated suit, we utilize the vibratory sensors. This way, the player can quickly translate the visual instructions from the simulated suit to following directions inside the *Minecraft* environment. To ensure the system was quick to learn for new users, the program only outputs four directions (which are given in relation to α as calculated in the previous section):

1. **Forward:** $\pi - 1/2 < \alpha < \pi + 1/2$. The user's abdomen and pectorals are stimulated, indicating to them that they should move forward.

2. **Left:** $\pi/2 - 1/2 < \alpha < \pi - 1/2$. The user's left shoulder is stimulated, indicating to them that they should turn left.

3. **Right:** $\pi + 1/2 < \alpha < 3\pi/2 + 1/2$. The user's right shoulder is stimulated, indicating to them that they need to turn right.

4. **Turn around:** If α is not within the previous three ranges, the user is not facing the correct direction. Thus, the user's back is stimulated, prompting them to turn around.

Once the program is running, the user is given a set of initial instructions. The directions relayed to the user are updated every 500 milliseconds given the user's yaw and location. Taking the conditions outlined in Figure 6 as an example, the user's suit output would appear on the screen as shown in Figure 1.

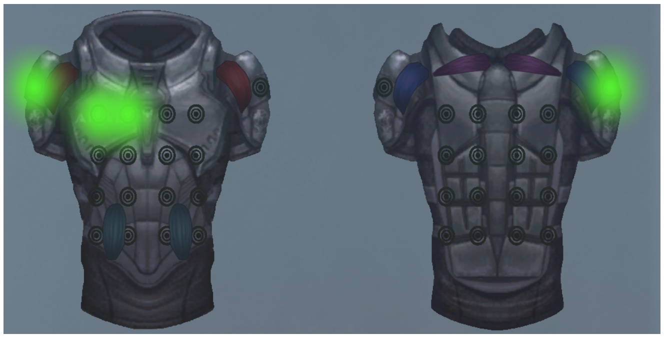
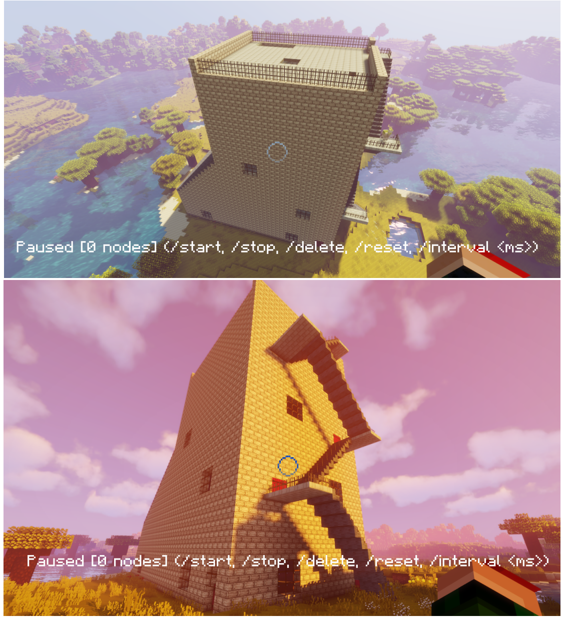


Figure 7: An example of how the ARAIG suit simulation software appears given the conditions in Figure 6.

The ARAIG simulation software integration has been omitted in the supplementary GitHub repository. If readers wish to use this software, they are encouraged to reach out to IFTech at [22]. Once readers receive the appropriate permissions to use the ARAIG visualization software and SDK, they are free to contact Cassandra Laffan or Robert Kozin for access to the full version of this mod and its functionalities.

9 User Testing

A small series of tests were designed to examine whether or not this system is intuitive and quick to learn for new users. The sample size for this study is limited by the non-disclosure agreement (NDA) which protects the ARAIG visualization tool. Consequently, users tested in this study are only those with access to the researchers' machines. As a result, we survey four users of varying backgrounds, all of whom have access to one of the computers with the visualization software available.



```
Paused [0 nodes] (/start, /stop, /delete, /reset, /interval <ms>)
```

```
Paused [0 nodes] (/start, /stop, /delete, /reset, /interval <ms>)
```

Figure 8: Screenshots of the burn house as constructed in *Minecraft*.

The user tests take place in a "burn house", which is a structure built to the specifications as outlined in [23]. Burn houses are standardized buildings in which firefighters may practice navigating structures and fighting fires in a physically simulated environment here in North America. This is a logical testing environment as the system is being designed with first responders in mind. The main goal of this "pilot study" is not necessarily to evaluate how quickly a user may exit a building given different circumstances, but to observe how the average user interacts with the system. We also gather feedback on possible system improvements in hopes of making it more intuitive to new users.

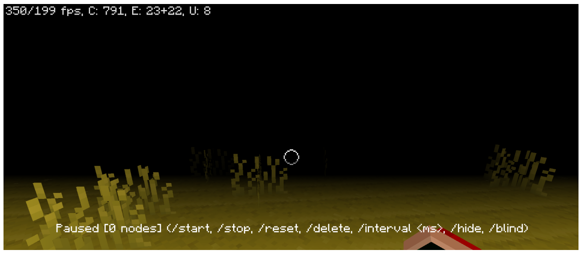
There are four categories of navigation tests, all of which have the same set up and goal: the user is tasked with navigating to the top of the tower, then retracing their pathway down. The navigation upward is not timed as it generally took 60 seconds±1 second; what was timed was the user navigating back to where they started. The users are told it is acceptable to both stray from the path if they were lost or quit for the same reason. The categories for testing are as follows:

1. **Control Run:** The control run times the users navigating to their starting points at the bottom of the tower with high visibility and no navigational assistance.

2. **Low Visibility:** This run is much like the control run, with no navigational assistance. However, distance of visibility for the users is greatly decreased, as per Figure 8.

3. **High Visibility with Path Recreation:** This category of testing allows the users their full field of vision. Their goal of retracing their path is assisted with output on the simulated ARAIG suit on a neighbouring screen. The users are informed that they could opt not to acknowledge the suit's output if they find it to be confusing or a hindrance to their task.

4. **Low Visibility with Path Recreation:** This set of tests has the users navigate up to the top of the tower and back down with low visibility. They are given the output of the simulated ARAIG suit on a neighbouring monitor to assist them in this task. Much like in the previous category, they are informed that if they felt the suit is acting as a hindrance to their task, they can opt to ignore its output.



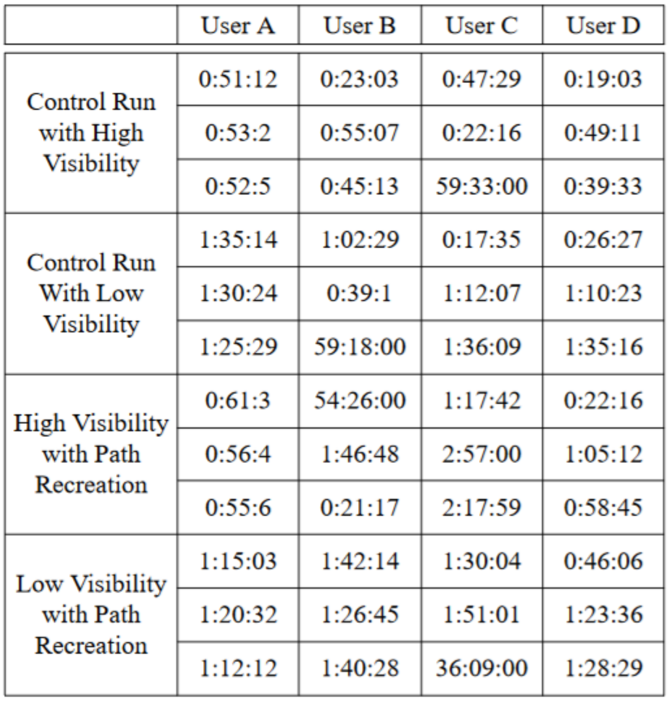
```
350/199 fps, C: 791, E: 23+22, U: 8
Paused [0 nodes] (/start, /stop, /reset, /delete, /interval <ms>, /hide, /blind)
```

Figure 9: An example of a low-visibility environment created by our mod.

Users are also asked to give any and all feedback they believe is pertinent to the experiment.

10 Results

Users were given time to practice controlling the player character in the game before running the tests. The results for each run are shown in Figure 10. Generally, the users were more efficient when not following the directions given by the suit in high visibility situations. However, whether or not the use of the suit made it easier for the user to find their way back to the start in low vision environments seemed entirely dependent on the user's experience with *Minecraft* in the past and the path they took. This will be further discussed below.



	User A	User B	User C	User D
Control Run with High Visibility	0:51:12	0:23:03	0:47:29	0:19:03
	0:53:2	0:55:07	0:22:16	0:49:11
	0:52:5	0:45:13	59:33:00	0:39:33
Control Run With Low Visibility	1:35:14	1:02:29	0:17:35	0:26:27
	1:30:24	0:39:1	1:12:07	1:10:23
	1:25:29	59:18:00	1:36:09	1:35:16
High Visibility with Path Recreation	0:61:3	54:26:00	1:17:42	0:22:16
	0:56:4	1:46:48	2:57:00	1:05:12
	0:55:6	0:21:17	2:17:59	0:58:45
Low Visibility with Path Recreation	1:15:03	1:42:14	1:30:04	0:46:06
	1:20:32	1:26:45	1:51:01	1:23:36
	1:12:12	1:40:28	36:09:00	1:28:29

Figure 10: The results of our short pilot study. Measurements are given in Minutes:Seconds:Milliseconds.

User feedback was as follows:

- Users A, B and D all remarked that having the suit in the low vision environment made navigating somewhat easier if they did not initially take erratic paths.
- Users B, C and D all said that dividing their attention between *Minecraft* and the ARAIG simulation software made navigating efficiently very difficult.
- Users B and C insisted that completing the task would be much easier while wearing the suit.
- User B suggested that instead of stimulating the user on the back to prompt him to turn around, to instead have it indicate that the user should go forward. This would emulate a "pushing" motion.

When we asked users if they would find the system useful when wearing the physical ARAIG suit, all of them responded that yes, it would be more helpful.

11 Discussion

On average, the time taken to navigate the high visibility portion is shorter unassisted versus navigating with the assistance of the simulated ARAIG suit. This is supported by the feedback from most users: splitting their attention between two monitors may be distracting and much more difficult than simply guessing their return path. The shortest pathways, as reflected in Table 1, are either the user taking advantage of the fact there were a few optimal routes to the top of the building from the ground, or them falling down multiple flights of stairs.

Users A and D both have previous experience playing *Minecraft*. They found the ARAIG suit's contributions to their navigation back to their starting point to be beneficial. Users B and C, on the other hand, expressed that the simulated suit detracted from their ability to navigate, as they were already focusing heavily on how to navigate in the low light environment. Shorter pathways up, which were generally just a simple race up the stairs, are reflected in the low visibility runs.

12 Conclusion

In this project, we continue the work we first presented in [2], where we propose a system for assisting first responders in navigating out of low visibility environments utilizing the ARAIG haptic suit. In order to circumvent various obstacles due to the pandemic and supply chain interruptions, we opt to simulate the ARAIG functionality and path recreation in the digital game *Minecraft*. To do so, a mod for the game integrating the ARAIG visualization software is written. This mod tracks a user's movement through the *Minecraft* worldspace. Once the player's points in space are recorded, two different implementations of creating a graph from these points are proposed: the naive approach, which directly compares every node to every other node, and an octree approach, which divides the worldspace into octants, allowing for efficient edge creation. Then, two search algorithms are implemented and compared for finding the most efficient egress path in the resulting graph. The first implementation is another naive approach, which simply jumps from a node to the last recorded node in its edge array. The second implementation is a basic A* algorithm, which, while having a higher time complexity in a worst-case scenario, does not fail in special circumstances.

Finally, we have four users test our software and give us constructive feedback based upon their experiences playing the game in conjunction with the visualization tool for the ARAIG suit. Users generally agree dividing their attention between two programs is difficult and the task would be much easier to complete if wearing the physical ARAIG suit. This feedback is useful for when our research evolves to include firefighters, as we do not want firefighters to feel as if they are disconnected from, or do not understand, the physical input from the suit.

12.1 Future Work

Succeeding this leg of our research are various steps we plan to implement. First and foremost, as we note above, we are looking to implement multi-user functionalities. In doing so, further usage of A* is necessary and the naive implementation for finding the egress path simply will not work. There are numerous variants and alternatives to A* [14], including iterative deepening A* (IDA*) and recursive best first search (RBFS). The next immediate step is thus finishing the *Minecraft* mod, implementing various search algorithms and comparing them on large datasets.

Before we begin testing in real-world environments, we intend on bridging the gap between these two concepts. As our *Results* section mentions, users lamented having to split their attention between two screens while navigating with the simulated ARAIG suit. Further testing in the *Minecraft* environment includes having a user wear the physical suit while navigating the game world. This would allow the user to focus his full visual attention on *Minecraft* while receiving instructions from the ARAIG garment.

Other steps in this research include integrating the algorithms with software which interfaces with the physical world, such as Google's AR Core API [24]. Using our algorithms in more controlled, less noisy real-world environments will allow us to refine these algorithms before we begin integrating them with noisier data such as LiDAR, sonar or a 3D camera data. Finally, we can eventually begin work on integrating these algorithms fully with the physical ARAIG suit and supplementary sensors, as first introduced in [2]. We will then test this technology in a physical burn house as presented in [23].

Conflict of Interest The authors declare no conflicts of interest.

Acknowledgements The researchers acknowledge the funding provided by the Toronto Metropolitan University in conjunction with the Natural Sciences and Engineering Research Council (NSERC).

A special thanks to Jinnah S. Ali-Clarke and Aaron Gill-Braun for providing invaluable support and insight for our research. Another special thanks to Erin MacLellan for her invaluable editing contributions to this paper.

References

[1] S. R., “Covid-19’s impact felt by researchers,” 2021.

[2] C. F. Laffan, J. E. Coleshill, B. Stanfield, M. Stanfield, A. Ferworn, “Using the ARAIG haptic suit to assist in navigating firefighters out of hazardous environments,” 11th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2020, doi:10.1109/iemcon51383.2020.9284922.

[3] C. F. Laffan, R. V. Kozin, J. E. Coleshill, A. Ferworn, B. Stanfield, M. Stanfield, “ARAIG And Minecraft: A COVID-19 Workaround,” in 2021 IEEE Symposium on Computers and Communications (ISCC), 1–7, 2021, doi: 10.1109/ISCC53001.2021.9631428.

[4] W. J. Ripple, C. Wolf, T. M. Newsome, P. Barnard, W. R. Moomaw, “World Scientists’ Warning of a Climate Emergency,” BioScience, **70**(1), 8–12, 2019, doi:10.1093/biosci/biz088.

[5] J. P. Tasker, “Elizabeth May says climate change, extreme events like Fort McMurray fire linked — CBC News,” 2016.

[6] S. Larson, “Massive fire north of Prince Albert, Sask., is threatening farms and acreages — CBC News,” 2021.

[7] W. Mora, Preventing firefighter disorientation: Enclosed structure tactics for the Fire Service, PennWell Corporation, Fire engineering Books & Videos, 2016.

[8] “Fabric,” 2015.

[9] “Fabric Mixin Framework,” 2015.

[10] M. Johnson, K. Hofmann, T. Hutton, D. Bignell, “The Malmo Platform for Artificial Intelligence Experimentation,” in Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI’16, 4246–4247, AAAI Press, 2016, doi:10.5555/3061053.3061259.

[11] S. Adams, I. Arel, J. Bach, R. Coop, R. Furlan, B. Goertzel, J. S. Hall, A. Samsonovich, M. Scheutz, M. Schlesinger, S. C. Shapiro, J. Sowa, “Mapping the Landscape of Human-Level Artificial General Intelligence,” AI Magazine, **33**(1), 25–42, 2012, doi:10.1609/aimag.v33i1.2322.

[12] D. Brewer, N. R. Sturtevant, “Benchmarks for Pathfinding in 3D Voxel Space,” in SOCS, 2018.

[13] F. W. Fichtner, A. A. Diakité, S. Zlatanova, R. Voûte, “Semantic enrichment of octree structured point clouds for multi-story 3d pathfinding,” Transactions in GIS, **22**(1), 233–248, 2018, doi:10.1111/tgis.12308.

[14] M. J. Atallah, M. Blanton, 22.4, CRC Press, 2010.

[15] S. J. Russell, P. Norvig, Artificial Intelligence: A modern approach, Pearson, 4th edition, 2022.

[16] “Protocol FAQ, URL: https://wiki.vg/Protocol FAQ.” .

[17] S. Team, “Spigot, URL: https://www.spigotmc.org/.” .

[18] “MinecraftForge Documentation, URL: https://mcforge.readthedocs.io/en/latest/.” .

[19] Notch, “Minecraft 0.0.11A for public consumption,” 2009 URL: https://web.archive.org/web/20150716115516/http://notch.tumblr.com/post/109000107/minecraft-0-0-11a-for-public-consumption.

[20] C. F. Laffan, R. Kozin, “Octree Path Finding Algorithm,” 2021. URL: https://github.com/cassLaffan/Minecraft Pathfinding.

[21] “ARAIG As Real As It Gets, URL: https://iftech-technologies.com/downloadthe- sdk/.” .

[22] B. Stanfield, M. Stanfield, “Download the ARAIG SDK, URL: https : / / www . firefacilities.com/fire-training-towers/tower-models/thecaptain/.” .

[23] “Fire Department Training Building - Multiple Fire Fighter Trainees,” 2020. URL: https : / / www . firefacilities.com/fire-training-towers/tower-models/thecaptain/.

[24] “Build new augmented reality experiences that seamlessly blend the digital and physical worlds, URL: https://developers.google.com/ar.” .

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 7, No. 4, 59-71 (2022)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

μPMU Hardware and Software Design Consideration and Implementation for Distribution Grid Applications

Ahmed Abdelaziz Elsayed[*], Mohamed Ahmed Abdellah, Mansour Ahmed Mohamed, Mohamed Abd Elazim Nayel

Faculty of Engineering, Department of Electrical Engineering, Assiut University, Assiut, 71511, Egypt

ARTICLE INFO

Article history:
Received: 05 June, 2022
Accepted: 23 July, 2022
Online: 30 July, 2022

Keywords:
Synchrophasor roadmap
Distribution grid
Real-time monitoring system
μPMU hardware and software design
Error and phasor estimation
Computational time minimization

ABSTRACT

This article presents a roadmap for distribution grid μPMU hardware and software design consideration and implantation to ensure high performance within limited computational time of sampling frequency 512 samples/cycle. A proposed 12 channels, multi-voltage level μPMU hardware and rules of voltage and current transducer, analog filter, analog-to-digital converter, sampling rate definition, and PCB design and selection are presented. From the software view, software minimization procedures are implemented to reduce the estimation time of the proposed μPMU to 18 μsec under high sampling frequency operation. Additionally, error estimation and compensation are used to ensure robust performance, while the computational burden of the error compensation stage is reduced by Taylor series linearization. The proposed μPMU is designed to provide traditional phasor, frequency and harmonics measurements besides a point-on-wave under dynamic operation mode. The proposed device is tested under IEEE Std C37. 118.1 and 118.2 and showed accurate phasor estimation up to 0.03% for the magnitude and angle accuracy up to 0.0036^o*, while the frequency is estimated with maximum variation of 0.032% under dynamic operation.*

1 Introduction

Distribution grids modern structure is more dynamic in nature due to the rapid integration of Distributed Energy Resources (DER) and electric vehicles [1]. In light of the presence of these factors, the traditional uni-directional power flow control and protection infrastructure of the distribution grid required to be integrated with adding smartness features, real-time monitoring, and two-way communications to ensure grid stability. A real-time monitoring synchrophasor measurement is a leading technology to observe the power system with reporting rates up to 2 measurements/cycle and via two-way communications.

Traditional Phasor Measurement Units (PMUs) are suitable for transmission systems application, with Total Vector Error (TVE) up to 1% and angle resolution up to 1^o, but could not be applied for distribution grid applications for the following reasons: 1) small distance between buses, 2) high Total Harmonic Distortion (THD), 3) unbalanced conditions, 4) fast load-changing and, 5) large R/X ratio. μPMU is developed as an upgraded version of the PMU with higher sampling frequency and specification up to TVE 0.1% and 0.01^o angle accuracy, respectively [2].

The high cost of the μPMU and the large number of distribution grid buses, which increases the number of installed $\mu PMUs$ to observe the grid, are the main factors constraint the use of synchrophasor technology in the distribution grid applications [3]. However, great investment and consideration are given to the distribution grids due to the rapid and critical changes of their rules over the last years [4]. Motivated to address these challenges, researchers seek to reduce the distribution grid real-time monitoring system cost and enhance the performance of synchrophasor technology. In one track, development and online-learning tools, high-performance frequency and phasor estimation algorithm for low-cost PMUs, and statistical performance and error analysis are studied to allow the researchers to test and improve the synchrophasor technology on one hand, and reduce the hardware specification needed with more accurate phasor estimation abilities on other hand [5]–[6].

In another avenue, the development of low cost μPMU is one of the research tracks to reduce monitoring system costs, by relying on the low Voltage Side (LVS) measurements and downscaling the accuracy and resolution specifications of the designed μPMU [7]–[8]. In [9], a low cost μPMU is proposed by modifying smart meter infrastructure, adding Global Positioning System (GPS) to synchronize the measurements and updating the calculation algorithms to

[*] Corresponding Author: Ahmed Abdelaziz Elsayed, Contact No +201114040626& Email Ahmed-A.Elaziz@aun.edu.eg

https://dx.doi.org/10.25046/aj070409

estimate the input signals phasors. The unit used a sampling frequency of 32 samples/cycle and the main measurements unit address magnitude TVE and angle accuracy of 0.2% and 0.3^0, respectively. The work is extended in [10], where the hardware and software are upgraded with a sampling frequency of 64 samples/cycle. The main measurement unit addressed magnitude TVE and frequency estimation accuracy 0.13% and 11mHz, while the angle estimation accuracy was not studied. The low sampling frequency mitigates both units from reaching high angle resolution and only measured harmonics content up to 16-32 order.

In [11], a low-cost PMU is designed with a simple main measurement unit structure comprised of GPS and Analog to Digital Converter (ADC), while the other components are ignored. The unit is designed using a modified Discrete Fourier Transform (DFT) to mitigate the off-nominal operation with a sampling frequency of 256 samples/cycle. The proposed unit is tested for frequency estimation only and showed a maximum variation of 10 mHz. In [8], Field Programmable Gate Array (FPGA) based μPMU using iterative-Interpolated DFT (i-IpDFT) is used to improve the μPMU latency and speed by relying on a parallel operation to estimate the phasor and the frequency within the sampling time of 512 sample/cycle. The μPMU hardware is considered as the GPS, ADC, and microcontroller process where a scaled signal of 1.25 volt is directly injected into the main measurements unit. The i-IpDFT has a TVE of 0.02% compared with DFT under steady-state, but it is faster. The phasor magnitude TVE, angle accuracy, and frequency variation were not tested and only the latency of the device is considered.

From the above literature the following gaps are remarked:

1. Only a few works considered the low cost μPMU design and implementation and studied uncompleted main measurement unit structure comprised of ADC, GPS, and microcontroller without taking into account the overall structure and Voltage and Current Transducers (VT and CT) underperformance, which limits the integration of the previous models in the actual distribution grid.

2. None of the previous works gave a clear roadmap to deciding the μPMU specifications and the main measurement unit design criteria.

3. Error estimation and compensation of each process inside μPMU were not implemented, which limits the previous design to reach the specification of the commercial μPMU [12, 13].

4. Increasing the sampling frequency, on one hand, improves the estimation of the μPMU while on the other hand, it reduces the computational time and limits using complex hardware, which required software to be optimized.

In this article, a proposed roadmap for μPMU design and implementation is discussed. The proposed roadmap describes the μPMU specification selection, structure design, hardware component, and software design to minimize the μPMU computational effort within a sampling time of 512 samples/cycle. In addition, μPMU source of errors are clarified, the impacts of the error on the estimated phasor are explained and the error calibration and compensation to ensure high phasor estimation are mentioned and implemented. The main contribution of this article is to provide a roadmap for the researcher to design and implement μPMU for distribution grid to further improve the distribution grid applications and boost their performance.

The key contribution of this article is summarized in the following points:

1. We propose robust hardware to ensure high performance of the μPMU under steady-state and dynamic operations that provide traditional and synchronized point-on-wave measurements.

2. We propose a μPMU software that has a very small computational time, which estimates all the output information within the sampling time and sends the information per half cycle.

3. We clarify, estimate, and compensate the μPMU error at each stage to ensure high accuracy of the μPMU measurements and ensure light and fast software for the μPMU using Taylor series linearization.

4. We propose a low-cost high-performance μPMU based on Multi-voltage level measurements.

In order to evaluate the performance of the proposed μPMU, the proposed μPMU is tested to check the measurement accuracy and communication performance using IEEE Std C37. 118.1 and 118.2 [14, 15]. Opal-RealTime Simulator-OP4510 is used as a reference and calibration device for the proposed μPMU to evaluate the accuracy and resolution of the estimated phasor and the frequency variation under steady-state and dynamic operation. Three tests are conducted in steady-state, 1) the magnitude changes by 0.1 p.u in the range of 0.1-1 p.u at nominal frequency, 2) the frequency changes by 0.1 Hz in the range of 49.5-50.5 Hz while the magnitude is 1 p.u and 3) the unit is tested to measure a phase difference of 1, 0.1 and 0.01 degrees to find the angle accuracy and resolution. For the dynamic operation, the phasor and frequency estimation performance is tested under 1) ramp frequency change of 0.2 Hz 2) sudden change in magnitude and phase of 0.1 and 10 degrees to estimate the step response of the unit, and 3) harmonics estimation under dynamic operation and point-on-wave recording.

The rest of this article is organized as follows: in Section 2, the μPMU specification selection and the unit structure design are clarified. In Section 3, the μPMU hardware and software design and implementation are discussed, while the error estimation and compensation are presented in Section 4. The test and validation environment of the unit is mentioned in Section 5 and the result and discussion of the proposed μPMU are presented in Section 6. Finally, the article is concluded in Section 7.

2 μPMU Specification Selection and Structure Design

In this section μPMU specifications selection of the main measurement unit is discussed considering accuracy, resolution, reporting rate, and modes of operations. Also, the sampling frequency selection is clarified from the power quality requirements view and the

μPMU structure design, including number of current and voltage channels and their locations, are described.

2.1 μPMU Specification selection

μPMU specifications are selected based on the list of applications to be addressed. Each application has a level of measurements accuracy and resolution limit, TVE, measurements types, and mode of operation (steady-state or dynamic operation mode), which require different performance and reporting rates. Also, other specifications are decided by the designer such as operation range of the voltage and the current, LCD reporting rate and SD card storing rate,...etc. A survey of 75 distribution grid applications is published by Quanta Technology in 2022, which described the existed distribution grid applications, importance, complexity, and the minimum specifications required for each application from the industrial and research perspective [16], which is useful to define the required specifications limit of the μPMU. Four applications are selected to design the proposed μPMU to meet their limits. A7 (Frequency monitoring) and A26 (Distribution state estimation) are selected as steady-state applications. While for dynamic applications, A42 (Faulted circuit identification) and A55 (Islanding detection for distribution generation) are used. Table.1 summarises the specifications of the four applications. Based on these requirements, the designed μPMU has two operation modes (steady-state and dynamic operations), and the accuracy and resolution limits are 1% and 0.1% for magnitude and 0.1^{o}% and 0.01^{o}% for angle, respectively with latency below 300 msec. For steady-state, the normal measurements are sent with a reporting rate of 1 measurement/cycle for steady-state application, while for the dynamic mode, the point-on-wave data are also sent to the (Phasor Data concentrates) PDC with a reporting rate of 2 measurement/cycle.

Table 1: List of specifications

Applications		***Steady-state***		***Dynamic***	
		A7	***A26***	***A42***	***A55***
Accuracy	Mag	*1%*	*1%*	*1%*	*1%*
	Ang	-	0.1^{o}	0.1^{o}	0.1^{o}
Resolution	Mag	*0.1%*	*0.1%*	*0.1%*	*0.1%*
	Ang	-	0.01^{o}	0.01^{o}	0.01^{o}
Measurements		**Phasor, Frequency ROFC, Harmonics**		**Point-on-wave**	
Latency		***2000msec***		***300msec***	***500msec***
Reporting rate		1 report/cycle		2 reports/cycle	

2.1.1 μPMU Modes of operation and reporting rate

Modes of operation are designed in the μPMU to adjust its specifications to a certain performance. Different modes required extra measurements such as point-on-wave data and a higher reporting rate for better and faster decisions. Prony analyses are used for the detection of the transient or dynamic events to switch the mode of operation [17].

2.1.2 μPMU Measurements type

$\mu PMUs$ are designed to measure the voltage and current phasors of the fundamental and harmonics, frequency, Rate of Frequency Change (ROFC) and provide the point-on-wave data. The Point-on-wave measurements are provided for dynamic applications to allow off-line investigation of the system state at PDC. Measurements types are identified by the group of applications needed to be addressed by the real-time monitoring system.

2.1.3 μPMU Measurements accuracy and resolution

Accuracy and resolution are indications of the measurement quality and the minimum change that could be detected by the μPMU. Both parameters depend on the sampling frequency, ADC bits number, phasor estimation algorithms, hardware deviation, and error compensation techniques, where the final values are required to meet the application requirements. Since these two parameters depend on different factors, the designed μPMU accuracy and resolution are defined under test. Factors that affect the accuracy and resolution are optimized during the design to increase them within the allowed computational and communication time limits.

2.1.4 μPMU TVE

The TVE is a measure of the overall accuracy reached by the μPMU. It is normalized per unit difference between the ideal samples of measurements and the measured value by the units under the test. Eq.1 describes the mathematical expression of the TVE.

$$TVE = \sqrt{\frac{[(X_r^{`} - X_r)^2 + (X_i^{`} - X_i)^2]}{{X_r}^2 + {X_i}^2}} \tag{1}$$

Here, $X_r^{`}$ and $X_i^{`}$ are the real and imaginary parts of estimated measurements by the unit under test.

2.1.5 μPMU Sampling Frequency

μPMU Sampling frequency f_s is one of the main critical factors in the μPMU design. On one hand, high sampling frequency increases the measured phasors, frequency, and harmonics content accuracy and resolution. While on the other hand, high sampling frequency reduces the computational time that existed for calculation, error compensation, and communication, which reduces the ability for stable and reliable operation. From the power quality view, the sampling frequency should be designed to be at least twice the frequency of the maximum harmonic required to be observed to meet the Nyquist theory [18]. This is represented by Eq.2, where h is the order of the maximum harmonic required to be measured and f_o is the nominal frequency of the system.

$$f_s \geq 2h \times f_0 \tag{2}$$

For the distribution grid Low and Medium Voltage sides (LVS and MVS) the range of 35th -50th harmonics content is required to be estimated [19], which required sampling frequency to be at least $f_s \geq 3 \times 50 = 150$ sample/cycle. From the DFT view, $f_s = 256$, 512 and 1024 samples/cycle can be used for the power quality monitoring purpose. For the sampling frequency 256, the estimated phasor magnitude and frequency showed resolution of 0.2 % and 0.2 Hz using FFT, which did not address the resolution limits required in Table.1. Therefore, f_s=512f_o is selected as the minimum sampling frequency to meet these limits. With a sampling frequency of 512f_o, the accuracy level is improved to 0.1% and the deviation in the estimated frequency is 0.1 Hz. However, working on a sampling frequency 1024f_o gives more accurate results, but it limits the computational time of the μPMU to 19.56/16.276 μsec for a 50/60 Hz system. Designing software to read ADC samples, estimate the phasor, frequency, ROFC, harmonics content, compensate for the error, and sent the information within this limited time is a very complex and requires an expensive set-up. Therefore, the sampling frequency is selected to be 512f_o, which allows a time operation of the μPMU to be 39.06/32.55μsec for a 50/60 Hz system, compared with 416.67/347.22 μsec for the traditional PMU. To meet the time requirements, the designed μPMU hardware and software are required to mitigate the excessive and unnecessary processes that consume high time.

2.2 μPMU structure

The structure of the μPMU is the term that describes the number of the μPMU measurement channels, their operation level, and their location in the power system. These parameters are calculated using Optimal μPMU placement ($O\mu PP$), which estimates the μPMU locations and the maximum number of current channels needed to ensure full observability for the studied grid and location of the voltage and current channels. In [20, 21], μPMU with two current channels one located at MVS and the other at LVS of the distribution transformer load with the voltage channels, is found to be the optimal structure that ensures full observability of IEEE 33 and 69 systems. This configuration minimized the overall cost, considering development and running cost, and the proposed μPMU is designed with the same structure. Each channel is comprised of four ports to measure the three phases and the neutral signal. In this case, 12 signals are measured from the power system and 13 steamers are sent to the PDC, which includes the information of the 12 signals and an extra steamer is sent for unit specifications (IP and time information). Figure.1 shows the proposed μPMU structure.

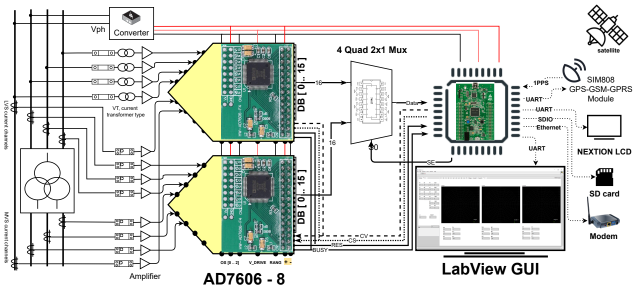


Figure 1: Proposed μPMU structure.

3 μPMU Hardware and Software Design and Implementation

In this section, the μPMU hardware design is clarified to select the appropriate components to meet the computational time and accuracy requirements including VT and CT, ADC, analog filter, GPS, and microcontroller selection. In addition, the μPMU software design to minimize the operation process within the limited time is clarified with time analysis to ensure stable operation.

3.1 Hardware Design

3.1.1 CT and VT selection

The CT and VT are used in μPMU to scale down the signal to the microcontroller level. Due to their critical rules, they are required to be of high accuracy. This is ensured by selecting a high class that meets the accuracy limits and/or robust error estimation and compensation of the CT and VT errors, which is discussed further in section 4. For the current measurements, the CT signals are transferred to a voltage by multiplying the secondary current with the secondary burden impedance. Then the voltage signal from the CT is amplified, so the 120% of the CT rate reflects the full range of the ADC voltage level. Eq.3 shows the current signal amplifier gain value.

$$A_{CT} = \frac{V_{ADC-max}}{1.2 Z_B \times I_{CTs}} \tag{3}$$

For the proposed μPMU a CT 200/1 A-11KV, class 0.5 split core is used for the MVS, and for the LVS a CT 800/1 A-400V class 0.5 split core is used. Both CTs use 0.5-ohm pure resistance as burden impedance and the secondary voltage of both CTs have the same range. So, their current channels have the same amplifier design structure in the μPMU main measurement unit. For the voltage measurement, current-type VT is implemented by using isolating CT 1 : 1 ZMPT101B. The input voltage signal is converted to a current signal using primary side limiting resistance that is selected to limit the primary current to 2mA at the maximum input voltage. The secondary current is multiplied by the burden impedance of 0.5-ohm and the secondary voltage signal is amplified to the ADC maximum voltage level. Eq.4 represents the VT amplifier gain.

$$A_{VT} = \frac{R_{P-limit} \times V_{ADC-max}}{V_{in-max} \times Z_{s-B}} \tag{4}$$

This structure provides a bi-polar voltage and current signals to the micro-controller. In regular cases, DC offset is added to shift up the sinusoidal wave above zero. However, to prevent the undefined DC offset and loading effect, which make the calibration of the unit is more difficult and less accurate as in [22], external bipolar ADC is preferred to be used to allow reading both positive and negative parts of the cycle with high accuracy. The amplifier LM324 is used, which is a general-purpose amplifier with high input resistance, linear operation and DC off-set uncertainty of 7mV.

3.1.2 ADC selection and requirements

ADC converts the analog signals of the voltage and current to digital signals to be analyzed by the microcontroller and extracts useful information from these signals. The ADC should be fast, has a linear operation mode in the operation region, and be of high resolution. Sigma-Delta, Flash, Pipeline, and Successive Approximation Registers are the four types of ADCs that are commonly used. However, SARs are the most used for general purposes due to their high accuracy, low power, zero-cycle latency, and ease to use.

In order to meet all these requirements, AD7606 SAR-type ADC is selected, which has the ability to simultaneous sampling 8-Channels-16 bits of data with a sampling frequency up to $200KHz$. The AD7606 is considered as one of the best selections for Data-Acquisition operation since it can operate with a 5V single supply and can accommodate a true Bipolar analog input $\pm 5V$ and $\pm 10V$, with On-chip 2.5V accurate reference and reference buffer, Analog input clamp protection, Input buffer with 1 $M\Omega$ analog input impedance, Second-order anti-aliasing analog filter that has a 3 dB cutoff frequency of 15 kHz and provides 40 dB anti-alias rejection when sampling at 100 kSPS, a track-and-hold amplifier, Oversampling capability with a pin driven flexible digital filter yields improvements in Signal to Noise Ratio (SNR) to 91.2 dB, reduces the 3 dB bandwidth, and high-speed serial and parallel interfaces to communicate (Serial - Byte - Parallel) [9]. The measured value by the ADC has a resolution of 30 PPM and accuracy of ±12 LSB. Two ADCs are used to measure the 12 μPMU signals using parallel mode operation. Figure.2 shows the timing diagram to read the 12 signals within $2\mu sec$. The following input pins are adjusted to ensure robust sampling with 512 samples/cycle:

1. RANGE: Two range pins are used to control the input operating range of AD7606 (±10,±5) V range. Range Pins can be tied to either supply (±10) or to the ground (±5). In the proposed μPMU the two pins are connected to the ground for (±5) voltage operation.

2. CONVERT: AD7606 has two active high Conversion start pins (A, B), one pin for start every four channels. Both the conversion pins can be tied together and controlled by a single PWM signal from the Micro Controller.

3. Chip select(CS): Active low input pin, which is set to low whenever the host controller wants to perform sampling and read data.

4. READ (RD/SCLK): Active low read signal is used during Parallel operation and for each pulse it will clock all the channel data out.

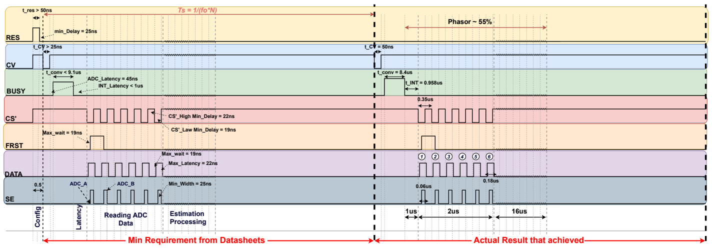


Figure 2: ADC input pins timing diagram.

3.1.3 Analog filters design

An Analog filter is used to filter the scaled signal from the CT and VT. Two low pass filters are used. The first filter removes harmonics content with a frequency above the half sampling frequency $(45 - 55\% f_s)$. The second is the surge filter, which removes the transient pulses due to switching events. A second-order Butter-Worth low pass filter is designed for this purpose. For the designed μPMU the cuff-off frequency of the filter is required to be in the range of 11-15 kHz, while the surge filter is designed with 100KHz. This function is made with the built-in analog filter of the selected AD7606 with a suitable cut-off frequency for this purpose. The transfer function of the filter is estimated from the data-sheet or estimating the filter frequency gain and phase response in a lab experiment.

3.1.4 Micro-controller selection

The microcontroller is the brain of the μPMU, which organizes and controls all the hardware systems to operate correctly, calculate and send the output information. The selected microcontroller should be capable to address the minimum hardware requirements, which allows correct ADC communication, accurate and minimum time of voltage and current phasor estimation using DFT, frequency estimation and ROFC, error compensation, communication with other parts, and send the data to the PDC. Based on these requirements, STM32F407 ARM Cortex M4 ultra-high performance microcontroller is selected. STM32F407 features addressed all the minimum requirements required by the proposed μPMU specification which are:

1. It provides a simultaneous conversion process.

2. Samples are stored in the RAM and shared with a communication buffer through the Direct Memory Access (DMA). In the realized tests, the platform demonstrated enough processing capacity for the calculus and delivery of packages to the PDC.

3. Besides that, it has an Ethernet interface, which is faster than the serial peripheral interface (SPI) and allows the data packages to transfer by directly accessing the memory via DMA without occupying the processor.

4. STM32F407 is based on the high-performance Arm® Cortex®-M4 32-bit RISC core operating at a frequency of up to 168 MHz. The Cortex-M4 core features a Floating point unit (FPU) single precision which supports all Arm single-precision data-processing instructions and data types. It also implements a full set of DSP instructions and a memory protection unit (MPU) which enhances application security.

3.1.5 GPS synchronization signal and 4G communication

The μPMU measurements are synchronized with high accuracy clock from the GPS, which consists of 24 satellites in six orbits and provides a Pulse Per Second (PPS) to all GPS modules. The PPS is used in the μPMU to enable the PWM that controls the ADC conversion process. The GPS module is selected to ensure low Synchronization Time Uncertainty (STU) to reduce the synchronization error. GPS SIM808 module is used in the proposed

μPMU. The module is a GSM, GPRS, and GPS three-in-one functions module, which uses the latest GSM/GPS module SIM808 from SIMCOM, supports GSM/GPRS Quad-Band network, combines GPS technology for satellite navigation, and provides communications abilities using 4G. The GPS module has an STU of equal to 10nsec and uses different GPS and 4G antennas to mitigate the excessive communication and overlap problems.

3.1.6 μPMU Additional parts

In order to allow the useful function of the proposed μPMU, powering circuits and accessories are added to the main measurements units. First, AC/DC center-tapped pulse transformer is used to power the circuit with ±5 V. Second, the Nextion LCD touch screen is used to display the output information to the user every second. In addition, built-in storage Ultra-high-speed SD card is added to record the information with communication protocol SDIO for a period of up to 48 hours in the dynamic model. To allow the user access to the μPMU output information from the PC use, A Graphic User Interface (GUI) is designed using LabVIEW to display, save and export the output information. All hardware components are placed on a Printed Circuit Board (PCB).

3.2 Software Design

The main challenges of the designed software are to meet the time requirements of 39.06/32.55 μsec to read the 12 signals from the ADC, update the 12 phasors, and estimate the frequency and ROFC, harmonics content estimation, digital filtering, and error compensation for each updated measurements. Keeping in mind the communication time to not exceed the half-cycle.

3.2.1 Phasor estimation

Phasors play a leading role in the measurement technologies of power systems. They represent the magnitude and phase angle of the fundamental component of the voltage and current in complex representation refer to a specific time reference. DFT is used to estimate the phasor using measured ADC samples. Non-Recursive DFT estimates the phasors with previous without looking to the last estimated phasors, where DFT is applied to all 512 samples. This process required 512 × 2 product and 512 summation operation. As a result, it consumes a very large time and failed to be implemented in the proposed software structure. Recursive DFT, on the other hand, updates the new phasor based on the previous phasor estimated. This process required two multiplication and one summation process, which reduced the phasors estimation time for the proposed μPMU [23]. The estimated phasor using the Recursive method is defined by Eq. 5.

$$X^h_{N+k} = X^h_{N+k-1} + ((x_{(N+k)} - x_{(k)}) \times [C^h_R(k) - jS^h_I(k)] \tag{5}$$

$$C^h_R = \frac{\sqrt{2}}{N}\begin{bmatrix} cos(\frac{2h\pi 0}{N}) \\ ... \\ cos(\frac{2h\pi(N-1)}{N}) \end{bmatrix}, S^h_I = \frac{\sqrt{2}}{N}\begin{bmatrix} sin(\frac{2h\pi 0}{N}) \\ ... \\ sin(\frac{2h\pi(N-1)}{N}) \end{bmatrix}$$

The DFT coefficient matrices C^h_R and S^h_I are precalculated and stored in the microcontroller to reduce the computational time of the DFT algorithms. It is important to note that the algorithms calculate $\mid X^h_{N+k} \mid^2$ to avoid the square process which consumes huge time. The angle of the phasor is measured using the linearization of tan^{-1} by the Tayler series described in [24]. Eq. (6) shows the phasor angle estimation using the linearization function Γ.

$$\theta_{N+K} = \Gamma[Imag(X^h_{N+k})/Re(X^h_{N+k})] \tag{6}$$

The 12 phasors estimation time is 14 μsec with these modified steps.

3.2.2 Digital filter (Simple Average Filter)

In order to compensate for the error due to leakage phenomena under the off-nominal operation, a Simple Average filter is used for both measured magnitude and angle with a window of half cycles. Therefore, the magnitude and phase of the last $N/2$ estimated phasor are stored in the microcontroller and the average value is calculated using Eq.(7).

$$\mid \overline{X}_{N+K} \mid^2 = \frac{2}{N}\sum_{K+1+\frac{N}{2}}^{K+N} \mid X(i) \mid^2, \overline{\theta}_{N+K} = \frac{2}{N}\sum_{K+1+\frac{N}{2}}^{K+N} \theta(i) \tag{7}$$

Here, $\mid \overline{X} \mid^2$ and $\overline{\theta}$ are the averaged values of both magnitude power 2 and the phasor angle.

3.2.3 Frequency estimation

In order to estimate the frequency and ROFC, the angle-based method is used. In this method, the change in the phasor angle is assumed to be a quadratic time function [13, 14]. The phasor angle is stored for the last N estimated phasors and frequency defined by Eq.(8)

$$f = f_o + \triangle f + \frac{df}{dt}t = f_o + \frac{a_1}{2\pi} + \frac{a_2}{\pi}t \tag{8}$$

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} N\Sigma b\Sigma b^{\cdot 2} \\ \Sigma b\Sigma b^{\cdot 2}\Sigma b^{\cdot 3} \\ \Sigma b^{\cdot 2}\Sigma b^{\cdot 3}\Sigma b^{\cdot 4} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ b \\ b^{\cdot 2} \end{bmatrix} \begin{bmatrix} \theta(0) \\ ... \\ \theta(N-1) \end{bmatrix}$$

Here, b is a vector of size N, which represents the corresponding time of each sample $b(n) = n/f_s$.In order to reduce the storage information inside the microcontroller, the b vector is precalculated and only this vector is stored in the memory. These modifications allow estimating the frequency within 2 μsec.

3.2.4 Communication Protocol

Inside the μPMU UART is used as a common communication bus to share the information between GPS, microcontroller, LCD and the directly connected PC GUI. For data export and communication with external PDC, the μPMU follows the IEEE Std C37.118.2 communication protocol to generate the data frames, uses CRC method to check this data and collect the frames in packets to send the information [15].

4 μPMU Error Estimation and Calibration

Errors are associated with the estimated phasors in each process starting from the CT and VT, analog filter, ADC sampling and digitization process, GPS synchronization, and off-nominal operation error. The output phasor calculated in this case is represented by Eq.(9)

$$\vec{X} = \vec{E}_{CT/VT}\vec{E}_{AF}\vec{E}_{ADC}\vec{E}_{GPS}\vec{P}\vec{X}^{act} \tag{9}$$

where $\vec{X}^{act}$ is the actual phasor without error. $\vec{E}_{CT/VT}$, $\vec{E}_{AF}$, $\vec{E}_{ADC}$, $\vec{E}_{GPS}$ and $\vec{P}$ are the complex gain error due to CT/VT scaling, analog filter error, ADC digitization error, GPS synchronization error and off-nominal complex gain P error, respectively. Each error has it is own nature and is required to be compensated, if possible, to ensure high estimation quality [13]. $\vec{E}_{ADC}$ and $\vec{E}_{GPS}$ depend on a random process and these errors are just minimized by the appropriate selection of the ADC and GPS.

4.1 CT and VT Error estimation and Compensation

CT and VT errors are external errors added to the main measurement unit of the μPMU. In traditional PMUs, CT and VT errors are small and negligible compared with the accuracy and resolution limits required by their applications. In contrast, CT and VT errors are higher in distribution since it is noisier. Also, the accuracy and resolution limits are higher in the distribution grid and transducer errors are required to be compensated to meet these limits. In this article, the CT error is the main focus, since both current and voltage measurements rely on CTs. To ensure high-quality measurements, 1) the selected CT should be of a class higher than the accuracy limits, 2) the CT error should be estimated and compensated. Two methods are existed to compensate for the CT error. In the first method, the CT is loaded under different conditions of pure and distorted input current to provide a set of experimental data. These data are used to rather define the empirical calibration equation as in [22] or learned by using an artificial neural network to compensate for the error as in [25]. In the second method, accurate modeling of the CT is used to estimate the primary current based on the CT secondary current waveform, which allows for estimating the complex gain error for the CT at different operation conditions. To achieve this, the Preisach model is used to simulate the CT B-H curve, which can represent the CT major and minor loop under different operating conditions as in [26]. For the proposed μPMU, the empirical equation is used to compensate for the CT error under steady-state. For the dynamic operation, the Discreet Preisach model is used, based on the point-on-Wave measurements provided to evaluate the CT performance under unpredicted waveform [27]. The CT and VT complex gain error estimated during the calibration is represented in Eq.(10).

$$\vec{E}_{CT/VT} = \vec{I}_s/\vec{I}_p \;\; p.u \tag{10}$$

4.2 Analog Filter Error

The analog filter error is represented by a lag phase shift and reduction in the magnitude. In order to estimate the analog filter error, the transfer function of the filter is estimated from the filter frequency and phase response in the experimental lab or from the datasheet information. Eq.(11) represents the analog filter complex gain error based on the filter parameters.

$$\vec{E}_{AF} = [(1-u^2)^2 + (2\zeta u)^2]^{-0.5} \angle -\Gamma[(2\zeta u)/(1-u^2)] \tag{11}$$

Here, $u = 2\pi f/\omega_n$, ζ and ω_n are the analog filter damping coefficient and natural frequency, respectively. The analog filter error magnitude $\mid \vec{E}_{AF} \mid$ is linearized using Taylor series in the range of ±5 Hz to increase the algorithm's speed and reduce the computational effort on the microcontroller.

4.3 ADC Error

The ADC error is represented by the digitization process when the sample is approximated to the nearest digital level and the sampling error. The estimated phasor is measured with samples including the digitization error. The ADC complex gain error is represented by Eq.(12).

$$\vec{E}_{ADC} = [DFT(x_{(N+k)} + \gamma_{ADC})]/DFT(x_{(N+k)}) \tag{12}$$

where γ_{ADC} is the ADC approximation value that was added to the samples due to the digitization process. γ_{ADC} is a uniform distribution function and limited between in ± half the interval between the two digital levels $\gamma_{ADC} = rand(\pm 0.5 \times LSB)$. The ADC error is randomly variated and can not be estimated or compensated. However, this error is reduced by using the oversampling process of the AD7606 selected in the proposed μPMU. As the number of bits increases the digitization error reduces. For the sampling error, the PWM control signal with high accuracy needs a high accuracy timer. Systick timer with 24-bit ($\approx 5.95 nsec$) is used, which is not perfect for the required sampling frequency because for a 50Hz system the sampling time is 39.06 μsec and the timer reload value is 6562.5≥ (SystemCoreClock/1000) / ($512 \times 50 \times 0.001$). To ensure accurate sampling accumulation and deaccumulation using variable sampling interval control strategy is used [28].

4.4 GPS Error

The μPMU GPS error is represented by the module response time response to the PPS signal. The STU is defined as the maximum response time of the GPS module. The response time τ of the module is assumed to be a Gaussian Distribution function with mean and standard deviation equal to μ= 0.5STU and δ = STU/6, respectively [13]. In the proposed μPMU GPS module SIM808 module is used, which has synchronization time accuracy equals $10 nsec$, which gives an error in the angle of 0.00018^o while GPS magnitude is approximately one. The GPS complex gain error is represented by Eq.(13), which represents a lead phase shift to the estimated phasor.

$$\vec{E}_{GPS} \approx 1\angle 360^o \tau f \tag{13}$$

4.5 Off-nominal Error P-gain

Due to the off-nominal operation, when the frequency deviates from the nominal value, the calculated phasor is multiplied by P and Q complex gains, due to the leakage phenomena. The Q gain, which represents a second harmonics ripple, is mitigated using a simple average digital filter discussed in section 3. The complex gain P error is represented by Eq.(14).

$$\vec{P} = \frac{Sin(\pi \Delta f N/f_s)}{NSin(\pi \Delta f/f_s)} e^{j\pi \Delta f (N-1)/f_s} \tag{14}$$

The P complex gain error magnitude and phase angle are linearized using Taylor series in the range of ±5 Hz to increase the algorithms speed and reduce the computational effort on the microcontroller.

4.6 White Noise

White noise is a natural error added to the measurements due to the environment and electronic distortion. Before exporting the output information, the white noise should be filtered. For this task, the Moving Average Filter (MAF) is used with a window of $N/2$. Eq.(15) shows the MAF representation.

$$y(i)^* = \frac{2}{N} \sum_{n=0}^{\frac{N}{2}-1} y(i-n) \tag{15}$$

where $y(i-n)$ is the output measurements and $y(i)^*$ is the filtered value. Table.2 shows the error contribution of each process on the estimated phasor and the theoretical TVE estimated using Monte Carlo Simulation of 1,000,000 simulations. Transducers error shows the major contribution when they loaded by 5% of the full rate. In order to meet the accuracy limits, the developed μPMU should be calibrated against these errors and compensate for them. Figure.3 shows the μPMU processes structure and operation sequence.

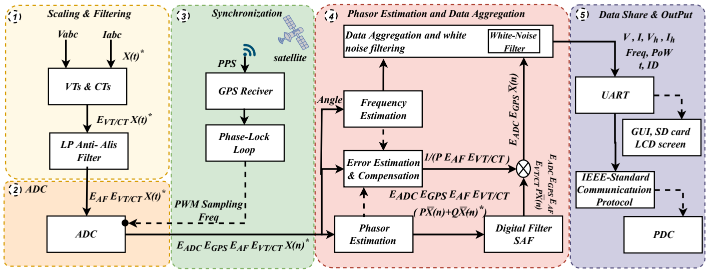


Figure 3: μPMU operation structure.

It is worth noting that in the design of the software and hardware, the measurement accuracy, software speed, and overall cost are the three main performance parameters that we care about during the design. We always looking to select the cheapest hardware that ensures the required level of accuracy with low computational time and reliable operation.

4.7 Taylor series linearization

Error compensation is an important step to improve the proposed μPMU accuracy over the previous work. However, the error compensation algorithm should be designed to be light and fast. Since the total μPMU computational time equals the sampling time, using the non-linear equations of the gain error limits the idea of error compensation. Therefore, all error gains discussed in section 4 are represented by a Taylor series. Eq.(16) shows the linearization of the complex gain error E to L order, which is mainly a function of the measured quantity d.

$$E(d) = a_0 + a_1 d + ... + a_L d^L + e \tag{16}$$

e is the error between the $E(d)$ values and the linearization function. The matrix form of the Eq. (16) for m values is represented by Eq. (17).

$$\begin{bmatrix} E(1) \\ : \\ E(m) \end{bmatrix} = \begin{bmatrix} 1 & d(1) & d(1)^L \\ : & : & : \\ 1 & d(m) & d(m)^L \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ : \\ a_L \end{bmatrix} \tag{17}$$

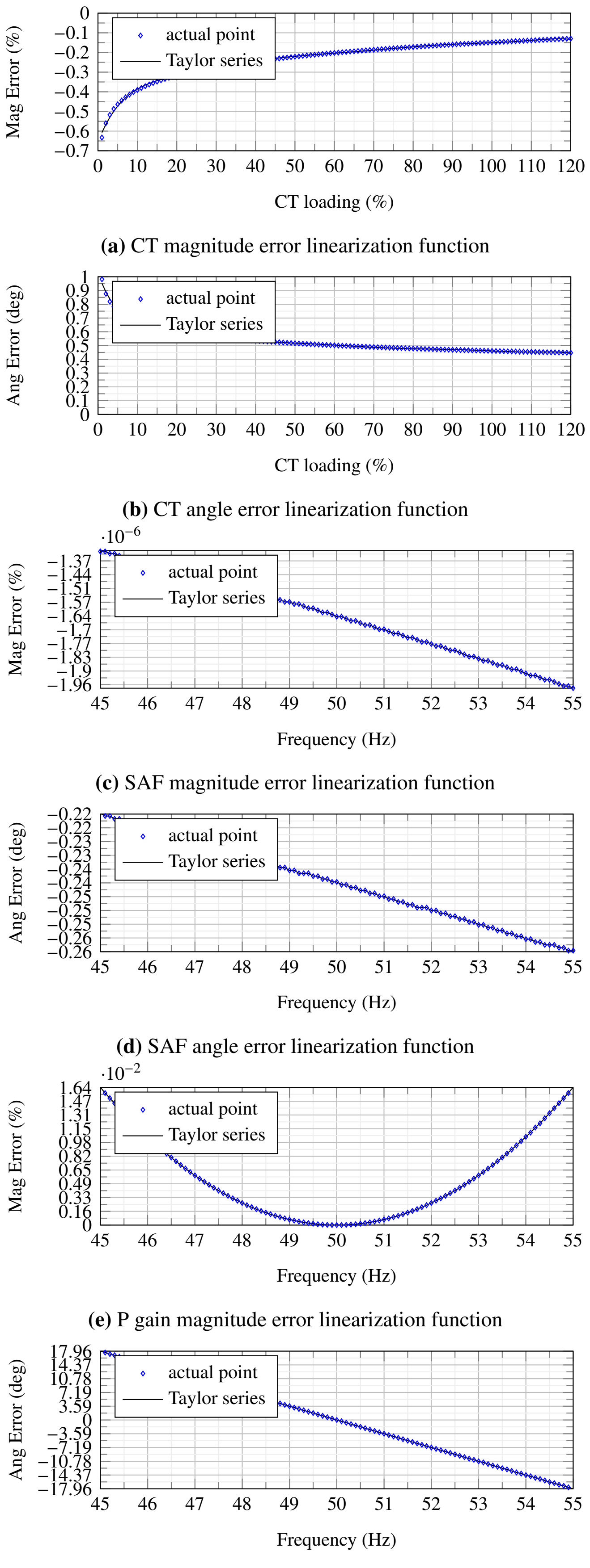
Eq. (17) is solved using Least Square in Eq. (18) to find the A matrix coefficients, using D and E matrix.

$$A = [(D^T . D)^{-1} . D^T].E, \ m \geq (L-1) \tag{18}$$

Using high order reduces the error e between the complex gain error and the linearization function. On the other hand, using high order increases the number of sum and multiplication operation, which increase the computational burden of the error compensation stage. Therefore, the linearization function order is selected as the smallest order to represent the complex gain error with an acceptable error range. It is important to note that for each complex gain two linearization functions are estimated for the magnitude and angle.

Figure.4 shows the estimated Taylor series function for the complex gain errors of the magnitude and angle for the MVS CT as an example of the power transducer, analog filter, and the complex gain P, respectively. The analog filter transfer function is estimated using the given data-sheet gain and phase frequency response. It is worth noting that there is no linearization function for the GPS error or the ADC error. Since the error produced by these elements is randomly generated, the errors of both stages are only reduced through the μPMU main measurement unit robust design such as using oversampling ADC properties, and high accuracy GPS module.

Table. 2 Shows the effect of error compensation, which shows a high improvement in the accuracy of the measurement. It is worth noting that to ensure low operation time, complex gain errors with low effect are not compensated. For example, the magnitude error due to the analog filter is very small compared with the required design limits needed. Adding error compensation for this small error, increase the computational burden on the microcontroller with the limited operation time. On the other hand, the angle error due to the analog filter is high and above the target limits of the proposed design. Therefore, analog filter angle error is considered in the error compensation stage. Gray elements in Table.2 show the error values that are required to be compensated to improve the μPMU accuracy to satisfy the target limits. Table.2 summarizes the performance of the enhancement stage, where the error compensation time reduces from 42 μsec to 0.8 μsec.



(a) CT magnitude error linearization function

(b) CT angle error linearization function

(c) SAF magnitude error linearization function

(d) SAF angle error linearization function

(e) P gain magnitude error linearization function

(f) P gain angle error linearization function

Figure 4: Complex error gains linearization

5 Testing and validation

In this section, the μPMU tests are discussed according to IEEE Std C37. 118.1 and 118.2 to evaluate the main measurement unit operation and performance [14, 15]. The unit steady-state and dynamic operation is examined using six tests and the result is compared with previous work. Opal-RT simulator-OP4510 is used to generate a reference signal to the μPMU in each test and the measurements are sent to the simulator to compare and estimate the error between the two signals. Figure.5 shows the experimental setups.

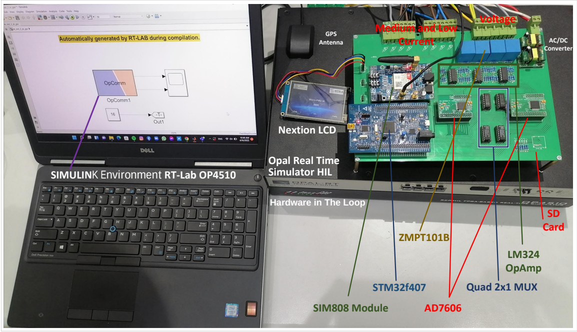


Figure 5: μPMU test and validation environment.

5.1 Steady-state tests

In steady-state operation, three tests are conducted. In the first test, the phasor magnitude estimation performance is tested in the range from 0.1-1 p.u with a variation step of 0.1 p.u. The test is made under nominal operation. In the second test, the frequency estimation performance is tested in the range of 49.5-50.5 Hz with a step change of 0.1 Hz. In the third test, the phasor angle estimation performance is tested using input signals with phase shifts 1, 0.1, and 0.01 degrees to estimate the angle accuracy and resolution.

5.2 Dynamic tests

In dynamic operation, three tests are conducted to test the performance of the μPMU under dynamic operation. In the first test, an input signal with a ramp frequency change of 0.2 Hz/sec is applied to the μPMU, the main purpose of this test is to evaluate the performance of estimating the ROFC and frequency. In the second test, two input signals are applied to the μPMU, one with a sudden change by 0.1 p.u and 10 degrees to estimate the phasor step response of the μPMU and the other signal with a step frequency change by ± 0.5 Hz to estimate the frequency step response. In the third test, a sinusoidal wave of fundamental and 40% harmonics is measured by the proposed μPMU. The test is made for harmonics content up to 16^{th} harmonic order. The μPMU is adjusted to operate on the dynamic mode to provide the point-on-wave measurement in this test. The main purpose of this test is to check the performance of the μPMU to estimate the harmonics content of the input signal and shows the point-on-wave data.

Table 2: Theoretical error limits using MCS and Taylor linearization performance

Error Source		Maximum Error Without Error compensation		Maximum Error With Error compensation		Taylor estimation performance							
		Mag %	Ang^o	Mag %	Ang^o	L		Error % in Taylor fun		Time μsec before Taylor total 48 μsec		Time μsec after Taylor total 0.8 μsec	
						Mag	Ang	Mag	Ang	Mag	Ang	Mag	Ang
CT /VT		1.5	1.5	0.00549	0.44	8	9	0.366	0.44	-	-	0.25	0.25
Analog Filter		6.5e-9	0.252	1.27e-14	1e-6	2	1	1.9e-6	1e-6	4	12	0.1	0.05
ADC digitization		±9.83e-5	±4.9e-5	± 9.83e-5	±4.9e-5	-	-	-	-	-	-	-	-
GPS		≈0	-1.8e-4	≈0	-1.8e-4	-	-	-	-	-	-	-	-
off-nominal	Q	±0.005	-	±0.005	-	-	-	-	-	-	-	-	-
	P	0.0164	±17.965	1.0496e-4	±4e-13	2	1	4.6e-3	4e-13	14	12	0.1	0.05

6 Result and Discussion

In this section the performance of the proposed μPMU is analyzed under the standard tests of IEEE- C37.118.1 and C37.118.2 is discussed.

6.1 Steady-state operation

6.1.1 Test1 (Magnitude step change)

The result of the magnitude step change is shown in Figure.6. The magnitude estimation under nominal operation shows a very accurate result, the TVE is 0.0318% including the current-type VT without error compensation. The bipolar structure shows more stability and less error compared with adding DC off-set as in [22]. The error in the estimated frequency showed a mean of 3.5 mHz and a standard deviation in the measurement of 1mHZ, which resulted in a maximum variation of 6mHz in a transient change in the voltage. However, compared with previously designed units in the literature the proposed μPMU shows lower variations.

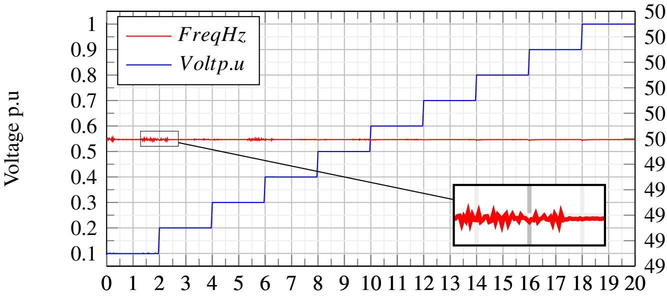


Figure 6: Magnitude step change.

6.1.2 Test2 (Frequency step change)

The result of the frequency step change is shown in Figure.7. The magnitude estimation under off-nominal operation shows degradation in the accuracy. For the magnitude, the TVE increased to 0.7% at the transition period shown by pulses in the graph. Otherwise, the TVE is less than 0.05%. The error in the estimated frequency showed a maximum variation of 0.1 Hz at the transient change, while inside the range stable operation, the frequency maximum variation is 30 mHz. The increased error in both the magnitude and frequency happened due to the off-nominal operation complex gain error P and Q.

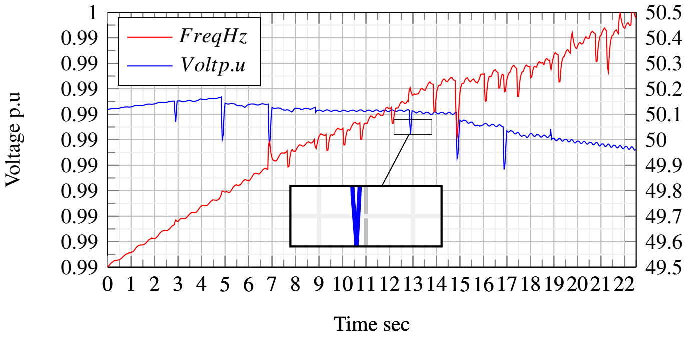


Figure 7: Frequency step change.

6.1.3 Test3 (Angle accuracy and resolution)

Three signals with a phase shift of 1, 0.1, and 0.01 degrees are applied to the μPMU channel. Figure.8 shows the measured angle in all tests. In the three cases, the μPMU showed robust estimation. The mean of the error is constant in the three cases and has a mean of 0.0036^o lag phase shift, due to the analog filter and the complex gain P, and a very low standard deviation equals $2.85^o \times 10^{-5}$. The μPMU angle resolution detected is 0.0036 degrees and reflects the ADC resolution with a sampling frequency of 512 samples/cycle.

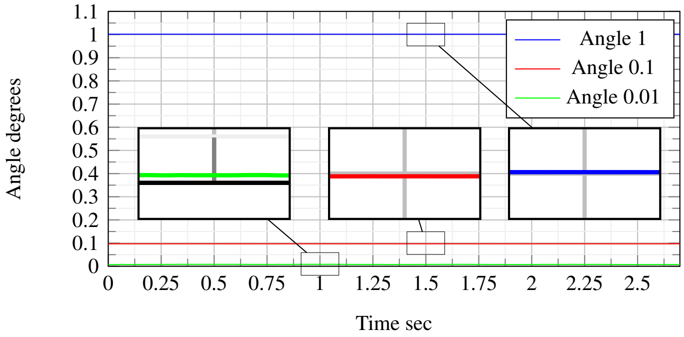


Figure 8: Angle variation.

6.2 *Dynamic Operation*

6.2.1 *Test4 (Ramp frequency change)*

When the ramp signal is applied to the μPMU, the estimated ROFC equals the input signal value of 0.2HZ. Figure.9 shows the reported estimated frequency during the test. At the nominal region, the frequency has low ripple and low error, as the frequency deviated with time (increase or decrease), the off-nominal caused a deviation in the measured frequency. As the difference from the nominal frequency increases, the deviation of the measured frequency also increases due to the mismatch of the analog filter and the off-nominal error due to P and Q complex gain frequency.

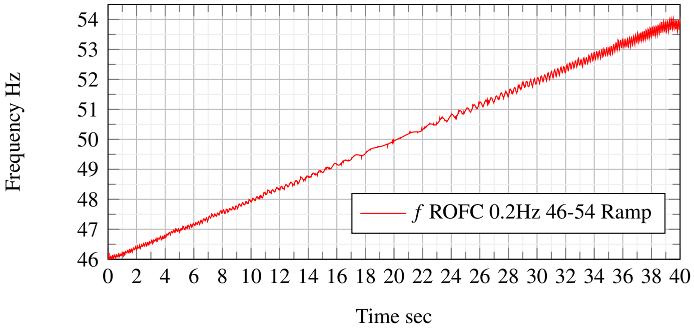


Figure 9: Ramp frequency change.

6.2.2 *Test5 (Phasor and frequency step response)*

In the case of phasor step change, the phasor magnitude and the angle have the same TVE, accuracy, and resolution in the steady-state. Both magnitude and phase response to the change at the same time and have a response time of 20 msec, which represents the reporting rate and it is also equal to the μPMU latency as shown by Figure.10. This value depends on the selected reporting rate due to the operation mode and can be reduced by adjusting the reporting rate of 2 measurements/cycle. In the case of frequency step change by ± 0.5 Hz, at the nominal operation, the variation of the measured frequency is within the range of 4 mHz, which results in a TVE of 0.008%. The frequency of the two signals is changed with a step of ± 0.5Hz and 49.5, which causes off-nominal operation. Therefore, the second harmonics ripple is added to the measured phasor causing a second harmonics to change in the angle and frequency. The SAF mitigates this ripple to a final variation of ± 16 mHZ. Figure.11 shows the reported frequency measurements.

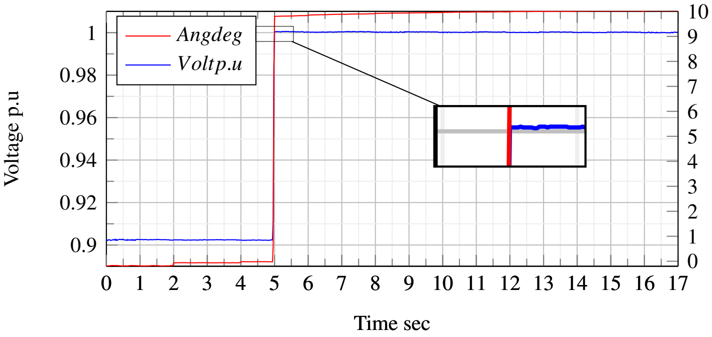


Figure 10: Phasor step response.

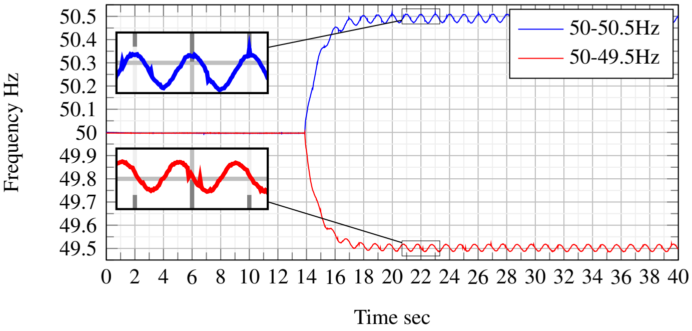


Figure 11: Frequency step response.

6.2.3 *Test6 (Harmonics content)*

In this test, a sinusoidal wave of 60% fundamental and 40% harmonics from 2^{nd}-16^{th} order is measured by the μPMU. Figure.12 represents the point-on-wave data recorded of the first four signals. Figure.13 the p.u error in the estimated harmonics is shown. The maximum error is found to be 1.36% in the 6^{th} harmonics. In all harmonics, the error is around 1.206%. This magnitude shift is made by the amplifier LM324 DC shift and uncertainty in the gain, which could be improved by using a better amplifier in future design. Table.3 shows a comparison between the proposed unit previous proposed units.

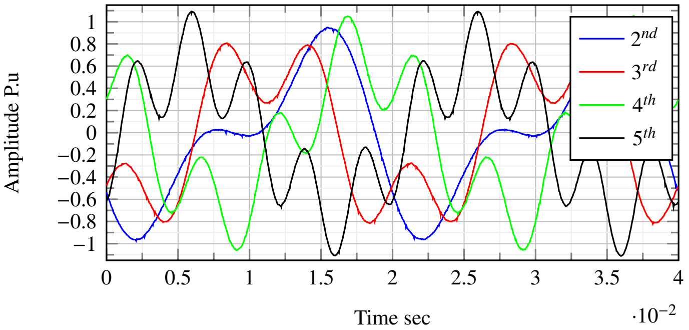


Figure 12: Harmonics Point-on-wave data 2^{nd}-5^{th}.

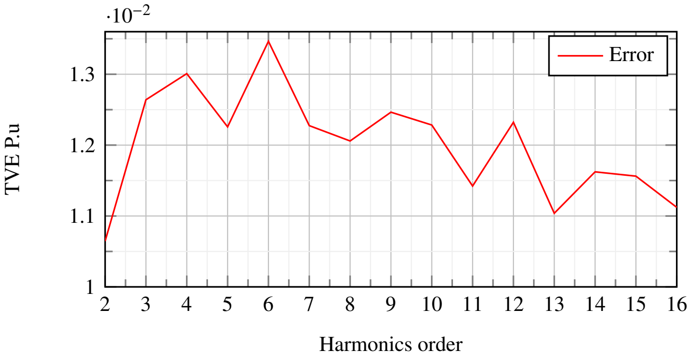


Figure 13: Harmonics estimation error.

Table 3: Comparison of μPMUs

Specifications	HW	Phasor-SW	fs	latency msec	steady state max variation			Dynamic max variation			Cost
					Mag %	Ang deg	Frq %	Mag %	Ang deg	Frq %	K$
Proposed unit	Full structure	Linearized DFT and Error	512	20	0.03	3.6e-3	6e-3	0.05	3.6e-3	0.032	0.25
[10]	Smart meters and GPS	DFT	32	-	0.1	0.3	0.01	0.8	-	0.06	0.3
[11]	Filter, GPS, Mc and communication	DFT	64	22.5	0.062	-	6e-3	0.13	-	0.022	0.4
[12]	ADC, GPS and Mc	Modified DFT	256	-	-	-	0.032	-	-	0.57	-
[13]	ADC, GPS and Mc	i-IpDFT	512	0.07 MC	-	-	-	-	-	-	0.8
PSL	Full structure	-	512	20	0.05	0.003	0.01	-	-	-	12

7 Conclusion

This article presents a roadmap to design distribution level μPMU to meet the requirements of distribution grid applications. The μPMU hardware and software are discussed to ensure the robust design and minimum software operation time. Proposed μPMU shows accurate performance during steady-state and dynamics operation and meets their application requirements with an accuracy and resolution up to 1% and 0.1% for the magnitude and 0.1^o and 0.01^o for the angle. The proposed μPMU address the required limits with a phasor estimation accuracy up to 0.03% for the magnitude and 0.0036 angle resolution, while the frequency is estimated with accuracy up to 0.006% under steady-state. For dynamic operation, both phasor and frequency measurements show a certain level of degradation due to the dynamic and off-nominal error. Power system frequency variation shows a higher effect over magnitude and phase variation. The magnitude and frequency of TVE are degraded to 0.05% and 0.032% under dynamic operation, while the angle variation is the same for both cases. By using linearization, the off-line calculation time is reduced to $18\mu sec$, which allows stable communication. The overall μPMU performance addressed the required specification and showed robust performance under steady-state and dynamic operation.

Acknowledgment

This work was done as part of the research project “ Smart Monitoring system for Distribution Power Grid ” supported by the Egyptian National Telecom Regulatory Authority (NTRA). Also, the authors would like to thank the Smart-Systems Company for providing the real-time simulator and the μPMU test environment.

References

[1] J. A. Momoh, Smart grid: fundamentals of design and analysis, volume 63, John Wiley & Sons, 2012.

[2] A. Von Meier, E. Stewart, A. McEachern, M. Andersen, L. Mehrmanesh, “Precision micro-synchrophasors for distribution systems: A summary of applications,” IEEE Transactions on Smart Grid, **8**(6), 2926–2936, 2017, doi: 10.1109/TSG.2017.2720543.

[3] F. C. Trindade, W. Freitas, “Low voltage zones to support fault location in distribution systems with smart meters,” IEEE Transactions on Smart Grid, **8**(6), 2765–2774, 2016, doi:10.1109/TSG.2016.2720544.

[4] P. De Oliveira-De Jesus, C. H. Antunes, “Economic valuation of smart grid investments on electricity markets,” Sustainable Energy, Grids and Networks, **16**, 70–90, 2018.

[5] X. Zhao, D. M. Laverty, A. McKernan, D. J. Morrow, K. McLaughlin, S. Sezer, “GPS-disciplined analog-to-digital converter for phasor measurement applications,” IEEE Transactions on Instrumentation and Measurement, **66**(9), 2349–2357, 2017.

[6] T. Ahmad, N. Senroy, “Statistical characterization of PMU error for robust WAMS based analytics,” IEEE Transactions on Power Systems, **35**(2), 920–928, 2019.

[7] D. Schofield, F. Gonzalez-Longatt, D. Bogdanov, “Design and implementation of a low-cost phasor measurement unit: A comprehensive review,” in 2018 Seventh Balkan Conference on Lighting (BalkanLight), 1–6, IEEE, 2018.

[8] P. Romano, M. Paolone, T. Chau, B. Jeppesen, E. Ahmed, “A high-performance, low-cost PMU prototype for distribution networks based on FPGA,” in 2017 IEEE Manchester PowerTech, 1–6, IEEE, 2017.

[9] M. C. Garcia, D. Dotta, L. Pereira, M. C. de Almeida, M. R. Paternina, O. L. dos Santos, L. C. da Silva, J. E. d. R. A. Junior, “A development PMU device for living lab applications,” Journal of Control, Automation and Electrical Systems, **32**(4), 1111–1122, 2021.

[10] M. C. Garcia, D. Dotta, L. Pereira, M. C. de Almeida, O. L. Santos, L. C. da Silva, “Design and development of D-PMU module for smart meters,” in 2020 IEEE Power & Energy Society General Meeting (PESGM), 1–5, IEEE, 2020, doi:10.1109/PESGM41954.2020.9281601.

[11] R. N. Rodrigues, L. H. Cavalcante, J. K. Zatta, L. C. M. Schlichting, “A Phasor Measurement Unit based on discrete fourier transform using digital signal processor,” in 2016 12th IEEE International Conference on Industry Applications (INDUSCON), 1–6, IEEE, 2016.

[12] PSL, “Micro-PMU Datasheet,” in [Online], Available:https://powerside.com/wpcontent/uploads /2020/01/Powerside-microPMU-DataSheet-.pdf, Accessed on: Aug.1, 2021.

[13] A. A. E. Elsayed, M. A. Mohamed, M. A. Nayel, “Distribution Grid Phasor Estimation and μPMU Modeling,” in 2022 IEEE International Conference on Power Electronics, Smart Grid, and Renewable Energy (PESGRE), 1–8, IEEE, 2022.

[14] I. Power, et al., “IEEE Standard for Synchrophasor Measurements for Power Systems–Amendment 1: Modification of Selected Performance Requirements,” IEEE Std C37. 118.1 a-2014 (Amendment to IEEE Std C37. 118.1-2011), **2014**, 1–25, 2014.

[15] IEEE, "IEEE standard for synchrophasor data transfer for power systems," IEEE Standard C37. 118.2-2011 (Revision of IEEE Std C37. 118-2005), 1–53, 2011.

[16] NASPI, "Distribution Synchronized Measurements Roadmap Final Report," in [Online], https://www.naspi.org/node/934, Accessed on: Apr.15, 2022.

[17] J. Zhao, G. Zhang, "A robust prony method against synchrophasor measurement noise and outliers," IEEE Transactions on Power Systems, **32**(3), 2484–2486, 2016.

[18] A. G. Phadke, J. S. Thorp, "Phasor measurement units and phasor data concentrators," in Synchronized Phasor Measurements and Their Applications, 83–109, Springer, 2017.

[19] "IEEE Draft Recommended Practices and Requirements for Harmonic Control in Electric Power Systems," IEEE P519/D6ba, September 2013, 1–26, 2013.

[20] A. A. E. Elsayed, M. A. Mohamed, M. Abdelraheem, M. A. Nayel, "Optimal μPMU Placement Based on Hybrid Current Channels Selection for Distribution Grids," IEEE Transactions on Industry Applications, **56**(6), 6871–6881, 2020, doi:10.1109/TIA.2020.3023680.

[21] A. A. Elaziez, M. A. Mohamed, M. AbdelRaheem, M. A. Nayel, "Optimal μPMU placement and current channel selection considering running cost for distribution grid," in 2020 IEEE International Conference on Power Electronics, Smart Grid and Renewable Energy (PESGRE2020), 1–8, IEEE, 2020.

[22] I. Abubakar, S. Khalid, M. Mustafa, H. Shareef, M. Mustapha, "Calibration of ZMPT101B voltage sensor module using polynomial regression for accurate load monitoring," Journal of Engineering and Applied Sciences, **12**(4), 1077–1079, 2017.

[23] A. G. Phadke, J. S. Thorp, Synchronized phasor measurements and their applications, volume 1, Springer, 2008.

[24] G. Daoud, H. Selim, M. M. AbdelRaheem, "Micro phasor measurement unit phasor estimation by off-nominal frequency," in 2018 IEEE International Conference on Smart Energy Grid Engineering (SEGE), 53–57, IEEE, 2018, doi: 10.1109/SEGE.2018.3023680.

[25] M. S. Ballal, M. G. Wath, H. M. Suryawanshi, "A novel approach for the error correction of CT in the presence of harmonic distortion," IEEE Transactions on Instrumentation and Measurement, **68**(10), 4015–4027, 2018.

[26] A. Rezaei-Zare, R. Iravani, M. Sanaye-Pasand, H. Mohseni, S. Farhangi, "An accurate current transformer model based on Preisach theory for the analysis of electromagnetic transients," IEEE Transactions on Power Delivery, **23**(1), 233–242, 2007.

[27] M. Andreev, A. Suvorov, N. Ruban, R. Ufa, A. Gusev, A. Askarov, A. Kievets, "Development and research of mathematical model of current transformer reproducing magnetic hysteresis based on Preisach theory," IET Generation, Transmission & Distribution, **14**(14), 2720–2730, 2020, doi:10.1049/iet-gtd. 2018.6796.

[28] W. Yao, L. Zhan, Y. Liu, M. J. Till, J. Zhao, L. Wu, Z. Teng, Y. Liu, "A novel method for phasor measurement unit sampling time error compensation," IEEE Transactions on Smart Grid, **9**(2), 1063–1072, 2016, doi: 10.1049/smg.2016.6796856.

Advances in Science, Technology and Engineering Systems Journal Vol. 7, No. 4, 72-78 (2022)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

A Machine Learning Model Selection considering Tradeoffs between Accuracy and Interpretability

Zhumakhan Nazir[*], Temirlan Zarymkanov, Jurn-Guy Park

School of Engineering and Digital Sciences, Computer Science Department, Nazarbayev University, Nur-Sultan, 010000, Kazakhstan

ARTICLE INFO

Article history:
Received: 31 May, 2022
Accepted: 14 July, 2022
Online: 16 August, 2022

Keywords:
Interpretable Machine Learning (IML)
Explainable ML
Model Selection techniques

ABSTRACT

Applying black-box ML models in high-stakes fields like criminology, healthcare and real-time operating systems might create issues because of poor interpretability and complexity. Also, model building methods that include interpretability is now one of the growing research topics due to the absence of interpretability metrics that are both model-agnostic and quantitative. This paper introduces model selection methods with trade off between interpretability and accuracy of a model. Our results show 97% improvement in interpretability with 2.5% drop in accuracy in AutoMPG dataset using MLP model (65% improvement in interpretability with 1.5% drop in accuracy in MNIST dataset).

1. Introduction

This paper is an extension of the work originally presented in ICITEE 2021 with 1) addition of classification problems and 2) more clarified outcomes (i.e. graphs, tables) [1].

ML models are widely used in various fields including public health and the judicial system. However, the majority of the state-of-the-art estimators could be categorized as 'black-box' models with poor accountability and transparency [2]. For instance, the CNN model learned to detect metal token on the corner of the radiology image instead of the image itself (no accountability) [3]; because the model is black-box it is hard to notice such behavior (no transparency).

In [4], research interest in interpretability in model building is rising. Unfortunately, because of the absence of quantitative assessment metrics, evaluating interpretability is not a trivial goal. According to [5], [6], interpretability is inversely proportional to accuracy. Therefore, one realistic approach is to trade accuracy for interpretability, specifically, is it possible to create simpler (easily interpretable) models with high enough accuracy (drop in accuracy to a certain threshold)?

To address the above problem, we acquire a simple and effective numerical interpretability metric-simulatability operation count (SOC) [7], following the major contributions of this work:

- Evaluate interpretability and accuracy of commonly used models for regression and classification tasks: tree-based models, multi-layer perceptron (MLP) and support vector machine (SVM).
- Propose and apply methodology for a trade-off between interpretability and accuracy to enhance interpretability of the models, by letting accuracy to drop up to certain limits.

2. Motivation and Related work

2.1. Motivation

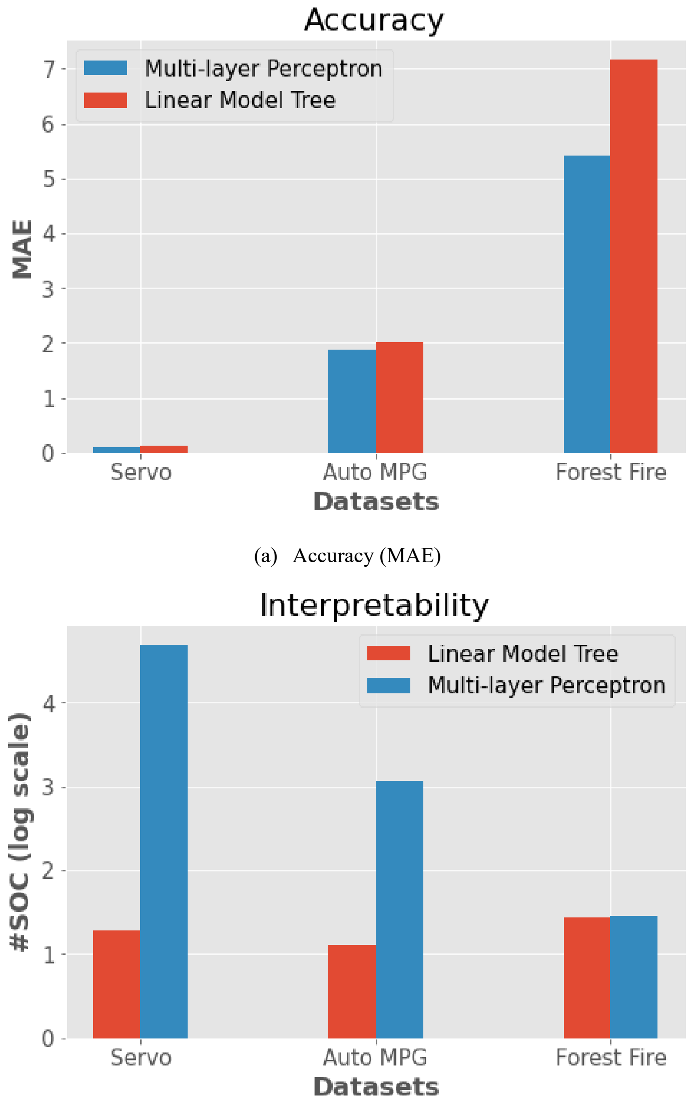
Even though supremacy of black-box models led to their extensive usage, they have lower interpretability compared to tree-based models (e.g., linear model tree), which can compete with other models on both regression and classification tasks. Complexity of black box models result in higher accuracy in general. However, they have lower interpretability with respect to tree-based models (e.g. decision tree) which can achieve competitive performance on both classification and regression tasks. As depicted in Figure 1a, the linear model tree (LMT), compared to MLP regression, has almost the same accuracy results (MAE) on AutoMPG and Servo datasets, and worse results (higher MAE) on Forest Fire dataset. While the interpretability level of LMT is remarkably higher (lower SOC) than MLP in the AutoMPG and Servo datasets (Figure 1b).

For comparatively simple datasets (AutoMPG and Servo), LMT model can be used to increase interpretability with a small accuracy degradation. On the other hand, when the degradation of

[*] Corresponding Author: Zhumakhan Nazir, zhumakhan.nazir@nu.edu.kz

https://dx.doi.org/10.25046/aj070410

accuracy cannot be neglected for complex datasets (i.e. Forest Fire), we can prune a complex model by hyper-parameter tuning in order to raise interpretability level (e.g. by decreasing the number of neurons and/or hidden layers in MLP) within practical accuracy range. As a by-product, size of the model may be reduced, training and inference speed may be increased.



(a) Accuracy (MAE)

(b) Interpretability (#SOC)

Figure 1: Motivating Example: Accuracy and interpretability of LMT and MLP Algorithms

2.2. Related Work

Interpretability/explainability of ML models and explainable AI are now emerging research areas due to the wide usage of AI technologies [8]. According to [2], it is favored using simple and interpretable models as they are capable of replacing sophisticated ’black box’ models. In [9], if-then-based rules are extracted from SVM using a two-step method: first run SVM on data and obtain the set of support vectors, then another interpretable model is trained. In [10], the author proposed a human-based proxy metric that is derived from evaluation of model interpretability by humans or a black-box model’s post-hoc interpretation. Authors of [11] studied a simulatability and a ‘what if’ local explainability of logistic regression, neural network and decision tree. They proposed the metric of interpretability as the run time Operation Count (OC). According to [5] the Simulatability Operation Count (SOC) evaluates interpretability for several regression models through the proposed formula. The experiments in this paper use SOC formulas for comparing interpretability of selected models in our experiments.

3. Methodology

3.1. SOC metric

Interpretability of algorithms can be evaluated in terms of simulatabilty. Simulatability Operation Count (SOC) - the number of arithmetic operations needed to execute an algorithm. According to [8], SOC can be a proxy metric for simulatability. For instance, a linear regression model with 10 variables does 10 multiplications and 9 additions, thus its SOC is 19. More detailed derivation of SOC of estimators can be found in [5].

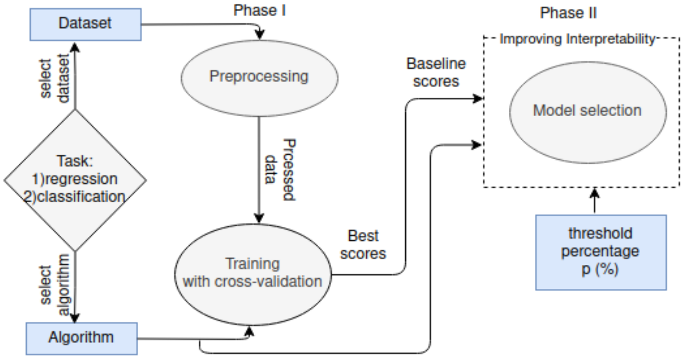


Figure 2: An Overview of the Trade-offs Methodology

3.2. Workflow of the experiment

The workflow of the experiment is shown in Figure 2. Phase I is divided into Data Preprocessing, followed by Model Training. In the first stage 1) categorical entries of datasets are converted to numerical with OrdinalEncoder [12]; 2) StandardScaler [12] is applied to reduce effects of entries on regression coefficients; 3) outliers which has z-score bigger than 3 are removed from the dataset [13]; 4) correlated entries are dropped to prevent multicollinearity (i.e. Variance Inflation Factor (VIF) is larger than 10) [14]. In the next stage (model training), hyper-parameters are selected by applying GridSearchCV implementation of sklearn [12].

Phase II consists of a model selection method that we are proposing. In the first step, the SOC scores of the chosen estimators at the previous training phase will be evaluated. Next, we repeatedly run a model selection process to decrease SOC scores by letting the accuracy to drop by up to a limit set by threshold percentage (*p%*) from the highest accuracy values achieved in the training stage (trade-offs between accuracy and interpretability). The threshold percentage is chosen arbitrarily between 0 and 15%, but in reality its optimality depends on the specifics of the task (i.e. error tolerance and requirements like transparency and accountability). For example, on Figure 1a LMT and MLP have almost similar accuracy, on Figure 1b LMT has much lower SOC. In such cases, LMT is a suitable candidate for the tasks that require interpretability of algorithms.

Such trade-off can be done by tuning parameters of models influencing the SOC according to Table 1.

Table 1: SOC formula of algorithms [7]

Estimator	K_t or A_t	SOC formula
LMT	N/A	2D + 2P + 1
DT	N/A	2D + 1
MLP	A_t	$2 \times N_{H+1} + \sum_{h=1}^{H}(2 \times N_h + A_t) \times N_{h+1}$
	Relu	$A_t = 1$
	Sigmoid	$A_t = 4$
	Tanh	$A_t = 9$
SVM	K_t	SV × (K_t+ 2)
	Linear	$K_t = (2P - 1)$
	Polynomial	$K_t = (2P + 1 + d)$
	Sigmoid	$K_t = (2P + 10)$
	RBF	$K_t = (3P + 1)$

Following the feature selection stage, the number of variables in the dataset (*P*) is fixed, the depth (*D*) could be decreased to reduce SOC in Decision Tree (DT) and in LMT.

The type of activation functions (A_t), the number of hidden layers (*H*) and neurons (*N*) can be tuned in MLP to obtain lower SOC values. Lastly, selecting a simpler kernel function (K_t , e.g. Linear) and decreasing the number of support vectors (*SV*) (using NuSVR and NuSVC [12]) and by tuning hyperparameters reduces SOC in SVM.

Table 2: Servo Features

Feature	Description
motor	A,B,C,D,E
screw	A,B,C,D,E
pgain	3,4,5,6
vgain	1,2,3,4,5
class	0.13 to 7.10

4. Experimental Results

4.1. Experimental Setup

Datasets for Regression: For our experiments three test datasets (from complex to simple) are used: Forest Fire (complex) [15], Auto MPG (medium) [16] and Servo (simple) [17].

1) *Servo*: There are 5 variables with target class, and 167 data instances. Value of each variable is discrete, except for the target, it is continuous within the range [0.13, 7.1] and is a servo-mechanism's raise time. Detailed descriptions are explained in the TABLE 2.

2) *Auto MPG*: There are 9 variables and 398 data instances, the target variable is 'mpg' (miles per gallon). Detailed descriptions are in the TABLE 3.

Table 3: Auto MPG features

Feature	Description
mpg	**miles per gallon, continuous output variable**
model year	version of a car
cylinders	power unit of engine
displacement	measure of the cylinder volume
horsepower	power of engine produces
weight	weight of car
acceleration	amount of time taken for car to reach a velocity of 60 miles per hour
origin	multi-valued discrete
name	name of the car

3) *Forest Fire*: There are 517 data instances and 13 vari- ables including a target class 'area'. Full descriptions are in the TABLE 4.

Table 4: Forest Fire features

Feature	Description
area	**in ha, 0 means less than** 1ha/100 (=100m2)
X,Y	coordinates of place of fire
month, date	categorical value from jan. to dec. and mon. to sun. correspondingly
temp, wind, rain	meteorological data
RH	relative humidity
FFMC,DMC,DC,ISI	components of Fire Weather Index (FWI) of the Canadian system

Datasets for Classification: The classification datasets include Iris (simple) [18], MNIST (medium) [19] and Pima Indian Diabetes (complex) [20].

4) Iris: The dataset contains 5 features with 1 target class and 150 instances. Features of Iris dataset are real values and described in TABLE 5.

Table 5: Iris Features

Feature	Description
sepal length	1.0 - 6.9 cm
sepal width	0.1 - 2.5 cm
petal length	4.3 - 7.9 cm
petal width	2.0 - 4.4 cm
class	**Setosa, Versicolour, Virginica**

5) MNIST: The dataset has 784 features and 70000 (60k training and 10k test images) instances, predicting one of 10 digits. Features of MNIST consists of a 28x28 array of real values of all pixels in the picture.

6) Diabetes: There are 768 instances of 9 features of real values, which are described in TABLE 6. The outcome is positive or negative for the diabetes test.

Table 6: Diabetes Features

Feature	Description
Pregnancies Glucose Blood Pressure Skin Thickness Insulin BMI Diabetes Pedigree Function Age **class**	0 - 17 0 - 199 0 - 112 0 - 99 0 -846 0.0 - 67.1 0.078 - 2.42 21 - 81 **0, 1**

Algorithms for Regression: LMT is a model [21] and its implementation was according to M5 design [17]; Scikit- learn library's [12] MLP Regressor and SVR were used in our experiment. Accuracy metrics is a Mean Absolute Error (MAE)

Algorithms for Classification: DT, MLP Classifier and SVM implementations of scikit-learn library [12] were used to deal with classification task. The percentage of correct predictions (Accuracy) is used as an accuracy metric.

4.2. Results and Analysis

Preprocessing for Regression: Preprocessing stage allowed to reduce Servo dataset to 152 data instances, Auto MPG to 367 ('horsepower' and 'displacement' are dropped because of collinearity issue, 'name' variables is not used), and Forest Fire to 468.

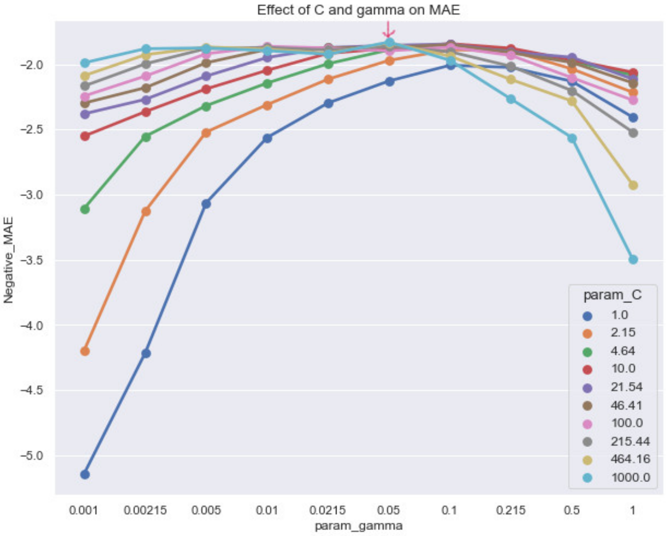


Figure 3: SVR training on Auto MPG dataset.

Model Training for Regression: As mentioned earlier, algorithms are trained with GridSearchCV allowing us to test a broad range of hyper-parameters. Figure 3 shows the process of training SVM on Auto MPG. The lowest error is at C = 1000 and gamma = 0.05 and sub-optimal configuration is obtained by concave down graph.

Table 7: Accuracy Performance (MAE) of Trained Models with references. (Lowest error values in bold).

	Servo	Auto MPG	Forest Fire
LMT	0.133	1.889	6.847
MLP	**0.096**	1.890	5.376
SVR	0.183	**1.830**	**5.212**
Lin. Reg.	0.863	2.304	6.723
Other Ref.	0.220 [22]	2.020 [23]	6.334 [24]

Overall outcomes of the model training phase are in Table 7. SVR performs better than other models on Auto MPG and Forest Fires datasets and MLP on Servo dataset. Performances of the Scikit Learn's Linear Regression and other references are provided for comparison.

Model Training for Classification: Like the regression training phase, GridSearchCV is used to find best combinations of hyperparameters for the models. One of the examples of parameter-tuning is shown in Figure 4, where optimal values are gamma = 0.00003 and C = 4.64.

Table 8: Accuracy Performance of Trained Models with references. (Highest accuracy values in bold).

	Iris	MNIST	Diabetes
DT	96.7%	79.0%	75.4%
Random Forest	97.7%	**97.2%**	76.5%
MLP	83.3%	94.9%	75.7%
SVM	**98.7%**	96.0%	**77.1%**
Other Ref.	98.7% [25]	99.7% [26]	76.0% [27]

Training stage's results are provided in Table 8. On MNIST dataset Random Forest has the best results, while SVM outperforms other models on IRIS and Diabetes dataset. As a comparison, the results of the Random Forest model from the Scikit Learn library were provided.

Model Selection for Regression: Following two approaches to improve interpretability are discussed in this paper: 1) model can be substituted by simpler model and 2) the same model is simplified by tuning its hyperparameters (e.g. reducing the number of neurons or layers in MLP).

Figure 5 shows the results of the first approach and the idea behind it is to demonstrate the behavior of models optimized for interpretability applying the trade-offs method. The point (0, 0) corresponds to the baseline accuracy and SOC score of the estimator on the given dataset. These are the first (top most) entries of each algorithm and dataset pair on Table 9. For example, for LMT and Servo, baseline score is (0.133, 19). The percentage of the increase or decrease is calculated with respect to the baseline score. Next entry in LMT and Servo on Table 9 is (0.134, 17), which corresponds to (0.75 = (0.134 − 0.133)/0.133 × 100, 10.52 = (19 − 17)/19 × 100) point on the blue graph (encircled with red) on Figure 5.

All estimators (MLP, SVR and LMT) behave similarly on regression task. For example, on the Servo dataset SOC is reduced significantly (approximately by 85%, 17% and 11%) for small raise in error (2.1%, 1.6% and 0.75% respectively). MLP is the most accurate estimator for Servo (with 0.096 MAE value). If 2% reduction in accuracy is feasible for Servo dataset, MLP's interpretability could be increased by 85%. Alternatively, if MAE value of 0.133 is acceptable for Servo dataset, LMT algorithm with SOC value of only 19 could be used instead of MLP.

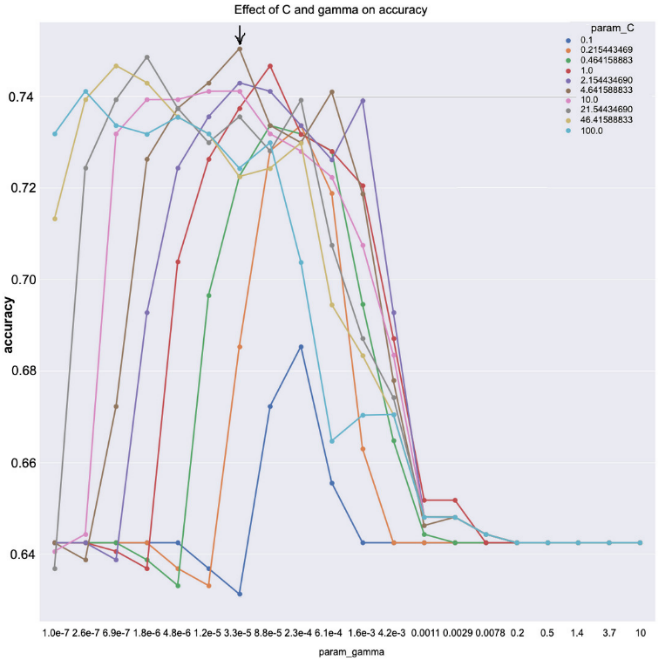


Figure 4: Training SVM on Diabetes dataset.

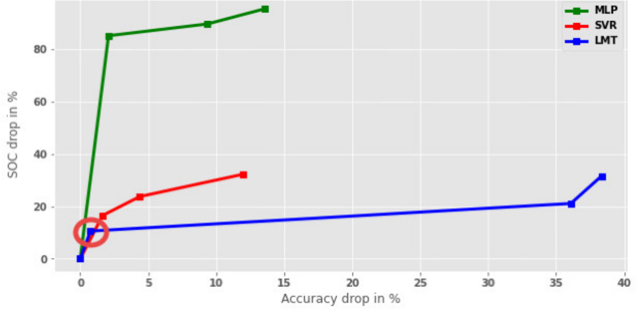


Figure 5: Comparison of Models in Accuracy and Interpretability on the Servo dataset.

Table 9 is a supplement of Figure 5 with additional results of Forest Fire and Auto MPG datasets. Same as in Figure 5, notable advancement is achieved in interpretability with a small degradation in accuracy. For the Auto MPG and Forest Fire datasets using LMT, the most accurate results are obtained with tree depth of 1 (see Figure 7), hence the model cannot be simplified further.

Figure 6 shows the results of the second approach with three datasets using MLP. The similar behavior is observed on all datasets - SOC of the MLP model is decreased notably with small degradation in accuracy. The elbow (turning) points are feasible candidate points for effective trade-off between interpretability and accuracy, since after these points (from left to right) the slopes of the graphs sharply drop. For instance, in Auto MPG dataset (line in red) interpretability is improved (reduction of SOC) by 97% with 2.5% reduction (raise in MAE) in accuracy. The similar pattern is observed in the rest of the datasets.

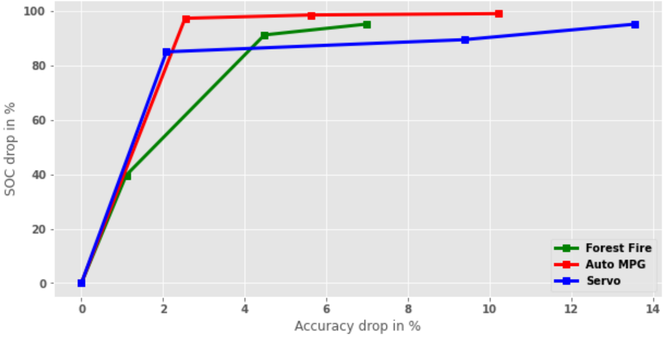


Figure 6: Trade-off between Accuracy and Interpretability in MLP estimator.

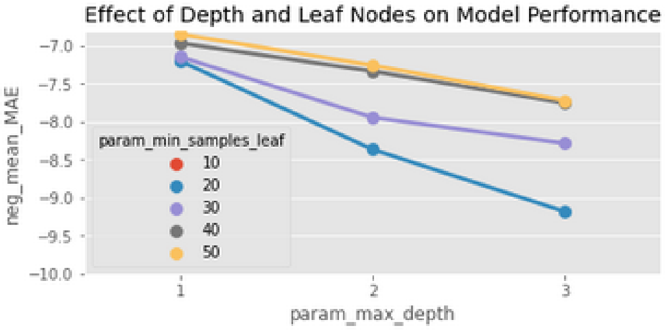


Figure 7: Performance of LMT on Forest Fire dataset.

Model Selection for Classification: For the classification task, similar approaches as for regression were applied; and Figure 8 summarizes the results of the first approach where trade-offs between accuracy and interpretability for all the models (DT, MLP and SVM) on MNIST dataset are depicted. It could be seen from Figure 8 that all graphs start at point (0, 0), which are the baseline scores for accuracy and interpretability. These baseline score correspond to the last entries (with highest accuracy and SOC values) of each algorithm and dataset pair on Table 10. For instance, for MLP and MNIST baseline score is (94.9%, 10719). Percentage change in SOC or accuracy are calculated with respect to the baseline score, for example, next entry in MLP and MNIST on Table X is (93.4%, 3879), and it gives a red point (1.5% = (94.9% - 93.4%), 64% = (1-(3879/10719))*100) on Figure 8. Overall, models perform in the same way on the classification task. For instance, interpretability could be increased dramatically (by 7.3%, 64%, and 40%) in exchange for a small decrease in accuracy (6.62%, 1.5%, and 3.2% correspondingly). SVM is the most accurate model (accuracy 96%) for MNIST dataset, since DT is interpretable intrinsically, its interpretability could not be improved effectively with trade-offs method.

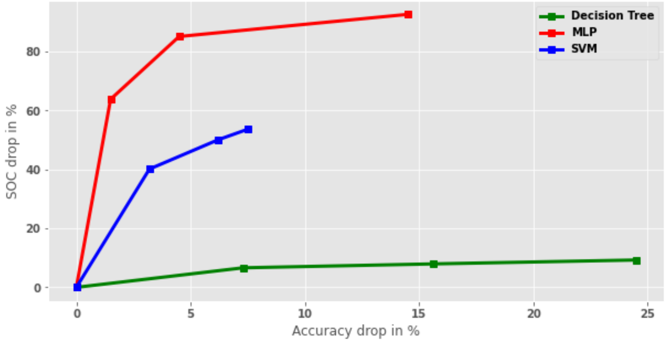


Figure 8: Comparison of Models in Accuracy and Interpretability on the MNIST dataset.

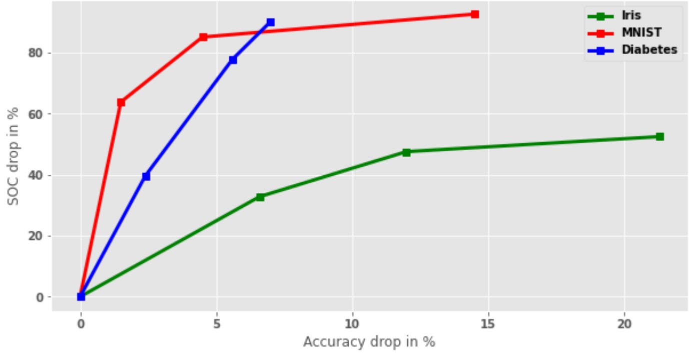


Figure 9: Trade-off between Accuracy and Interpretability in MLP estimator for classification.

Table 10: Comparison of models in terms of accuracy and interpretability (acc. is short for accuracy) for classification.

	DT acc., SOC	MLP acc., SOC	SVM acc., SOC
Iris	66.7%, 11 93.3%, 13 96.0%, 15 96.7%, 17	62.0%, 29 71.3%, 32 76.7%, 41 83.3%, 61	96.7%, 117 97.3%, 162 98.0%, 198 98.7%, 252
Mnist	54.5%, 137 63.4%, 139 71.7%, 141 79.0%, 151	80.4%, 794 90.4%, 1599 93.4%, 3879 94.9%, 10719	88.5%, 45220 89.8%, 48720 92.8%, 58380 96.0%, 97500
Diabetes	72.3%, 19 73.7%, 21 73.8%, 23 75.4%, 27	68.7%, 103 70.1%, 231 73.3%, 627 75.7%, 1038	75.7%, 5185 76.6%, 5338 77.1%,5712 - -

Table 10 is a more detailed version of Figure 8, and it shows significant increases of interpretability by allowing some drops in accuracy.

The second approach was to test one of the models on three datasets, for example MLP (Figure 9). The model performs similarly on all datasets and SOC could be improved at cost of lowering accuracy. For example, 64% increase of interpretability would require 1.5% reduction in accuracy on MNIST.

5. Conclusion

We introduced a methodology for trade-offs between interpretability and accuracy by inheriting the quantitative and model-agnostic metric - SOC. The LMT model, through its powerful but simple architecture (combination of linear regression and decision tree models), is the most interpretable estimator amongst the considered regression estimators; it has comparable accuracy to MLP in simple-medium datasets like Auto MPG and Servo.

The Decision Tree algorithm has the highest interpretability compared to other evaluated models due to its simplicity on classification task. It outperforms MLP on Iris and shows competitive results on Diabetes dataset. However, it has the lowest accuracy on MNIST with a large gap from other algorithms.

This paper demonstrates the tradeoff method between accuracy and interpretability using SOC metric. SOC is a model agnostic quantitative metric, hence it allows fair comparison between different types of estimators. In our experiments decreasing SOC leads to a simpler model with less memory requirement and faster inference speed. However, in general lower SOC may not always result in models with small memory requirement (i.e. replacing complex operation with simpler one) and faster inference speed (parallelizable model with high SOC can be faster than purely sequential model with low SOC on parallel hardwares).

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work was partly supported by the Nazarbayev Univer- sity (NU), Kazakhstan, under FDCRGP grant 021220FD0851.

References

[1] Z. Nazir, D. Kaldykhanov, K.K. Tolep, J.G. Park, "A Machine Learning Model Selection considering Tradeoffs between Accuracy and Interpretability," 2021 13th International Conference on Information Technology and Electrical Engineering, ICITEE 2021, 63–68, 2021, doi:10.1109/ICITEE53064.2021.9611872.

[2] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," Nature Machine Intelligence, **1**(5), 206–215, 2019, doi:10.1038/s42256-019-0048-x.

[3] J.R. Zech, M.A. Badgeley, M. Liu, A.B. Costa, J.J. Titano, E.K. Oermann, "Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study," PLoS Medicine, **15**(11), 1–17, 2018, doi:10.1371/journal.pmed.1002683.

[4] C. Molnar, G. Casalicchio, B. Bischl, "Interpretable Machine Learning – A Brief History, State-of-the-Art and Challenges," Communications in Computer and Information Science, **1323**(01), 417–431, 2020, doi:10.1007/978-3-030-65965-3_28.

[5] U. Johansson, C. Sönströd, U. Norinder, H. Boström, "Trade-off between accuracy and interpretability for predictive in silico modeling," Future Medicinal Chemistry, **3**(6), 647–663, 2011, doi:10.4155/fmc.11.23.

[6] T. Mori, N. Uchihira, "Balancing the trade-off between accuracy and interpretability in software defect prediction," Empirical Software Engineering, **24**, 779–825, 2019, doi:10.1007/s10664-018-9638-1.

[7] J.-G. Park, N. Dutt, S.-S. Lim, "An Interpretable Machine Learning Model Enhanced Integrated CPU-GPU DVFS Governor," ACM Trans. Embed. Comput. Syst., **20**(6), 2021, doi:10.1145/3470974.

[8] A. Adadi, M. Berrada, "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)," IEEE Access, **6**, 52138–52160,

2018, doi:10.1109/ACCESS.2018.2870052.

[9] M.A.H. Farquad, V. Ravi, S.B. Raju, “Support vector regression based hybrid rule extraction methods for forecasting,” Expert Systems with Applications, **37**(8), 5577–5589, 2010, doi:https://doi.org/10.1016/j.eswa.2010.02.055.

[10] F. Doshi-Velez, B. Kim, “Towards A Rigorous Science of Interpretable Machine Learning,” ArXiv E-Prints, arXiv:1702.08608, 2017.

[11] D. Slack, S.A. Friedler, C. Scheidegger, C.D. Roy, “Assessing the Local Interpretability of Machine Learning Models”, NeurIPS Workshop on Human-Centric Machine Learning, 2019, doi:10.48550/ARXIV.1902.03501.

[12] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., “Scikit-learn: Machine Learning in Python,” Journal of Machine Learning Research, **12**, 2012.

[13] E-handbook of statistical methods, NIST/SEMATECH, 2012, doi: https://doi.org/10.18434/M32189.

[14] R. O’Brien, “A Caution Regarding Rules of Thumb for Variance Inflation Factors,” Quality & Quantity, **41**, 673–690, 2007, doi:10.1007/s11135-006-9018-6.

[15] P. Cortez, A. Morais, “A Data Mining Approach to Predict Forest Fires using Meteorological Data,” 2007.

[16] J.R. Quinlan, “Combining Instance-Based and Model-Based Learning,” in Proceedings of the Tenth International Conference on International Conference on Machine Learning, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA: 236–243, 1993.

[17] J.R. Quinlan, “Learning With Continuous Classes,” World Scientific: 343–348, 1992.

[18] R.A. FISHER, “THE USE OF MULTIPLE MEASUREMENTS IN TAXONOMIC PROBLEMS,” Annals of Eugenics, **7**(2), 179–188, 1936, doi:https://doi.org/10.1111/j.1469-1809.1936.tb02137.x.

[19] L. Deng, “The MNIST Database of Handwritten Digit Images for Machine Learning Research [Best of the Web],” Signal Processing Magazine, IEEE, **29**, 141–142, 2012, doi:10.1109/MSP.2012.2211477.

[20] J. Smith, J. Everhart, W. Dickson, W. Knowler, R. Johannes, “Using the ADAP Learning Algorithm to Forcast the Onset of Diabetes Mellitus,” Proceedings - Annual Symposium on Computer Applications in Medical Care, **10**, 1988.

[21] L. Dillard, lmt.py, 2017, Link: https://gist.github.com/logandillard/lmt.py.

[22] P. Cortez, M.J. Embrechts, “Using sensitivity analysis and visualization techniques to open black box data mining models,” Information Sciences, **225**, 1–17, 2013, doi:https://doi.org/10.1016/j.ins.2012.10.039.

[24] P.E. Utgoff, ed., “Machine Learning, Proceedings of the Tenth International Conference, University of Massachusetts, Amherst, MA, USA, June 27-29, 1993,” Morgan Kaufmann, 1993, doi:10.1016/c2009-0-27798-1.

[24] A. Stanford-Moore, “Wildfire Burn Area Prediction,” 2019.

[25] Z. Hussain, H. Ibraheem, M. Aljanabi, A. Ali, M.A. Ismail, S. Kasim, T. Sutikno, “A new model for iris data set classification based on linear support vector machine parameter’s optimization,” International Journal of Electrical and Computer Engineering (IJECE), **10**, 1079, 2020, doi:10.11591/ijece.v10i1.pp1079-1084.

[26] D.C. Cireşan, U. Meier, L.M. Gambardella, J. Schmidhuber, “Deep, Big, Simple Neural Nets for Handwritten Digit Recognition,” Neural Computation, **22**(12), 3207–3220, 2010, doi:10.1162/NECO_a_00052.

[27] B. Chandra, V.P. Paul, “A Robust Algorithm for Classification Using Decision Trees,” in 2006 IEEE Conference on Cybernetics and Intelligent Systems, 1–5, 2006, doi:10.1109/ICCIS.2006.252336.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 7, No. 4, 79-91 (2022)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

On the Prediction of One-Year Ahead Energy Demand in Turkey using Metaheuristic Algorithms

Basharat Jamil[1,*], Lucía Serrano-Luján[1,2], José Manuel Colmenar[1]

[1]*Departamento de Ciencias de la Computación, Arquitectura de Computadores, Lenguajes y Sistemas Informáticos y Estadística e Investigación Operativa, Universidad Rey Juan Carlos, Móstoles-28933, Madrid, Spain*

[2]*Departamento de Electrónica, Tecnología de Computadoras y Proyectos, Universidad Politécnica de Cartagena, Campus Muralla del Mar. C/Doctor Fleming s/n, 30202 Cartagena, Spain*

ARTICLE INFO

Article history:
Received: 04 April, 2022
Accepted: 05 July, 2022
Online: 16 August, 2022

Keywords:
Energy Demand
Metaheuristics
Prediction
Evolutionary Algorithms
Turkey

ABSTRACT

Estimation of energy demand has important implications for economic and social stability leading to a more secure energy future. One-year-ahead energy demand estimation for Turkey has been proposed in this paper, using the metaheuristics method with GDP, the total population, and the quantities of imports and exports, as inputs variables. The records obtained from historical data were bifurcated into training and test datasets, where the training dataset is used by the algorithm in the process of generating models, while the test dataset was used to evaluate the performance of the algorithm. Here, two particular approaches have been proposed: Grammatical Evolution alone, and an ensemble of Grammatical Evolution with Differential Evolution. Under these four different forms are developed, viz, Grammatical Evolution with a recursive grammar (M1), an ensemble of Grammatical evolution executed on a linear grammar and Differential Evolution (M2), an ensemble of Grammatical evolution executed on a quadratic grammar and Differential Evolution (M3), and, Grammatical Evolution with a recursive grammar and Differential Evolution (M4). Moreover, the present approaches were also compared for estimation accuracy against the previously published DE models. It was substantiated that the M4 proposal exhibited the best performance towards estimation. It is therefore established that the current approach exhibits a better estimation capability (with RMSE of 2.2002), compared to the models previously available in the literature. M4 approach is then employed to predict the future energy demand using the same set of socio-economic inputs and the results demonstrated high prediction accuracy with an RMSE of 2.2278.

1. Introduction

Energy requirements of a country are met through proper planning, execution, and prediction based on the infrastructure and the availability of resources. It is no surprise that energy consumption has been rising aggressively across the globe with the increasing energy issues, while more attention must be focused on reducing environmental pollution, together with keeping up the economic growth [1,2]. Energy planning for the future is of utmost importance as it can have a significant impact on the economy of a country [3]. Broadly, energy demand has been mainly influenced by the price of energy and earnings, distributional effects, literacy of energy use among society, and the demand for energy in industry and transportation. Particularly the aspects that affect the demand for energy in the country are the climate of the region, urbanization rate [4], the sectorial energy consumption [5], use of technology [6], transition to renewable energy [7], and adherence and enforcement of environmental law and governmental policies [8].

The evolution of the energy sector and the uncertainties related to it have been a topic of interest to strategize energy secure future and layout policies for future energy demand. In this view, recently, a lot of focus has been realized on the modeling tools, frameworks, and assessment procedures that can provide more accurate insight into the future energy demand of a country [9]. Various approaches can be found in the literature dealing with the modeling energy demand in total of a country such as Turkey [10], China and India [11], Spain [12], etc. The recent ones involve the use of metaheuristics for estimating and predicting the energy

[*] Corresponding Author: Basharat Jamil, +34-631608231, basharat.jamil@urjc.es

https://dx.doi.org/10.25046/aj070411

demand for the future. We now look at various approaches presented in the literature for estimating and predicting country-wide energy demand.

In [11], the authors analyzed the performance of time series forecasting methods (using a grey model) to predict China's and India's future energy demand (1990-2016). The reported techniques were deemed to increase the prediction of the future energy demand of the two considered countries. Trend map, error measure, and fit method were used to analyze the accuracy and the mean absolute error of single-linear, hybrid-linear, and non-linear techniques were reported as 1.30-3.08%, 0.80-2.57%, and 2.06-2.19%, respectively. In [12], the authors presented a one-year-ahead estimation of energy demand based on historical data for the years 1980 to 2011 for 14 socio-economic indicators. Modified Harmony Search (HS) and an Extreme Learning Machine (ELM) were used and compared for prediction. Reported results by the proposed approaches made very accurate energy demand predictions with a mean absolute error of 3.21-4.63%. In a similar work [13], the authors analyzed the performance of the robust hybrid approach composed of a Basic Variable Neighbourhood Search (BVNS) algorithm with ELM. BVNS was employed to select the most suitable macroeconomic features from the large group of considered ones, whereas the ELM performed the energy demand predictions on the considered features. The proposed algorithm was reported to have the best Mean absolute error lower than 1.66% and an average Mean absolute error of 3.90%. In [14], the authors proposed a new framework for energy demand estimation by combining an adaptive Genetic Algorithm (GA) and a co-integration analysis for China. Model weights were optimized using Artificial Intelligence techniques (GA, ACO, and hybrid algorithms) together with co-integration analysis. It was reported that the proposed models have significantly better performance (Mean Absolute Percentage Error, of 8.68%). In [15], the authors proposed a hybrid algorithm formed by Bat Algorithm, Gaussian Perturbations, and Simulated Annealing Energy Demand (BAG-SA EDE) for China data. The analysis of the relationship between the energy demand and the input factors was carried out using a stationary test, co-integration test, and Granger causality test. It was reported that the quadratic model performs better prediction (mean absolute percentage error of 0.28%) of future energy demand than the multiple linear regression (mean absolute percentage error of 0.88%). Grammatical Evolution (GE) was proposed by [16] for developing new models using 14 macro-economic parameters for a year-ahead estimation of country-wide energy demand for Spain and France. The Differential Evolution (DE) algorithm was used to optimize each model's parameters. The proposed algorithms exhibited excellent accuracy (best Mean absolute error 1.9, and average Mean absolute error 3.33) for energy prediction. Bees Algorithm technique was suggested in [17] for estimating total energy demand in Iran based on population, GDP, import, and export data. Exponential and linear models were proposed for estimation and were also deployed to predict energy demand for up to the year 2030. It was concluded that the linear Bees Algorithm technique has the best prediction accuracy (relative error of 1.07%). In [18], the authors demonstrated a Mix-encoding Particle Swarm Optimization and Radial Basis Function (MPSO-RBF) based energy demand forecasting for China until the year 2020. The input data (for the years 1980-2009) consists of GDP, population, industry proportion, urbanization rate, and share of coal energy. It was reported that the proposed MPSO-RBF has four nodes of hidden layers and better accuracy in terms of the errors (Mean absolute percentage error of 0.78%) when compared to other Artificial Neural Networks (ANN)[19] based models. In [20], Particle Swarm Optimization and Genetic Algorithm optimal Energy Demand Estimating (PSO-GAEDE) model was proposed to improve the estimation efficiency for future projection in China. Linear, quadratic and exponential forms of models were proposed based on historical data from 1990 to 2009. The proposed PSO-GAEDE algorithm was reported to perform better than other algorithms with a Mean absolute percentage error of 0.54%. In [21], the authors deployed an ANN (with feed-forward multilayer perceptron model, coupled with an error back-propagation technique, FF-BP-ANN) to estimate the energy demand for Korea. Multiple linear regression models, exponential model, and ANN models [22] were analyzed and it was concluded that the ANN model outperforms the multiple linear regression models (linear and exponential) with accuracy in terms of Root Mean Squared Error, RMSE=5.7803. In [23], the authors forecasted the energy demand of China using a hierarchical Bayesian approach. Static and dynamic models were proposed with variable input parameters and forecasts were made for years up to 2030. It was reported that the hierarchical Bayesian approach has better performance (RMSE = 0.025) in model fitting than the fixed effects method (RMSE = 0.030). An analysis of energy structure and carbon emissions for China was performed in [24] using an optimized mixed data sampling model (ADL-MIDAS) model involving quarterly GDP, quarterly added value, and annual energy demand as the input variables. It was reported that compared to previous studies, the prediction error (RMSE) was in general below 0.1%, and the smallest error was reported to be 0.02%, thereby energy demand prediction was significantly improved. The forecast of energy demand for the Hunan province of China was presented in [25] between the years 2012 to 2030. Autoregressive integrated moving average and vector autoregressive models were employed for forecasting, and the resulting uncertainties analyzed using the Monte-Carlo method were reported to be under 15%. In [26], the authors employed a swarm intelligence-based Adaptive Firefly Algorithm (AFA) to improve the energy demand estimation in Turkey based on economic parameters. Linear and quadratic forms of models were proposed for estimation and historical data for Turkey. The proposed models were compared with Ant Colony Optimization (ACO) and Particle Swarm Optimization (PSO), each involving linear and quadratic models. It was concluded that the AFA-quadratic model resulted in the best accuracy with an accuracy of 99.24%. The determinants for energy demand in Turkey were discussed in [27]. ACO was used to optimize the energy demand problem. Three scenarios of energy demand growth were discussed and values of energy demand were determined for Turkey for 2006-2025. The quadratic equation optimized by ACO was suggested to be the most accurate with a deviation of 2.83%. Similar works were also reported by [28] and [29] for Turkey. Other authors used Differential Evolution (DE)[10], Artificial Algae Algorithm (AAA)[30], PSO [31], Ridge Regression (RR), and Partial Least Squares Regression (PLSR)[32] for estimating future energy demand in Turkey.

Based on previous research, it is elucidated that the long-term dependence of energy demand on economic growth factors certainly exists [33].

In this paper, a one-year-ahead energy demand prediction for Turkey is proposed involving a combination of evolutionary metaheuristic algorithms. These algorithms are based on Grammatical Evolution (GE) and Differential Evolution (DE), where GE has the capacity to develop flexible model forms based on the information provided in the grammar, while DE optimizes the model's parameters. Utilizing these two algorithms in combination results in better estimation accuracy and the predictions for energy demand can be made with greater reliability. The input variables selected are similar to the previous studies so that the comparison of algorithms is fair and the advantage of the present approach can be highlighted. These variables consist of the historical data of socio-economic parameters: GDP, population, import, and export quantities, which were used as inputs, while Energy demand was the target or output variable.

The novelty of the work is further justified by the comparison of algorithms and the grammars used. In the current scenario, two particular approaches have been proposed: Grammatical Evolution alone, and an ensemble of Grammatical Evolution with Differential evolution. Based on these algorithms, four different proposals were studied:

i. *M1*: Grammatical Evolution with a recursive grammar.

ii. *M2*: Ensemble of Grammatical Evolution with linear grammar and Differential Evolution.

iii. *M3*: Ensemble of Grammatical Evolution with quadratic grammar and Differential Evolution.

iv. *M4*: Grammatical Evolution with a recursive grammar and Differential Evolution.

The performance of these proposals is compared by assessing their estimation accuracy using RMSE as the objective function. The average error, R^2, absolute error, and relative error metrics were calculated to compare the performance of the proposals.

The present algorithm combination is also compared to the previous research works reported in literature involving Differential Evolution to validate their accuracy. Further, the application of these metaheuristics is utilized to predict the future energy demand for Turkey. One year ahead energy demand prediction is made using the combination of the previously described two algorithms (GE and DE). Thus, the approach with the combination of the Grammatical Evolution algorithm provided with a highly recursive grammar unified with the optimization strength of the Differential Evolution algorithm (M4) presents a unique and highly accurate approach to making energy demand predictions.

The rest of the paper is organized as follows. *Section 2* provides the outline of the methodology used consisting of the data obtained, problem definition, and the objective function. *Section 3* discusses the results obtained dealing with the estimation of energy demand and prediction of the year ahead of future energy demand for Turkey. Finally, *Section 4* presents the conclusions of the study.

2. Research Methodology and Algorithms

In this section, the proposed methodology is described. Firstly, the used dataset is described as well as the curation, training, and test selection processes. Then, the algorithmic proposals and quality metrics are detailed.

2.1. Data

The required data for the study consisted of the historical energy demand as well as the corresponding data for socio-economic parameters for Turkey, obtained from literature [10]. This data was compiled from MENR, the energy reports, and some of the previously reported studies in the literature [32]. The data consisted of Turkey's GDP ($\$10^9$), Population ($10^6$), and the amounts of Import ($\$10^9$) and Export ($\$10^9$) for the period from 1979 to 2011 (as annual values). These four elements are generally considered to have the biggest impact on the energy demand. Therefore these are considered as the input variables for the present study as well. The target variable for estimation (as well as prediction), is the energy demand (MTOE). Data obtained can be found in the appendix (Table A.1). Notice that the input variables are also named *X1* (GDP), *X2* (Population), *X3* (Quantity of Imports) to *X4* (Quantity of Exports) while the energy demand is denoted as *E*.

From the historical dataset of the energy demand for Turkey, an interesting trend can be deduced for the period studied. The energy demand growth rate ranges from a minimum of -6.34% (for the year 2001) to a maximum of 10.38% (for the year 1987) with an average growth rate of 4.26%. While the input parameters growth rate are: GDP (Average 8.41%, minimum -27.00% and maximum 39.81%), population (Average 1.57, minimum -3.26 and maximum 2.95), import (Average 14.81%, minimum -30.22% and maximum 56.02%), and export quantities (Average 14.58%, minimum -22.64% and maximum 61.51%). It will be, therefore, interesting to find out how these socio-economic parameters correlate to energy demand and further their impact on its future.

In order to avoid the influence on a model of the particular amount of each variable (due to the units used in each one of them), all the data have been normalized by the maximum value of each corresponding parameter. This way, every column has been transformed into the values that correspond to the ratio with the maximum value of the data:

$$\left(X_{i,n} = {}^{X_{i,j}}\!\big/_{X_{i,max}}\right) \tag{1}$$

where i = 1, 2, 3, 4. j = 1, 2…n, n being the number of years, and *max* being the maximum value of the i^{th} series.

Therefore, the maximum value of a given column is 1, while the rest of the values will be lower. This normalization is also done for the output variable. The data which spans 33 years were divided into two datasets i.e., *training* and *test*, where the training dataset consists of data for 17 years and the test dataset consists of the dataset for the remaining 16 years. The selection for each dataset was made randomly.

2.2. Problem definition and objective function

The search for a mathematical expression that captures the behavior of a target variable can be tackled as an optimization problem. In particular, this approach aims to find the expression with the minimum error, which in this case, is modeling the energy demand.

To assess the quality of a model, the root mean squared error (RMSE) metric, as shown in Equation (2), is used as the objective function.

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n}\left(E_{est,i} - E_{act,i}\right)^2\right]^{\frac{1}{2}} \qquad (2)$$

where, $E_{act,i}$ and $E_{est,i}$ in Equation (2) are the actual and estimated values of energy demand.

While other statistics viz, average error (AE), coefficient of determination (R^2), absolute error (ABS), and relative error (RE) were also used to evaluate the accuracy as follows:

$$AE = \frac{1}{n}\sum_{i=1}^{n}(E_{est,i} - E_{act,i}) \qquad (3)$$

$$R^2 = \frac{\sum_{i=1}^{n}(E_{est,i} - E_{est,avg})(E_{act,i} - E_{act,avg})}{\sqrt{\sum_{i=1}^{n}(E_{est,i} - E_{est,avg})^2 \sum_{i=1}^{n}\left(E_{act,i} - E_{act,avg}\right)^2}} \qquad (4)$$

$$ABS = \frac{1}{n}\sum_{i=1}^{n}\left|E_{est,i} - E_{act,i}\right| \qquad (5)$$

$$RE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|E_{est,i} - E_{act,i}\right|}{E_{act,i}} \qquad (6)$$

2.3. Algorithmic Methods

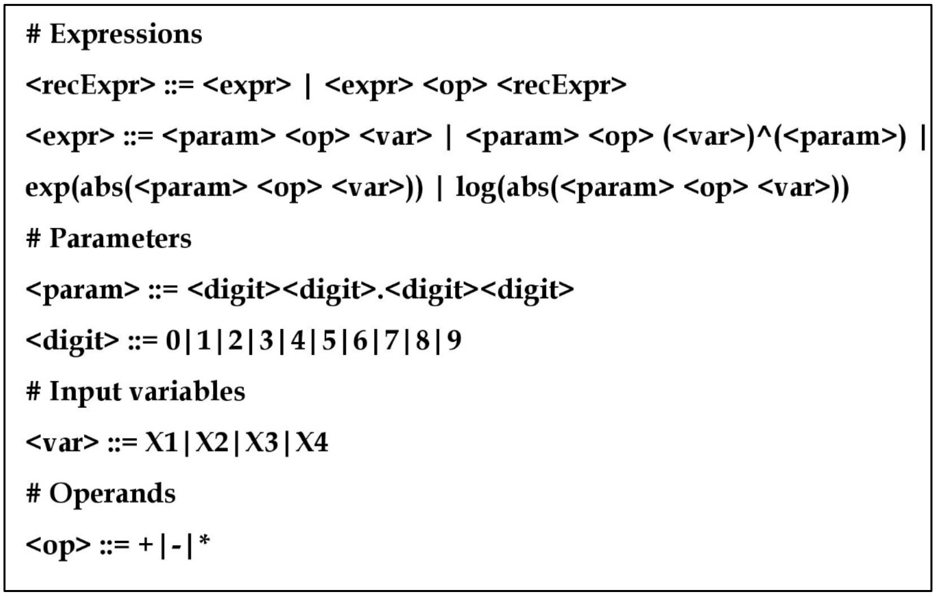
In this work, the deployment of Grammatical Evolution (GE) and Differential Evolution (DE) is proposed for the estimation of energy demand. GE is a metaheuristic algorithm belonging to the family of Genetic Programming whose main advantage is the ability to direct the search of the algorithm using grammar [34]. This way, researchers may introduce knowledge about the problem into the grammar, reducing the search space. On the other hand, DE is a metaheuristic algorithm better suited for problems where some parameter values need to be found [35].

Despite the combination of GE and DE has been proven to be effective in the past for this problem [16],[36,37], one of the main contributions of this work is the comparison with different combinations of GE and DE determined by different grammars using a reduced number of input variables. The tool used in this work is WebGE, which is an open-source optimization tool that implements both GE and DE algorithms described before. We refer the reader to [38] for further information.

In particular, four different approaches are proposed in this work. The first one is the use of recursive grammar in GE with no DE intervention (M1). Figure 1 shows the grammar applied to this approach. Elements on the left-hand side of each "::=" symbol are non-terminal values, which must be decoded using any of the productions on the right-hand side, separated by "|" symbols. As seen in the figure, this grammar can generate mathematical expressions using the four input variables (X1 to X4), constant values, addition, subtraction, and product arithmetic operators, and the exponential, power, and logarithmic functions. It is important to note that the <param> non-terminal symbol is devoted to generating constant numbers in the range [0.00,99.99].

The second approach, named M2, is directed by the grammar shown in Figure 2. The main idea is to produce expressions where arithmetic combinations of a parameter multiplied by an input variable are generated. In this case, DE is in charge of finding out the best parameter values (w_i) for each expression generated by GE. Notice that the <digit> rule is not needed since the parameters, represented by *wi*, are found by DE.

The third approach is similar to the previous one but includes the use of the quadratic values of input variables. This approach is termed M3 and is guided by the grammar shown in Figure 3. Again, DE is in charge of finding out the values of the parameters.



```
# Expressions
<recExpr> ::= <expr> | <expr> <op> <recExpr>
<expr> ::= <param> <op> <var> | <param> <op> (<var>)^(<param>) |
exp(abs(<param> <op> <var>)) | log(abs(<param> <op> <var>))
# Parameters
<param> ::= <digit><digit>.<digit><digit>
<digit> ::= 0|1|2|3|4|5|6|7|8|9
# Input variables
<var> ::= X1|X2|X3|X4
# Operands
<op> ::= +|-|*
```

Figure 1: GE executed on a recursive grammar (M1)

```
# Expressions
<expr> ::= (<param> <op> <var>) | (<param> * <var>) <op> (<param> * <var>) |
(<param> * <var>) <op> (<param> * <var>) <op> (<param> <op> <var>) |
(<param> * <var>) <op> (<param> * <var>) <op> (<param> <op> <var>) <op>
(<param> <op> <var>)
# Parameters
<param> ::= w1|w2|w3|w4
# Input variables
<var> ::= X1|X2|X3|X4
# Operands
<op> ::= +|-|*
```

Figure 2: Ensemble of GE executed on a linear grammar and DE (M2)

```
# Expressions
<expr> ::= (<param> <op> <var>) | (<param> * <var>) <op> (<param> * <var>) |
(<param> * <var>) <op> (<param> * <var>) <op> (<param> <op> <var>) |
(<param> * <var>) <op> (<param> * <var>) <op> (<param> <op> <var>) <op>
(<param> <op> <var>) | (<param> * <var> * <var>) | (<param> * <var> * <var>)
<op> (<param> * <var> * <var>) | (<param> * <var> * <var>) <op> (<param> *
<var> * <var>) <op> (<param> * <var> * <var>) | (<param> * <var> * <var>) <op>
(<param> * <var> * <var>) <op> (<param> * <var> * <var>) <op> (<param> *
<var> * <var>) | (<param> * <var> * <var>) <op> (<param> * <var> * <var>) <op>
(<param> * <var> * <var>) <op> (<param> * <var> * <var>) <op> (<param> *
<var> * <var>) | (<param> * <var> * <var>) <op> (<param> * <var> * <var>) <op>
(<param> * <var> * <var>) <op> (<param> * <var> * <var>) <op> (<param> *
<var> * <var>) <op> (<param> * <var> * <var>)
# Parameters
<param> ::= w1|w2|w3|w4
# Input variables
<var> ::= X1|X2|X3|X4
# Operands
<op> ::= +|-|*
```

Figure 3: Ensemble of GE executed on quadratic grammar and DE (M3)

Finally, the last approach is the ensemble of GE and DE using recursive grammar. In this case, first, the GE approach is taken, but allowing DE to look for the parameter values. Figure 4 shows the proposed grammar for this approach.

```
# Expressions
<recExpr> ::= <expr> | <expr> <op> <recExpr>
<expr> ::= <param> <op> <var> | <param> <op> (<var>)^(<param>) |
exp(abs(<param> <op> <var>)) | log(abs(<param> <op> <var>))
# Parameters
<param> ::= w1|w2|w3|w4
# Input variables
<var> ::= X1|X2|X3|X4
# Operands
<op> ::= +|-|*
```

Figure 4: Ensemble of GE executed a recursive grammar and DE (M4)

3. Results and discussion

The experimental experience consisted of the execution of the proposed algorithms over the training dataset under the four configurations presented in the previous section. In particular, 20 runs were executed for each one of the four approaches over the training dataset, obtaining a total amount of 80 different models. Table 1 shows the values of the parameters for the GE and DE algorithms, selected after preliminary experimentation.

Table 1: Details of the experiments

GE Parameters	
Generations	50
Crossover Probability	0.65
Population	20
Mutation Probability	0.1
Max wraps	3
Number of Codons	100
Tournament	2
Number of Runs	20
DE Parameters	
Recombination Factor	0.88
Mutation Factor	0.47
Population Size	20

The results presented in the following sections first describe the estimation followed by the prediction. On one hand, for the estimation problem, the values of the input variables of a given year are used to estimate the energy demand of the same year. On the other hand for prediction, the models have been adapted for the prediction of future energy demand, where the values of the input variables of a given year are used to predict the energy demand of the following year.

3.1. Estimation of Energy Demand

The energy demand estimations were obtained for all the years from 1979 to 2011 using the models generated by the proposed algorithms. In particular, for each one of the four approaches developed under the present study (M1, M2, M3, and M4), an average estimation using the 20 generated models has been evaluated for the target period.

The results of the estimation for the whole dataset are presented in Table 2 considering the average prediction of the 20 models obtained for each one of the four proposed approaches. The previously reported [10] two models Linear (DEL) and Quadratic (DEQ) are also included in this comparison.

Table 2: Estimations of the previously reported differential algorithms based on linear and quadratic [10] with the four model forms of the present approach i.e., M1, M2, M3 and M4

Year	Actual Energy (MTOE)	Linear (DEL) [10]	Quadratic (DEQ) [10]	M1	M2	M3	M4
1979	30.71	32.28	34.48	32.07	30.06	31.88	32.11
1980	31.97	30.89	31.26	29.79	30.71	30.99	31.11
1981	32.05	32.52	33.33	32.09	32.99	32.54	32.67
1982	34.39	34.15	34.84	34.50	34.48	34.01	34.12
1983	35.7	36.53	35.85	36.95	36.29	36.24	36.28

1984	37.43	38.73	37.01	39.50	38.84	38.39	38.39
1985	39.4	40.95	40.10	42.11	40.89	40.43	40.45
1986	42.47	43.27	42.78	44.48	42.27	42.50	42.52
1987	46.88	45.30	45.28	46.89	45.47	44.60	44.63
1988	47.91	46.89	47.68	49.38	47.12	46.02	46.09
1989	50.71	49.76	50.43	51.91	49.40	48.75	48.82
1990	52.98	54.02	53.08	54.55	53.86	53.22	53.25
1991	54.27	55.33	54.49	56.95	54.77	54.25	54.34
1992	56.68	57.52	56.42	59.30	56.85	56.41	56.49
1993	60.26	61.79	61.00	61.71	60.53	61.00	60.92
1994	59.12	60.08	60.48	64.20	59.63	58.79	58.87
1995	63.68	65.28	65.29	66.76	65.26	64.80	64.70
1996	**69.86**	**69.71**	**70.00**	**69.41**	**68.90**	**69.69**	**69.41**
1997	**73.78**	**72.31**	**73.17**	**72.13**	**71.64**	**72.54**	**72.24**
1998	**74.71**	**73.30**	**74.92**	**74.95**	**72.50**	**73.05**	**72.92**
1999	**76.77**	**74.18**	**75.47**	**78.44**	**72.74**	**73.30**	**73.29**
2000	**80.5**	**80.71**	**81.04**	**80.89**	**77.40**	**80.84**	**80.43**
2001	**75.4**	**75.71**	**74.60**	**83.28**	**75.13**	**74.81**	**74.78**
2002	**78.33**	**79.13**	**80.32**	**85.65**	**79.05**	**78.89**	**78.86**
2003	**83.84**	**82.36**	**84.01**	**88.05**	**84.57**	**83.47**	**83.54**
2004	**87.82**	**87.19**	**88.04**	**90.45**	**91.57**	**90.62**	**90.68**
2005	**91.58**	**93.10**	**92.84**	**95.31**	**97.36**	**97.83**	**97.90**
2006	99.59	96.25	56.00	95.31	101.70	102.69	102.57
2007	107.63	92.76	4.97	88.98	103.98	102.43	102.66
2008	106.27	94.19	-136.98	90.40	109.84	106.49	107.15
2009	106.14	90.17	-495.07	96.02	103.81	96.02	95.47
2010	109.27	103.09	-2.75	99.43	110.96	113.23	113.12
2011	114.48	114.50	64.97	100.12	118.99	129.30	128.27

Highlighted values represent the period of consideration in the study conducted by Beskirli et a [10]

Notice that the previous comparison of algorithms presented in the state of the art was made for a small range of years, corresponding to the period from 1996 to 2005 as highlighted in Table 2. However, in this work, the comparison is made for the entire range of the years (1979-2011) as previously described. Correspondingly, Table 3 compares the estimation performance over the test dataset of the four proposed approaches with the two methods found in the state of the art: a linear model (DEL) and a quadratic model (DEQ). For all of them, 5 error metrics have been obtained: root mean squared error (RMSE), Average Error (AE), R^2, Absolute Error (ABS), and Relative Error (RE). Notice that the depicted values correspond to the average value of the 20 models in the case of the four proposed approaches.

Table 3: Error metrics for the compared models over the test data. Best results are depicted in bold fonts.

Year	Linear (DEL) [10]	Quadratic (DEQ) [10]	M1	M2	M3	M4
RMSE	4.6403	116.5278	6.1913	3.7255	3.6481	**2.2002**
AE	2.4685	35.5655	4.0700	2.0997	2.1008	**1.6850**
R^2	0.9771	0.9812	0.9504	0.9807	0.9812	**0.9931**
ABS	81.4605	1173.6609	134.1800	69.2909	69.3275	**55.6063**
RE	0.0316	0.3408	0.0520	0.0269	0.0270	**0.0238**

As seen in the table, all the proposed approaches except M1 (GE with no DE) obtain superior results than the algorithms from the state of the art in terms of the error metrics. Further as observed, the M4 proposal exhibits the top performance among the proposed algorithms with the lowest values of RMSE (2.2002), Average Error (1.6850), Absolute Error (55.6063), Relative Error (0.0238), and the highest value of R^2 (0.9931). This can be attributed to the fact that the flexible model structure of the recursive grammar combined with the DE efficiency can obtain better parameter values, resulting in a much better performance in comparison to those approaches where the model structure is fixed. Therefore, the ensemble of GE and DE with a recursive grammar (M4) is the most appropriate algorithm for the estimation of energy demand. Hence, this appropriate combination of metaheuristic algorithms (M4) will be used to predict the energy demand for the future.

3.2. Prediction of Future Energy Demand

The confirmation of the performance improvement achieved from the proposed methods has been made in the previous section. It is established that Model M4 which represents an ensemble of a recursive GE combined with DE has the most accurate outcomes when applied to energy demand problems.

We now move on to the problem of predicting the year-ahead energy demand. To this aim, the set of input parameters remains the same as they were previously discussed. However, the algorithm took the previous year's data as input variables to predict the energy demand for the current year. This way a year ahead energy demand approach is established. In a similar fashion, the training and test datasets were separated, where the training dataset was used to train the algorithms while the test dataset was used as independent data to test the performance of the algorithms.

The experimentation then followed the same pattern: 20 runs were executed for the GE with a recursive grammar together with DE combination (with the same properties as previously discussed in Table 1) over the training dataset which produced the 20 models (which can be found in Appendix Table A.2). The results of the performance of algorithms in predicting the future energy demand are depicted in figures 5 (a) and (b) for training and test dataset, respectively, as an average of the 20 runs executed for energy prediction. It must be again noted here that since the input to the algorithm is normalized the output that is received is also in the normalized form. This can also be observed in figures 5(a) and (b).

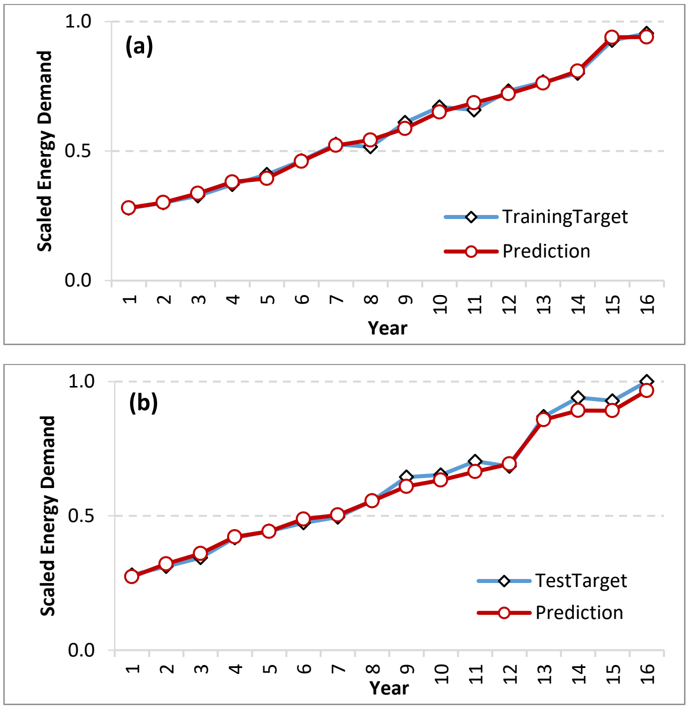


Figure 5: Predicted and actual energy demand (normalized) on (a) training and (b) test datasets. The X-axis denotes the ordinal number of the year in the dataset.

From the figure, it is observed that the predictions match the data very precisely and the errors are small. Table 4 shows the error metrics obtained for the prediction of a year ahead energy demand with M4. It is observed that the algorithm has high accuracy and performs well in terms of prediction accuracy as justified by the small errors produced on the test dataset. Also, since the training and test datasets were independent the issue of overfitting a model on the data is avoided, resulting in a more flexible and accurate model. The predictions of energy demand are presented below in Figure 6 with actual MTOE units for the complete dataset. As seen, the proposed method shows high accuracy on a one-year ahead prediction.

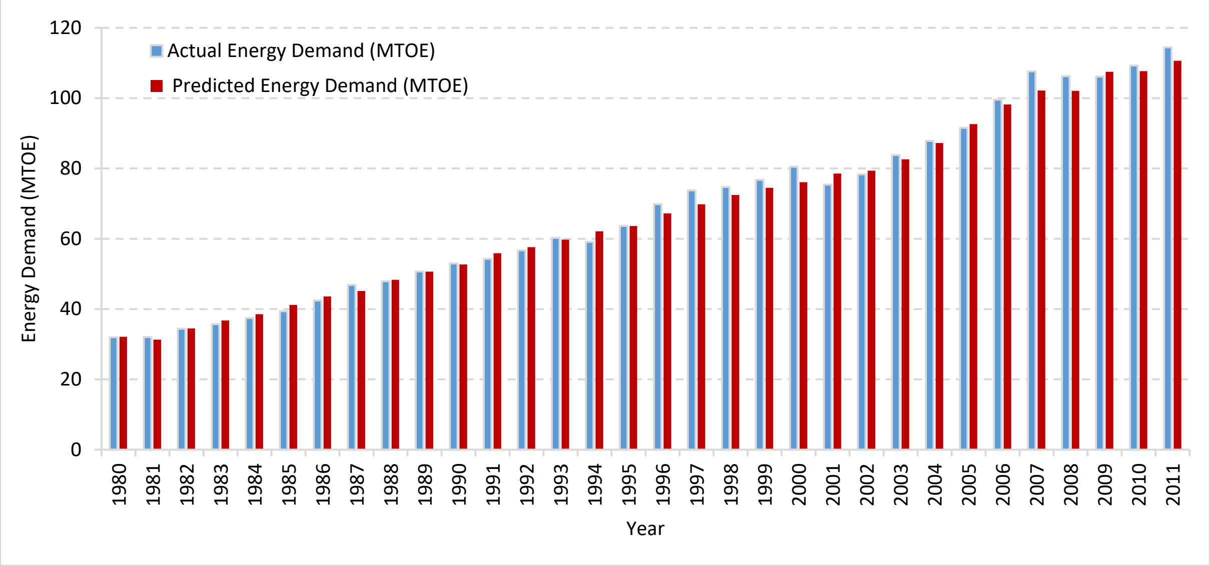


Figure 6: Evolution and comparison of actual and predicted energy demand (using the average of the 20 experimental runs)

The error computation is also made based on Figure. 6 and the obtained results were the following: RMSE=2.2278, AE=-0.6099, R^2= 0.9970, ABS=1.7263, and RE =0.0244. As seen, they are similar to the results with separated training and test data.

Table 4: Errors between predicted values from M4 and actual data for energy demand (normalized)

Error	Training	Test
RMSE	0.0146	0.0233
AE	0.0001	0.0107
R^2	0.9952	0.9959
ABS	0.1917	0.2908
RE	0.0210	0.0279

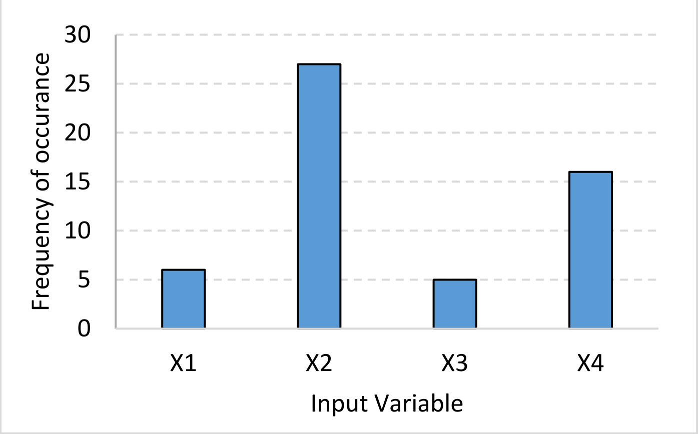


Figure 7: Histogram obtained from 20 runs of GE-DE

The details of prediction models are provided in the appendix (Table A.2). On further analysis of the obtained models, the number of terms appearing in the expressions ranges from 2 to 5, where some of the input parameters appear several times in the different model as well as different expressions. This leads us to find out the most influential parameter among the set of input parameters. To this aim, we calculate the frequency of occurrence of the four input parameters in the 20 models. The occurrence of the feature or the input variable in the equations is depicted in Figure 7. It can be observed that X2 (Population) appears the maximum number of times in the equations, followed by the X4 (quantity of export), the input X1 (GDP), and lastly the X3 (quantity of import). This means that energy demand primarily grows in correlation to the population of the country which in fact is imperative since the larger the population, the more is energy demand.

4. Conclusions

Country-wide energy demand predictions have been made considering socio-economic parameters based on metaheuristic algorithms for Turkey. Four methods were considered initially for the estimation of energy demand: (M1) Grammatical Evolution with a recursive grammar alone, (M2) Grammatical Evolution with a linear grammar together with Differential Evolution, (M3) Grammatical Evolution with a quadratic grammar together with Differential Evolution, and finally, (M4) Grammatical Evolution with a recursive grammar combined with Differential Evolution. The ensemble of GE with a recursive grammar and DE together (M4) performed very well in comparison to others with the least value of objective function RMSE=2.2002. Further, this M4 proposal was deployed to predict a year-ahead energy demand for Turkey. The algorithm again performed very well obtaining the predictions for a year ahead, and the results of prediction were deemed to be exceptionally well with the objective function to be RMSE=2.2278. Therefore, it is recommended to use the combination of Grammatical Evolution with recursive grammar to provide for more flexible model forms. Further, the use of Differential Evolution together with the recursive algorithm of GE provides optimization of the model weight leading to accurate predictions of the energy demand for a country. It was also inferred that population greatly affects the structure of the prediction models and thus is a significant parameter for energy demand models. It is suggested that for future studies long-term causal relationships are established along with the elasticities of algorithms. This will be significant and reflect a more certain behavior to predict the energy demand. Further, it is advised to investigate the behavior of algorithms with more input parameters to elucidate influential and non-influential parameters for energy demand estimation/prediction. This can also implicate the direct and indirect relationship between the input parameters and the energy demand.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The work reported under the manuscript has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 754382. The authors would like to thank Prof. Abraham Duarte for extending his support towards the research work carried out in the article.

References

[1] S. Katircioglu, C. Köksal, S. Katircioglu, "The role of financial systems in energy demand: A comparison of developed and developing countries," Heliyon, **7**(6), e07323, 2021, doi:https://doi.org/10.1016/j.heliyon.2021.e07323.

[2] S. Di Leo, P. Caramuta, P. Curci, C. Cosmi, "Regression analysis for energy demand projection: An application to TIMES-Basilicata and TIMES-Italy energy models," Energy, **196**, 117058, 2020, doi:https://doi.org/10.1016/j.energy.2020.117058.

[3] C. Huang, Z. Zhang, N. Li, Y. Liu, X. Chen, F. Liu, "Estimating economic impacts from future energy demand changes due to climate change and economic development in China," Journal of Cleaner Production, **311**, 127576, 2021, doi:https://doi.org/10.1016/j.jclepro.2021.127576.

[4] Y. Yu, N. Zhang, J.D. Kim, "Impact of urbanization on energy demand: An empirical study of the Yangtze River Economic Belt in China," Energy Policy, **139**, 111354, 2020, doi:https://doi.org/10.1016/j.enpol.2020.111354.

[5] S. Qiu, T. Lei, J. Wu, S. Bi, "Energy demand and supply planning of China through 2060," Energy, **234**, 121193, 2021, doi:https://doi.org/10.1016/j.energy.2021.121193.

[6] J. Huang, H. Zhang, W. Peng, C. Hu, "Impact of energy technology and structural change on energy demand in China," Science of The Total Environment, **760**, 143345, 2021, doi:https://doi.org/10.1016/j.scitotenv.2020.143345.

[7] K. Oshiro, S. Fujimori, Y. Ochi, T. Ehara, "Enabling energy system transition toward decarbonization in Japan through energy service demand reduction," Energy, **227**, 120464, 2021, doi:https://doi.org/10.1016/j.energy.2021.120464.

[8] P. Späth, H. Rohracher, "Local Demonstrations for Global Transitions—Dynamics across Governance Levels Fostering Socio-Technical Regime Change Towards Sustainability," European Planning Studies, **20**(3), 461–479, 2012, doi:10.1080/09654313.2012.651800.

[9] M.A. Islam, H.S. Che, M. Hasanuzzaman, N.A. Rahim, Chapter 5 - Energy demand forecasting, Academic Press: 105–123, 2020, doi:https://doi.org/10.1016/B978-0-12-814645-3.00005-5.

[10] M. BESKIRLI, H. HAKLI, H. KODAZ, "The energy demand estimation for Turkey using differential evolution algorithm," Sādhanā, **42**(10), 1705–1715, 2017, doi:10.1007/s12046-017-0724-7.

[11] Q. Wang, S. Li, R. Li, "Forecasting energy demand in China and India: Using single-linear, hybrid-linear, and non-linear time series forecast techniques," Energy, **161**, 821–831, 2018, doi:https://doi.org/10.1016/j.energy.2018.07.168.

[12] S. Salcedo-Sanz, J. Muñoz-Bulnes, J.A. Portilla-Figueras, J. Del Ser, "One-year-ahead energy demand estimation from macroeconomic variables using computational intelligence algorithms," Energy Conversion and Management, **99**, 62–71, 2015, doi:https://doi.org/10.1016/j.enconman.2015.03.109.

[13] J. Sánchez-Oro, A. Duarte, S. Salcedo-Sanz, "Robust total energy demand estimation with a hybrid Variable Neighborhood Search – Extreme Learning Machine algorithm," Energy Conversion and Management, **123**, 445–452, 2016, doi:https://doi.org/10.1016/j.enconman.2016.06.050.

[14] J. Huang, Y. Tang, S. Chen, "Energy Demand Forecasting: Combining Cointegration Analysis and Artificial Intelligence Algorithm," Mathematical Problems in Engineering, **2018**(5194810), 1–13, 2018, doi:10.1155/2018/5194810.

[15] Q. Wu, C. Peng, "A hybrid BAG-SA optimal approach to estimate energy demand of China," Energy, **120**, 985–995, 2017, doi:https://doi.org/10.1016/j.energy.2016.12.002.

[16] J.M. Colmenar, J.I. Hidalgo, S. Salcedo-Sanz, "Automatic generation of models for energy demand estimation using Grammatical Evolution," Energy, **164**, 183–193, 2018, doi:https://doi.org/10.1016/j.energy.2018.08.199.

[17] M.A. Behrang, E. Assareh, M.R. Assari, A. Ghanbarzadeh, "Total Energy Demand Estimation in Iran Using Bees Algorithm," Energy Sources, Part B: Economics, Planning, and Policy, **6**(3), 294–303, 2011, doi:10.1080/15567240903502594.

[18] S. Yu, Y.-M. Wei, K. Wang, "China's primary energy demands in 2020: Predictions from an MPSO–RBF estimation model," Energy Conversion and Management, **61**, 59–66, 2012, doi:https://doi.org/10.1016/j.enconman.2012.03.016.

[19] X. Yin, Q. Zhang, H. Wang, Z. Ding, "RBFNN-Based Minimum Entropy Filtering for a Class of Stochastic Nonlinear Systems," IEEE Transactions on Automatic Control, **65**(1), 376–381, 2020, doi:10.1109/TAC.2019.2914257.

[20] S. Yu, Y.-M. Wei, K. Wang, "A PSO–GA optimal model to estimate primary energy demand of China," Energy Policy, **42**, 329–340, 2012, doi:https://doi.org/10.1016/j.enpol.2011.11.090.

[21] Z.W. Geem, W.E. Roper, "Energy demand estimation of South Korea using artificial neural network," Energy Policy, **37**(10), 4049–4054, 2009, doi:https://doi.org/10.1016/j.enpol.2009.04.049.

[22] J. Liu, Y. Liu, Q. Zhang, "A weight initialization method based on neural network with asymmetric activation function," Neurocomputing, **483**, 171–182, 2022, doi:https://doi.org/10.1016/j.neucom.2022.01.088.

[23] X.-C. Yuan, X. Sun, W. Zhao, Z. Mi, B. Wang, Y.-M. Wei, "Forecasting China's regional energy demand by 2030: A Bayesian approach," Resources, Conservation and Recycling, **127**, 85–95, 2017, doi:https://doi.org/10.1016/j.resconrec.2017.08.016.

[24] Y. He, B. Lin, "Forecasting China's total energy demand and its structure using ADL-MIDAS model," Energy, **151**, 420–429, 2018, doi:https://doi.org/10.1016/j.energy.2018.03.067.

[25] R. Chen, Z. Rao, G. Liu, Y. Chen, S. Liao, "The long-term forecast of energy demand and uncertainty evaluation with limited data for energy-imported cities in China: a case study in Hunan," Energy Procedia, **160**, 396–403, 2019, doi:https://doi.org/10.1016/j.egypro.2019.02.173.

[26] H. Wang, Z. Chen, W. Wang, Z. Wu, K. Wu, W. Li, "Improving Energy Demand Estimation Using an Adaptive Firefly Algorithm BT - Computational Intelligence and Intelligent Systems," in: Li, K., Li, W., Chen, Z., and Liu, Y., eds., Springer Singapore, Singapore: 171–181, 2018.

[27] M. Duran Toksarı, "Ant colony optimization approach to estimate energy demand of Turkey," Energy Policy, **35**(8), 3984–3990, 2007, doi:https://doi.org/10.1016/j.enpol.2007.01.028.

[28] A. Ünler, "Improvement of energy demand forecasts using swarm intelligence: The case of Turkey with projections to 2025," Energy Policy, **36**(6), 1937–1944, 2008, doi:https://doi.org/10.1016/j.enpol.2008.02.018.

[29] O. ERSEL CANYURT, H. CEYLAN, H. KEMAL OZTURK, A. HEPBASLI, "Energy Demand Estimation Based on Two-Different Genetic Algorithm Approaches," Energy Sources, **26**(14), 1313–1320, 2004, doi:10.1080/00908310490441610.

[30] A. Beşkirli, M. Beşkirli, H. Haklı, H. Uğuz, "Comparing Energy Demand Estimation Using Artificial Algae Algorithm : The Case of Turkey," Journal of Clean Energy Technologies, **6**(4), 349–352, 2018, doi:10.18178/jocet.2018.6.4.487.

[31] T. Paksoy, G. Weber, "Particle Swarm Optimization Approach for Estimation of Energy Demand of Turkey," Global Journal of Technology & Optimization, **3**(June), 1–9, 2012.

[32] Y.M. Bulut, Z. Yildiz, "Comparing energy demand estimation using various statistical methods: The case of Turkey," Gazi University Journal of Science, **29**(2), 237–244, 2016.

[33] A. Löschel, S. Managi, "Recent Advances in Energy Demand Analysis—Insights for Industry and Households," Resource and Energy Economics, **56**, 1–5, 2019, doi:https://doi.org/10.1016/j.reseneeco.2019.04.001.

[34] M. O'Neill, C. Ryan, "Grammatical evolution," IEEE Transactions on

Evolutionary Computation, **5**(4), 349–358, 2001, doi:10.1109/4235.942529.

[35] R. Storn, K. Price, "Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces," Journal of Global Optimization, **11**(4), 341–359, 1997, doi:10.1023/A:1008202821328.

[36] N. Lourenço, J.M. Colmenar, J.I. Hidalgo, S. Salcedo-Sanz, "Evolving energy demand estimation models over macroeconomic indicators," GECCO 2020 - Proceedings of the 2020 Genetic and Evolutionary Computation Conference, 1143–1149, 2020, doi:10.1145/3377930.3390153.

[37] B. Jamil, L. Serrano-Luján, J.M. Colmenar, "Modelling energy consumption in Spain with metaheuristic methods," in 2021 6th International Conference on Smart and Sustainable Technologies (SpliTech), 1–3, 2021, doi:10.23919/SpliTech52315.2021.9566391.

[38] J.M. Colmenar, R. Martín-Santamaría, J.I. Hidalgo, "WebGE: An Open-Source Tool for Symbolic Regression Using Grammatical Evolution BT - Applications of Evolutionary Computation," in: Jiménez Laredo, J. L., Hidalgo, J. I., and Babaagba, K. O., eds., Springer International Publishing, Cham: 269–282, 2022.

Appendix

Table A. 1 provides the historical data on energy demand and the socio-economic parameters used in the study. Table A.2 gives the models generated by the algorithm.

Table A.1: Data of Turkey's actual energy demand, GDP, population, imports, and exports [10]

Year	Energy Demand (MTOE)	GDP ($\$10^9$)	Population (10^6)	Import ($\$10^9$)	Export ($\$10^9$)
	E	X1	X2	X3	X4
1979	30.71	82	45.53	5.07	2.26
1980	31.97	68	44.44	7.91	2.91
1981	32.05	72	45.54	8.93	4.70
1982	34.39	64	46.69	8.84	5.75
1983	35.70	60	47.86	9.24	5.73
1984	37.43	59	49.07	10.76	7.13
1985	39.40	67	50.31	11.34	7.95
1986	42.47	75	51.43	11.10	7.46
1987	46.88	86	52.56	14.16	10.19
1988	47.91	90	53.72	14.34	11.66
1989	50.71	108	54.89	15.79	11.62
1990	52.98	151	56.10	22.30	12.96
1991	54.27	150	57.19	21.05	13.59
1992	56.68	158	58.25	22.87	14.72
1993	60.26	179	59.32	29.43	15.35
1994	59.12	132	60.42	23.27	18.11
1995	63.68	170	61.53	35.71	21.64
1996	69.86	184	62.67	43.63	23.22
1997	73.78	192	63.82	48.56	26.26
1998	74.71	207	65.00	45.92	26.97
1999	76.77	187	66.43	40.67	26.59
2000	80.50	200	67.42	54.50	27.78
2001	75.40	146	68.37	41.40	31.33

2002	78.33	181	69.30	51.55	36.06
2003	83.84	239	70.23	69.34	47.25
2004	87.82	299	71.15	97.54	63.17
2005	91.58	361	72.97	116.77	73.48
2006	99.59	483	72.97	139.58	85.54
2007	107.63	531	70.59	170.06	107.27
2008	106.27	648	71.13	201.96	132.03
2009	106.14	730	73.23	140.93	102.14
2010	109.27	615	74.47	185.54	113.88
2011	114.48	731	74.72	240.84	134.91

Table A.2: Equations produced from the execution of ensemble of GE executed on a recursive grammar and DE (M4) for prediction of energy demand.

S.No.	Model
1.	$\frac{E}{E_{max}} = \left[0.194\left(\frac{X4_i}{X4_{max}}\right)^{1.443}\right] + \left[-0.194 + \left(\frac{X2_i}{X2_{max}}\right)^{1.443}\right]$
2.	$\frac{E}{E_{max}} = \left[-0.047 + \left(\frac{X4_i}{X4_{max}}\right)^{0.180}\right] - \left[-0.047\left(\frac{X2_i}{X2_{max}}\right)\right] \times \left[-0.047 - \left(\frac{X4_i}{X4_{max}}\right)^{-0.047}\right] + \left[e^{\left\lvert -0.047\left(\frac{X4_i}{X4_{max}}\right)\right\rvert}\right]$
3.	$\frac{E}{E_{max}} = \left[-0.638 + \left(\frac{X2_i}{X2_{max}}\right)\right] + \left[-0.638 - \left(\frac{X1_i}{X1_{max}}\right)\right] \times \left[\log\left\lvert -0.638 - \left(\frac{X2_i}{X2_{max}}\right)\right\rvert\right] \times \left[-0.638 + \left(\frac{X4_i}{X4_{max}}\right)^{0.014}\right]$
4.	$\frac{E}{E_{max}} = \left[e^{\left\lvert 0.360+\left(\frac{X2_i}{X2_{max}}\right)\right\rvert}\right] \times \left[0.360 + \left(\frac{X4_i}{X4_{max}}\right)^{0.152}\right] - \left[e^{\left\lvert 0.360\left(\frac{X2_i}{X2_{max}}\right)\right\rvert}\right]$
5.	$\frac{E}{E_{max}} = \left[-0.242 - \left(\frac{X4_i}{X4_{max}}\right)\right] \times \left[-0.242 + \left(\frac{X2_i}{X2_{max}}\right)^{1.283}\right]$
6.	$\frac{E}{E_{max}} = \left[1.417\left(\frac{X2_i}{X2_{max}}\right)^{0.508}\right] + \left[log\left\lvert 1.417\left(\frac{X2_i}{X2_{max}}\right)\right\rvert\right] \times \left[0.508\left(\frac{X4_i}{X4_{max}}\right)^{1.417}\right] + \left[log\left\lvert 0.508\left(\frac{X2_i}{X2_{max}}\right)\right\rvert\right] \times$ $\left[log\left\lvert 1.417 + \left(\frac{X2_i}{X2_{max}}\right)\right\rvert\right]$
7.	$\frac{E}{E_{max}} = \left[0.999\left(\frac{X3_i}{X3_{max}}\right)^{0.199}\right] \times \left[0.999\left(\frac{X2_i}{X2_{max}}\right)\right]$
8.	$\frac{E}{E_{max}} = \left[0.547\left(\frac{X3_i}{X3_{max}}\right)^{0.390}\right] + \left[log\left\lvert 0.547 + \left(\frac{X2_i}{X2_{max}}\right)\right\rvert\right]$
9.	$\frac{E}{E_{max}} = \left[-2.443\left(\frac{X2_i}{X2_{max}}\right)^{0.187}\right] + \left[e^{\left\lvert 0.187-\left(\frac{X2_i}{X2_{max}}\right)\right\rvert}\right] + \left[e^{\left\lvert 0.187\left(\frac{X3_i}{X3_{max}}\right)\right\rvert}\right]$
10.	$\frac{E}{E_{max}} = \left[4.946\left(\frac{X2_i}{X2_{max}}\right)\right] \times \left[0.203\left(\frac{X4_i}{X4_{max}}\right)^{0.203}\right]$
11.	$\frac{E}{E_{max}} = \left[0.356\left(\frac{X4_i}{X4_{max}}\right)^{0.626}\right] + \left[-0.356 + \left(\frac{X2_i}{X2_{max}}\right)\right]$
12.	$\frac{E}{E_{max}} = \left[0.201\left(\frac{X4_i}{X4_{max}}\right)^{0.201}\right] \times \left[log\left\lvert 0.887\left(\frac{X4_i}{X4_{max}}\right)\right\rvert\right] \times \left[0.887 + \left(\frac{X2_i}{X2_{max}}\right)\right]$

13.	$\frac{E}{E_{max}} = \left[-0.189\left(\frac{X4_i}{X4_{max}}\right)^{-0.189}\right] - \left[e^{\left\lvert 0.716\left(\frac{X3_i}{X3_{max}}\right)\right\rvert}\right] \times \left[log\left\lvert 0.716 + \left(\frac{X1_i}{X1_{max}}\right)\right\rvert\right] \times \left[-0.189 + \left(\frac{X2_i}{X2_{max}}\right)^{0.716}\right]$
14.	$\frac{E}{E_{max}} = \left[-0.427 + \left(\frac{X2_i}{X2_{max}}\right)\right] + \left[0.400\left(\frac{X4_i}{X4_{max}}\right)^{0.400}\right]$
15.	$\frac{E}{E_{max}} = \left[0.356\left(\frac{X4_i}{X4_{max}}\right)^{0.626}\right] - \left[0.356 + \left(\frac{X2_i}{X2_{max}}\right)\right]$
16.	$\frac{E}{E_{max}} = \left[-0.323 + \left(\frac{X2_i}{X2_{max}}\right)\right] - \left[-0.327\left(\frac{X1_i}{X1_{max}}\right)\right]$
17.	$\frac{E}{E_{max}} = \left[log\left\lvert 2.222\left(\frac{X2_i}{X2_{max}}\right)\right\rvert\right] \times \left[log\left\lvert 2.433 + \left(\frac{X1_i}{X1_{max}}\right)\right\rvert\right]$
18.	$\frac{E}{E_{max}} = \left[log\left\lvert 2.159\left(\frac{X2_i}{X2_{max}}\right)\right\rvert\right] - \left[log\left\lvert 0.291 - \left(\frac{X4_i}{X4_{max}}\right)\right\rvert\right] \times \left[2.159\left(\frac{X1_i}{X1_{max}}\right)\right] \times \left[0.291\left(\frac{X3_i}{X3_{max}}\right)^{2.159}\right]$
19.	$\frac{E}{E_{max}} = \left[-0.534\left(\frac{X2_i}{X2_{max}}\right)^{-0.264}\right] - \left[-0.264 - \left(\frac{X1_i}{X1_{max}}\right)\right] \times \left[-0.264 + \left(\frac{X2_i}{X2_{max}}\right)\right]$
20.	$\frac{E}{E_{max}} = \left[log\left\lvert 2.584 + \left(\frac{X1_i}{X1_{max}}\right)\right\rvert\right] - \left[\left(-0.711 - \left(\frac{X2_i}{X2_{max}}\right)\right)^{-0.711}\right]$

Advances in Science, Technology and Engineering Systems Journal Vol. 7, No. 4, 92-97 (2022)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Metamaterial-Inspired Compact Single and Multiband Filters

Ampavathina Sowjanya[*], Damera Vakula

ECE Department, National Institute of Technology, Warangal, 506004, India

ARTICLE INFO

Article history:
Received: 25 May, 2022
Accepted: 26 July, 2022
Online: 16 August, 2022

Keywords:
Bandpass Filter
Metamaterial
Fractional Bandwidth
Roll-off
Insertion loss

ABSTRACT

In this paper, Compact bandpass filters have been designed. A single bandpass filter was designed using novel triple concentric complementary split-ring resonators placed along the microstrip line in the ground plane. Gaps and via were placed on the microstrip line to control electromagnetic characteristics, resulting in a single bandpass filter. In turn, spiral resonators were attached to the microstrip transmission line at the gaps in the transmission line to obtain a compact dual passband filter. Stepped impedance microstrip line and T-shaped stubs were attached to the microstrip line in between spiral resonators. The structure designed resulted in a Triple bandpass filter. A fractional bandwidth of 3% was achieved at the center frequency of 3GHz. The filter had a 1.5dB insertion loss which is the minimum value in the operating frequency band. The filter resonance frequency was 1.32 GHz and 2.47GHz which have a fractional bandwidth of 7.5% and 4.85% respectively and the corresponding insertion loss was 1.3dB and 1.8dB respectively. The triple bandpass filter had a fractional bandwidth of 1.16%, 11.4%, and 1.86%, centered at 1.29 GHz, 2.27 GHz, and 3.21GHz with 1.6dB, 1.3dB, and 1.8 dB insertion loss at the respective frequencies. The proposed bandpass filters are useful for GPS, WLAN, WiMAX, and radar applications.

1. Introduction

In modern wireless communication systems, there is an urgent need for compact filters with low insertion loss, and high selectivity to transmit a specific band of frequencies. Bandpass filters allow desired frequencies and reject unwanted frequencies. The basic requirement for wireless communications is portability. For wireless devices to be portable, internal components must be as compact as possible. Metamaterial provides the opportunity to exploit its unusual properties to control the electromagnetic characteristics of microwave devices such as filters, antennas, couplers, power dividers, etc., to design compact devices with better performance characteristics. The controlling of electromagnetic parameters is done by the attainment of negative permittivity and/or permeability over a certain range of frequencies by engineering materials such as altering their dimensions or creating gaps for better performance of microwave devices. In [1], backward wave propagation with split ring resonators which are loaded along the transmission line is demonstrated. In [2], a bandpass filter is designed using a high-temperature superconducting structure based on SRRs, but the filter was large. In [3] S-shaped complementary resonant structures were used to design a bandpass filter, but insertion loss was high in the passband. In [4], a complementary split-ring resonator (CSRR) based dual-mode patch bandpass filter is proposed, but the filter was large. In [5], a hybrid bandpass filter using substrate integrated waveguide and CSRR is proposed, but it had high insertion loss. In [6], a bandpass filter using multiple SRR and CSRRs was designed, but with large size. In [7], a microstrip bandpass filter for X band applications is presented. The filter is designed with SRRs but had high insertion loss with large size. In [8], a bandpass filter using an optimized coupling matrix synthesis method is designed, which was large. In [9], a bandpass filter based on half mode substrate integrated waveguide, periodic CSRRs, and defected ground structures is proposed, but is large. In [10], a low-profile metamaterial bandpass filter loaded with a four-turn complementary spiral resonator for wireless power transfer applications is designed. In [11], a full mode substrate integrated waveguide bandpass filter using CSRR is designed, that has high insertion loss.

Multiband filters are in demand to avoid the crowding of multiple devices for different applications. To accommodate multiple applications within a device and have better performance characteristics becomes challenging for the researchers. Numerous methods have been proposed. In [12], two dual bandpass filters one using CSRRs, complementary spiral resonators (CSR), and tapered substrate integrated waveguide;

[*] Corresponding Author: Ampavathina Sowjanya, sowjanya6754@gmail.com

https://dx.doi.org/10.25046/aj070412

another one with CSRRs, CSR, and substrate integrated waveguide are designed. In [13], a filter with dual passbands using tapered SRRs which consist of asymmetric shape square rings with interlinking is developed. In [14], a bandpass filter with CSRR and defected ground structures using substrate-integrated waveguide technology is proposed. In [15], a dual bandpass filter based on coupling between two identical SRRs and a CSRR is presented. In [16], varactor diodes were loaded on split ring slots to achieve a unit cell for frequency tunability is designed. In [17], a bandpass filter was designed for three bands using stub loading on resonators. The resonators were further coupled with internal resistors for dual mode operation. In [18], a compact three stubs loaded open-loop resonator-based triple bandpass filter is presented. In [19], resonators with open loops for filters with three pass bands are designed. In [20], a bandpass filter for triple bands using metamaterials is proposed. Rectangular stubs with a meander line were used in the design. The filters mentioned are large [12-20].

In this paper, metamaterial-inspired LC resonant structures were implemented to design a single bandpass filter. Firstly, gaps were formed along the microstrip transmission line. Via filled with copper was introduced along the microstrip transmission line. Triple CSSRs were etched in the ground plane just beneath the microstrip transmission line. Bandpass response with better sensitivity and insertion loss was achieved using the above configuration. The dual bandpass filter was designed by adding spiral resonators connected to the gaps introduced in the microstrip line to the proposed single bandpass filter. Stepped impedance resonant microstrip line and T-shaped stubs were added between spiral resonators attached to the microstrip line to the proposed dual bandpass filter to design a triple bandpass filter that helps in improving return loss.

2. Design Methodology

The microstrip transmission line allows all frequencies to pass through it. An electric field was induced from the microstrip line to a ground plane and a magnetic field around the microstrip transmission line. A bandpass filter was designed in this paper first by stopping all frequencies using gaps along the microstrip transmission line. By introducing a via filled with copper along the microstrip transmission line between the microstrip line and ground plane through the substrate, a passband was observed with high insertion loss and low return loss. Triple complementary concentric split-ring resonators were etched in the ground plane along the microstrip transmission line to act as LC resonant structure was excited by a time-varying electric field parallel to its axis and provided a single negative permittivity medium which introduced a passband over a certain range of frequencies but with low return loss in the passband and better insertion loss; sensitivity from passband to stopband transition and vice versa was also observed. This combination of gaps, vias, and complementary structures was introduced along with the transmission line resulting in passband response with better insertion loss and sensitivity by choosing appropriate dimensions for concentric rings, microstrip lines, vias, and gaps. Spiral resonators with optimum dimensions were connected to the gaps of the microstrip transmission lines which are coupled electrically to the triple concentric split ring resonators introducing one more passband, which helps in reducing insertion loss. Thus, a compact dual bandpass filter is designed with better insertion loss and high selectivity in the first band, and better insertion loss and selectivity in the second band. The stepped impedance resonant microstrip line for the proposed dual bandpass filter generated one more passband and the T-shaped stubs attached to the microstrip line in between spiral resonators helped to improve return loss. Thus, a compact triple bandpass filter was designed.

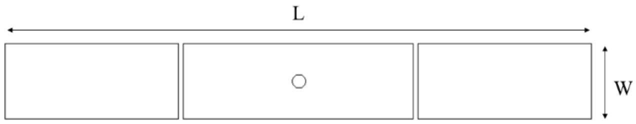


Figure 1: Top view of the single bandpass filter

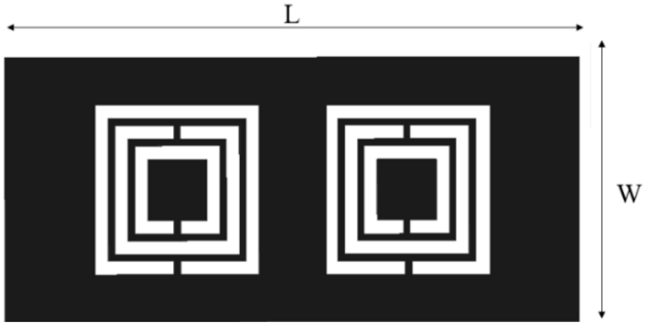


Figure 2: Bottom view of the single bandpass filter

The proposed filters were realized on RT Duroid 3010 substrate that has a height of 0.13 mm, and a dielectric constant of 10.2. Microstrip line made with 35 µm thickness copper material had a width of 2.25 mm for better impedance matching with 50Ω transmission line. Gaps along the microstrip line were 0.15 mm while via filled with copper had a diameter of 0.4 mm, and placed through the substrate from the microstrip line to the ground plane. Complementary split rings had a width of 0.4 mm and gaps in them had a width of 0.2 mm with gaps between concentric rings 0.2 mm. The outer ring had a length and width of 5 mm on each side. The top and bottom views of the simulated single bandpass filter are shown in Figure.1 and 2 respectively. The dimensions of the proposed filter were 17.5 mm×7.81 mm which was approximated by the guided wavelength λ_g at the center frequency as 0.55 λ_g ×0.25 λ_g.

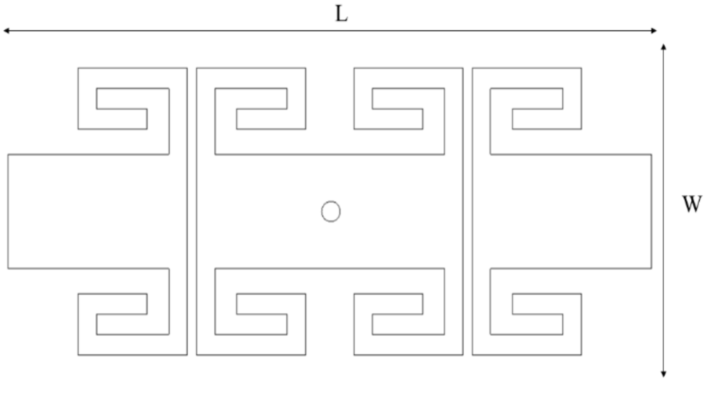


Figure 3: Top view of the dual bandpass filter.

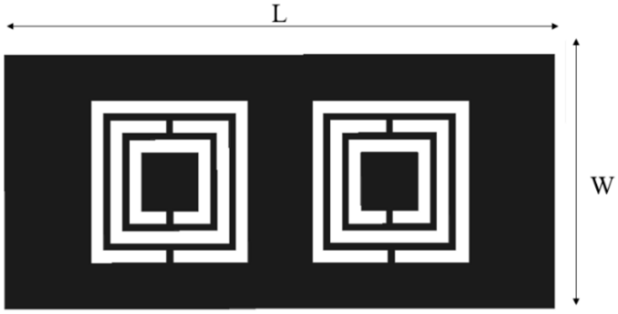


Figure 4: Bottom view of a dual bandpass filter

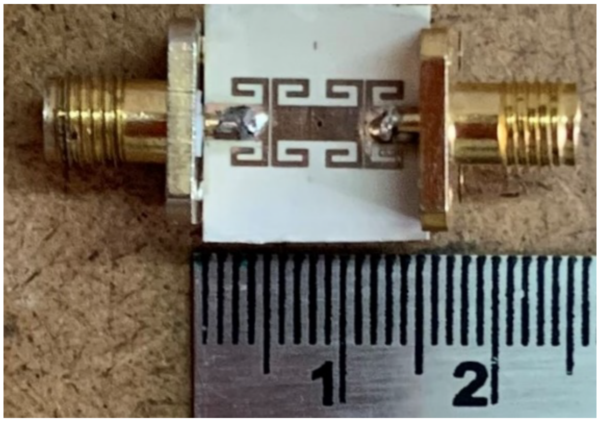


Figure 5: Fabricated top view of a dual bandpass filter

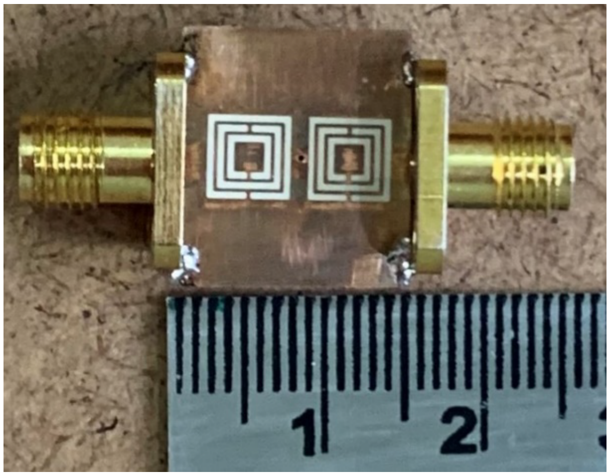


Figure 6: Fabricated bottom view of the dual bandpass filter

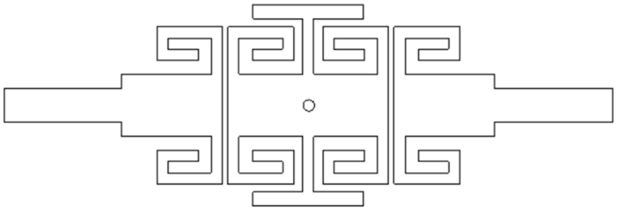


Figure 7: Top view of the triple bandpass filter

The dimensions of the filter were 14 mm × 15.81mm which is approximately 0.19 λ_g × 0.21 λ_g, where λ_g is the guided wavelength at the lower resonant frequency. The top and bottom views of the simulated dual bandpass filter are shown in Figures.3 and 4 while that of fabricated prototypes are shown in Figures.5 and 6, respectively.

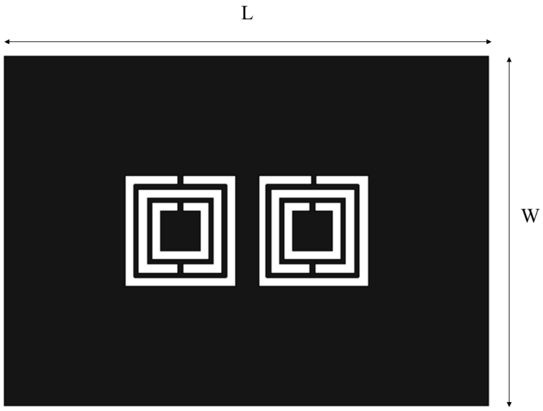


Figure 8: Bottom view of the triple bandpass filter

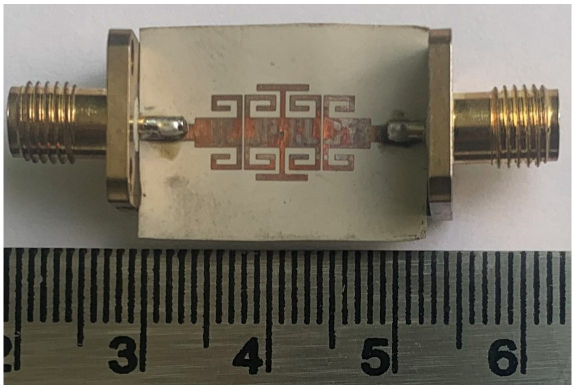


Figure 9: Fabricated top view of Triple bandpass filter

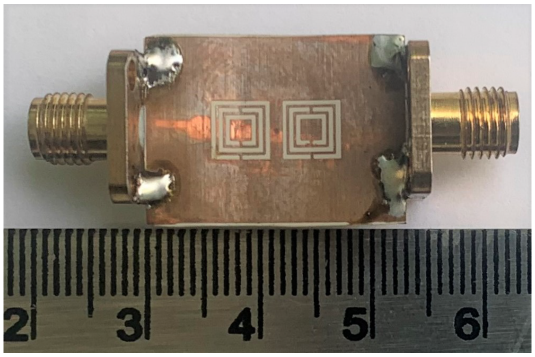


Figure 10: Fabricated bottom view of Triple bandpass filter

The dimensions of the designed filter were 22mm × 15.81mm which can also be represented as 0.3 λ_g × 0.2 λ_g, where the guided wavelength (λ_g) is calculated at the lowest resonant frequency. The structure of the triple bandpass filter is shown in Figures 7 and 8, while that of the fabricated filters is shown in Figures 9 and 10.

3. Results and Discussion

The proposed bandpass filters were designed and simulated with the help of the Ansys HFSS simulator using the finite element method (FEM). Ansys HFSS is three-dimensional electromagnetic simulation software. This software is used for designing and simulating high-frequency electronic products such as antennas, filters, etc., The most important component of FEM is mesh. Mesh can handle inaccuracies in the 3D model and produce reliable results consistently. The simulation results for the single bandpass filter are shown in Figure.11. The fractional bandwidth of the proposed filter is 3% at the 3GHz center frequency. The minimum insertion loss of the filter is 1.5 dB. The roll-off rate of the filter is 58.8 dB/GHz on the lower side and 67.8 dB/GHz on the upper side.

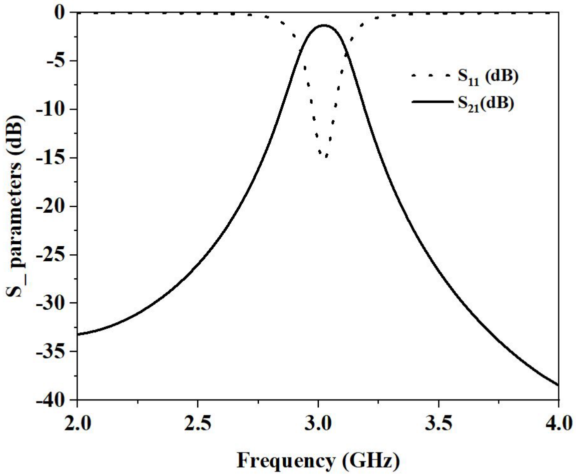


Figure 11: Simulation S-parameters of the single bandpass filter

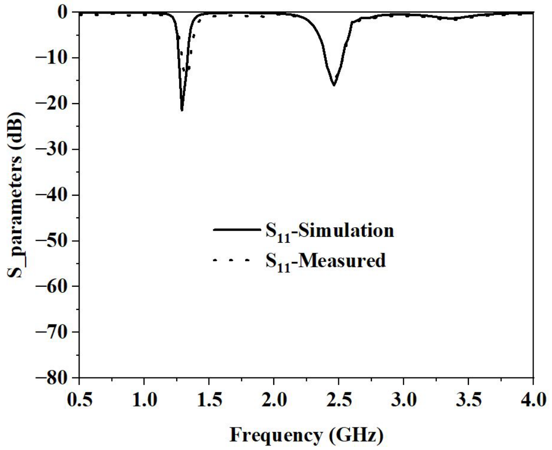


Figure 12: Simulation and measured return loss S_{11} of the dual bandpass filter

The performance of the proposed filters was validated by measuring S_parameters using vector network analyzer N5222A. The simulation and measured S_{11} and S_{21} results for the dual bandpass filter are shown in Figure.12 & Figure.13 respectively. The simulated dual bandpass filter has center frequencies at 1.3 GHz and 2.46 GHz. The fractional bandwidth was 8% for the first band, and 4.85% for the second band. The minimum insertion loss obtained was 1.15 dB and 1.7 dB for both bands respectively. The return loss was 21 dB for the first band and 17 dB for the second band. The measured dual bandpass filter had center frequencies at 1.32 GHz and 2.47 GHz. The fractional bandwidth of the proposed filter was 7.5% for the first band and 4.85% for the second band. The minimum insertion loss obtained was 1.3 dB for the first band and 1.8 dB for the second band. The return loss was 15 dB for the first band and 16 dB for the second band. The roll-off rate for the first band was 203 dB/GHz on the lower side, 64 dB/GHz on the upper side, for the second band 35 dB/GHz on the lower side, and 94.6 dB/GHz on the upper side. The electrical size of the dual bandpass filter is 0.039 λ_g^2.

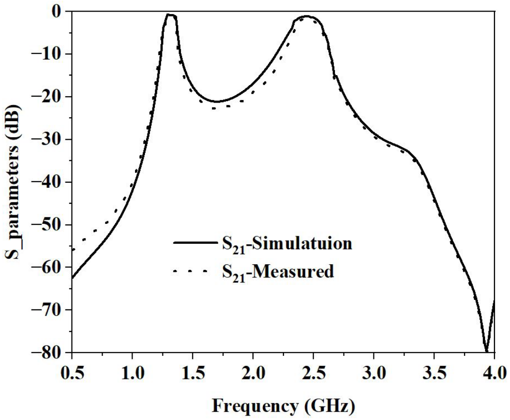


Figure 13: Simulation and measured insertion loss S_{21} of the dual bandpass filter

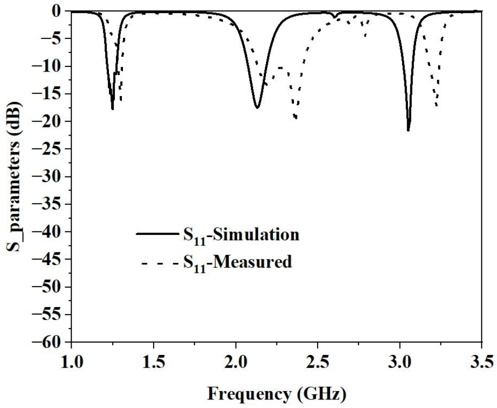


Figure 14: Simulation and measured return loss S_{11} of the triple bandpass filter.

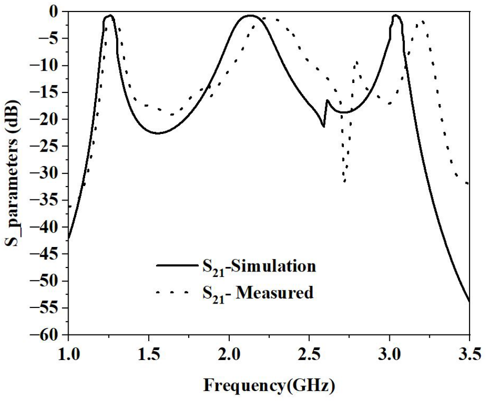


Figure 15: Simulation and measured insertion loss S_{21} of the triple bandpass filter.

The simulation results and those determined experimentally for S_{11} and S_{21} of the triple bandpass filter are shown in Figure.14 & Figure.15 respectively. The simulated triple bandpass filter has center frequencies at 1.27GHz, 2.2GHz, and 3GHz. The fractional bandwidth for the first band is 1%, for the second band is 6%, and for the third band is 1.8%. The minimum insertion loss obtained was 1.2 dB, 0.99 dB, and 1.5 dB respectively. The return loss was 18 dB,17 dB, 22 dB respectively. The measured triple bandpass filter had center frequencies at 1.29 GHz, 2.27 GHz, and 3.21 GHz. The fractional bandwidth was 1.16%, 11.4%, and 1.86% for all three bands respectively. The minimum insertion loss obtained was 1.6 dB, 1.3 dB, and 1.8 dB respectively in the passband. The return loss was 17 dB,13 dB,18 dB respectively. The roll-off rate for the first band was 262 dB/GHz on the lower side and 115 dB/GHz on the upper side, for the second band it was 56 dB/GHz on the lower side, 49 dB/GHz on the upper side, for the third band it was 47 dB/GHz on the lower side and 180 dB/GHz on the upper side. The electrical size of the triple bandpass filter is 0.06 λ_g^2.

Table 1: Single bandpass filter comparison table

References	**Center frequency (GHz)**	**Fractional Bandwidth (%)**	**Insertion loss (dB)**	**Electrical size (λ_g^2)**
4	3.41	3.5	<1	0.38
5	10	12	1.59	0.39
6	9.1	3	0.04	0.58
7	9	<1	2.24	0.92
8	2.4	12.5	0.7	0.39
Single bandpass filter	3	3	1.5	0.13

The measured dual bandpass filter operating at center frequencies of 1.32 GHz and 2.47 GHz with a fractional bandwidth of 7.5% and 4.85% respectively has a 0.039 λ_g^2 size is compact and useful for GPS and WLAN applications. The measured triple bandpass filter operates at center frequencies of 1.29 GHz,2.27 GHz, and 3.21 GHz with a fractional bandwidth of 1.16%, 11.4%, and 1.86% has 0.06 λ_g^2 is compact and useful for GPS, WLAN, WiMAX applications. The simulation and measurement results were in good agreement. The difference in measured results is due to dimensional tolerances during the etching of the copper material at the time of fabrication. The right shift in the measurement results may be due to a delta reduction in the dimensions of the design during fabrication.

From table.1., it is observed that the single bandpass filter is compact. From table 2, it is observed that the dual bandpass filter is compact. From table.3., it is observed that the triple bandpass filter is compact.

Table 2: Dual bandpass filter comparison table

References		Center frequency (GHz)	Fractional bandwidth (%)	Insertion loss (dB)	Electrical size (λ_g^2)
12	Type 1	4.32,5.52	5.76,4.98	2.79,2.92	0.22
	Type 2	3.84,4.96	6.69,5.28	1.65,3.33	0.07
13		8.89,11.04	<1, <1	2.68,2.61	3.3
14		5.57,7.84	6.8,4.1	1.8,2	0.22
15		5.02,8.92	<1, <1	6.22,5.23	0.17
Dual bandpass filter		1.32, 2.47	7.5,4.85	1.3,1.8	0.039

Table 3: Triple bandpass filter comparison table

References	Center frequency (GHz)	Fractional bandwidth (%)	Insertion loss (dB)	Electrical size (λ_g^2)
16	8.63,9.86,11.62	2.4,3.3,2.5	0.8,1.4,1.7	0.97
17	1.575,2.4,3.45	10,5,11	0.7,1.14,0.3	0.54
18	1.93,3.6,4.89	19.2, 11.6,2.86	-	0.076
19	1.96,2.6,3.9	5,11,3	1.5,0.6,1.83	0.41
20	2.1,5.7,7.3	57.1,8.7,4.1	0.8,1.6,2.4	0.073
Triple bandpass filter	1.29,2.27,3.21	1.16,11.4,1.86	1.6,1.3,1.8	0.06

4. Conclusion

A compact single bandpass filter using novel Triple concentric complementary split-ring resonator structure, gaps, and via along the microstrip line was proposed. The proposed single bandpass filter is useful for s band radar applications. The proposed filter has better insertion loss and sensitivity with a compact size. By adding spiral resonators to the proposed single bandpass filter design configuration, the dual bandpass filter was designed, fabricated, and measured resulting in a compact size, better insertion loss, and sensitivity. The proposed dual bandpass filter is useful for GPS and WLAN applications. The compact triple bandpass filter was designed using a stepped impedance microstrip line and T-shaped stubs for the dual bandpass configuration. The Triple bandpass filter is useful for GPS, WLAN, and WiMAX applications.

References

[1] R. Marques, F. Martin, and M. Sorolla, "Metamaterials with Negative Parameters: Theory, Design, and Microwave Applications". Hoboken, NJ: Wiley, 2008.

[2] Cheng Tan, Yu Wang, Zhong Ming Yan et.al., "Superconducting filter based on split ring resonator structures", IEEE Transactions on Applied Superconductivity, **29**(4), 1-4, 2019, doi:10.1109/TASC.2019.2891017.

[3] A. K. Horestani, M. D. Sindreu, J. Naqui et.al., "S-Shaped complementary split ring Resonators and their application to compact differential bandpass filters with common-mode suppression", IEEE Microwave and Wireless Components Letters, **24**,(3), 149-151, 2014, DOI: 10.1109/LMWC.2013.2291853.

[4] Y. Cheng, L. Zeng, W. Lu, "A compact CSRR- based dual-mode patch bandpass filter", IEEE conference, November 2015, DOI: 10.1109/IMWS-AMP.2015.7325010.

[5] Y. Zheng, Y. Zhu, Yuandan "Compact hybrid bandpass filter using SIW and CSRRs with wide Stopband rejection", IEEE conference, 2020, DOI: 10.1109/APMC47863.2020.9331361.

[6] S. Chandra "X-Band Metamaterial Bandpass Filter Design", International Journal of Engineering Research & Technology (IJERT), **10** (5), 1004-1006, 2021, ISSN: 2278-0181.

[7] R. L. Defitri and A. Munir, "X-band microstrip narrowband BPF composed of the split ring resonator", Progress in Electromagnetic Research Symposium (PIERS), 3468-3471, 2016, DOI: 10.1109/PIERS.2016.7735347.

[8] A. A. Ibrahim, M. A. Abdalla, A. B. Abdelrahman, "Wireless bandpass Filters build on metamaterials" Microwaves & RF.**57**(5), 1-7, 2018, ISSN: 07452993.

[9] B. Fellah, N. Cherif, M. Abri et.al., "CSRR-DGS bandpass filter based on half mode substrate integrated waveguide for X-Band applications", Advanced Electromagnetics, **10**(3), 39-42, 2021, doi:10.7716/aem.v10i3.1782.

[10] R. Keshavarz and N. Shariati, "Low profile metamaterial bandpass filter loaded with 4-Turn complementary spiral resonator for WPT applications", IEEE conference, 50-53, 2020, ISBN: 9781728160443.

[11] S. Moitra, S. Nayak, R. Regar et.al., "Circular complementary split-ring resonators (CSRR) based SIW BPF", Second International Conference on

Advanced Computational and Communication Paradigms (ICACCP), 2, 1-5, 2019, ISBN: 9781538679890.

[12] H.Y. Gao, Z. X. Tang, X. Cao, et.al., "Compact dual-band SIW filter with CSRRs and complementary spiral resonators", Microwave and Optical Technology Letters, **58**(1), 1-4, 2016, DOI: 10.1002/ mop.29482.

[13] A. Y. Rouabhi, M. Berka, A. Benadaoudi et.al., "Investigation of dual-band bandpass filter inspired by a pair of square coupled interlinked asymmetric tapered metamaterial resonator for X-band microwave applications", Indian Academy of Sciences, 2022, DOI: 10.1007/s12034-022-02693-6.

[14] G. Soundarya and N. Gunavathi, "Compact dual-band SIW bandpass filter using CSRR and DGS structure resonators", Progress in Electromagnetics Research Letters, 101, 79–87, 2021.

[15] M. Berka, H. A. Azzeddine, A. Bendaoudi et.al., "Dual-band bandpass filter based on Electromagnetic coupling of twin square metamaterial resonators (SRRs) and complementary resonator (CSRR) for wireless communications", Journal of Electronic Materials, **50**(8), 4887-4895,2021, DOI: 10.1007/s11664-021-09024-1.

[16] F. Gongora, A. E. Martynyuk, J.R.Cuevas et.al, "Independently tunable closely spaced triband frequency selective surface unit cell using the third resonant mode of split ring slots", IEEE Access, **9**,105564-105576, 2021, DOI: 10.1109/ACCESS.2021.3100325.

[17] M.U. Rahman and J.D. Park, "A compact tri-band bandpass filter using two stub-loaded dual-mode resonators", Progress in Electromagnetics Research M, **64**, 201–209, 2018, DOI: 10.2528/pierm17120404.

[18] N. Kumar and Y. K. Singh, "Compact tri-band bandpass filter using three stub-loaded open-loop resonators with wide stopband and improved bandwidth response", Electronics Letters, **50**(25), 1950–1952, 2014.

[19] M. Weng, M. H. Hsu, C. W. Lin et.al, "A simple method to design a tri-bandpass filter using open-loop uniform impedance resonators", Journal of Electromagnetic Waves and Applications **34**(1), 103-115, 2020, DOI: 10.1080/09205071.2019.1689181.

[20] A. Kumar, D. K. Choudhary, and R. K. Chaudhary, "Metamaterial tri-band bandpass filter using meander-line with rectangular-stub", Progress in Electromagnetics Research Letters, **66**, 121–126, 2017.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 7, No. 4, 98-104 (2022)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Performance Adjustment Factor for Fixed Solar PV Module

Kelebaone Tsamaase*[,1], Japhet Sakala[1], Kagiso Motshidisi[2] , Edward Rakgati[3] , Ishmael Zibani[1], Edwin Matlotse[1]

[1]*Department of Electrical Engineering, Faculty of Engineering and Technology, University of Botswana, P/Bag UB0061,Gaborone, Botswana*

[2]*School of Computing and Information Systems, Botswana Accountancy College, P/Bag 00319, Gaborone, Botswana*

[3]*Energy Division, Research and Innovation, Botswana Institution for Technology Research and Innovation, P/Bag 0082, Gaborone, Botswana*

ARTICLE INFO

Article history:
Received: 10 May, 2022
Accepted: 30 July, 2022
Online: 16 August, 2022

Keywords:
Photovoltaic cell
Solar energy
Single diode model

ABSTRACT

There are different factors which contribute to the amount of output power which can be delivered by solar photovoltaic (PV) module at any time of the year. The factors include but not limited to solar irradiation, ambient temperature, relative humidity, wind velocity, position of sun in the sky, geographical position of installed solar PV module and others. Apparent position of the sun in the sky contribute to the amount of electromagnetic radiation from the sun reaching the module's surface area. With the sun further away from the module and irradiance reaching the module surface area at an angle non perpendicular to the surface leads to low output power delivered by the module. In southern hemisphere the PV module experience high output power around November/December which are summer months and low output power around June/July which are winter months. This paper develops performance adjustment factor of fixed solar PV module to adjust PV module output power such that the PV system can deliver required amount of power during winter season. The results show that the value of performance adjustment factor for fixed solar PV module or system was established and can be used to adjust performance or output power for winter periods.

1. Introduction

Performance of Solar photovoltaic (PV) module plays an important part in the overall performance of solar photovoltaic energy system or solar array in particular. To improve performance of the system various studies have confirmed that many factors need to be looked into. This include system design whereby a PV tracking system was found to yield more output power than fixed installation system. However, due to other technical considerations of the system design the fixed installation system is the most commonly preferred. Other considerations are geographical position of the system expressed in terms of latitude and longitudinal points. Weather conditions such as availability of sunlight, wind velocity, temperature, rainfall, technology applied during manufacture of modules, which lead to different types of modules such as monocrystalline, polycrystalline, thin film, and others. To research further on factors and their effects on PV module performance this paper is an extension of work originally presented in 2021 10th International Conference on Renewable Energy Research and Application (ICRERA) which was dealing with maximum power output and voltage profiles as a response to varying irradiance and ambient temperature [1]. Single diode model shown in Figure 1 [2-6] was simulated using Simulink model shown in Figure 2 to generate voltage profile, power profile and profile for maximum power of a fixed installation as it varies from January to December. The outcome of the maximum power profile showed that the maximum power output is lowest during the months of June/July which is winter season in southern hemisphere and increased gradually when summer months (November/December) approached. While the temperature and irradiance were considered when producing profiles, the output power profile followed closely the profile of solar irradiance during the same period, indicating that irradiance is a dominant

[*] Corresponding Author: Kelebaone Tsamaase, tsamaase@ub.ac.bw

https://dx.doi.org/10.25046/aj070413

factor in generating output power by PV module. For the northern hemisphere, the opposite takes place whereby solar irradiation profile picture was observed with the maximum monthly global irradiance recorded during May, June and July months [7]. In this paper two profiles of maximum output power are established. One profile is for the lowest values of maximum power determined in a five year period. The other profile is for the highest values of maximum power profiles established in the same five year period. The two profiles are used to come up with the performance adjustment factor for fixed PV module or PV system. The performance adjustment factor is used to adjust PV system design output to a higher value, such that the system meets the expected output throughout the year irrespective of the time of the year at which it is designed or at which data used in the design was recorded. The rest of the sections of this paper are arranged in the following manner: Section 2 deals with definitions of terms. Section 3 deals with establishment of profiles for maximum output power and voltages. Section 4 presents methodology for generating profiles. Section 5 deals with the results and discussions and section 6 gives conclusions.

2. Nomenclature

k	Boltzmann constant (1.381×10^{-23} J/K)
q	Electron charge (1.602×10-19 C)
I_{pv}	photovoltaic output current
I_{ph}	photon current, at standard test conditions (G, 1000 W/m^2, Temperature $25^oC, AM$ 1.5)
G	Solar irradiance
G_0	solar irradiance (W/m^2) at standard test conditions (G, 1000 W/m^2, Temperature $25^oC, AM$ 1.5))
(I_{SCr})	short-circuit current of a PV module at reference condition, ((STC) 1000 W/m^2, Temperature $25^oC, AM$ 1.5)
K_i	short circuit temperature coefficient at module reference conditions
I_s	reverse saturation current
I_{sh}	shunt current
V_{PV}	output voltage of PV module
T	ambient temperature (Kelvin)
n	Diode ideality constant
N_S	number of cells in series
I_D	diode current
R_S	series resistance
R_{Sh}	shunt resistance
F_p	performance adjustment factor
P_{max}	Maximum power output of PV module
$P_{\max\,(\max)}$	Highest value of maximum-power output of the PV module recorded in a year
$P_{\max\,(\min)}$	Minimum value of maximum-power output of the PV module recorded in a year
$P_{out(initial)}$	Output power of the PV module or PV system Design
$P_{out(final)}$	Adjusted output power of the MV module or PV systems.

3. Establishing Maximum Power Profiles for Winter and Summer Seasons

Maximum power output of the PV module for the winter and summer periods are computed by use of single diode model shown in Figure 1 [2]-[6].

The parameters in the model are current source which represents photon current (I_{ph}), diode current (I_d) passing through the diode which is connected across the source terminals, a shunt resistor (R_{sh}) which is connected across the current source terminals and represents leakages in form of leakage current (I_{sh}) experienced by the cell, and a series resistor (R_s) representing resistances due to solder bonds, metallization of cell, terminations of junction box, and emitter and base regions [2, 8, 9].

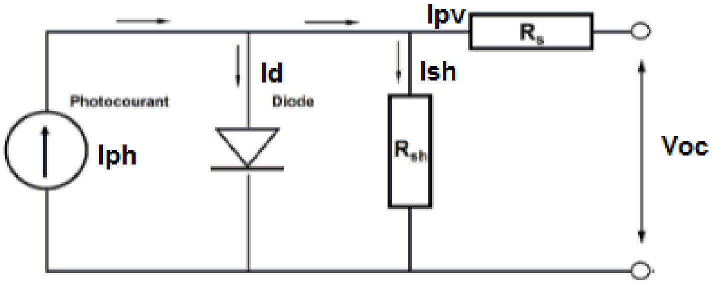


Figure 1: Single diode model representation of a PV cell

The cell model was used to represent PV module because the module is made up of electrical and mechanical interconnection of solar cells. The model in Figure 1 was simulated using Simulink and its corresponding Simulink model is shown in Figure 2 [1].

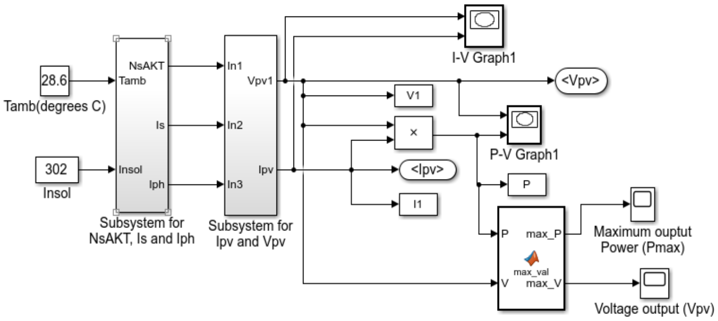


Figure 2: Simulink model for computing maximum power output

Considering expressions of different electrical parameters of the model the value of output current I_{pv} expressed in mathematical form is presented in (1) [10, 11, 12, 13, 14, 15, 16].

$$I_{pv} = I_{ph} - I_D - I_{sh} \tag{1}$$

Where I_{ph}, I_D and I_{sh} are shown in (2), (3) and (4) respectively.

$$I_{ph} = \left[I_{SCr} + K_i(T - 298)\frac{G}{G_0}\right] \tag{2}$$

$$I_D = I_S\left[e^{\frac{q(V_{PV}+IpvR_S)}{N_S nTk}} - 1\right] \tag{3}$$

$$I_{sh} = \frac{V_{PV}+R_S I_{PV}}{R_{sh}} \tag{4}$$

4. Methodology for generating maximum power output (P_{max}) profiles

The following data was used to establish P_{max} profiles. PV module characteristics, ambient temperature and irradiance from five different weather stations in Botswana [17]. Dataset for the two weather variables covers a five year period, from 2009 to 2013.

The characteristic parameters of a photovoltaic module were obtained from a module installed at a weather station called Renewable Energy Research Centre (CERC) at the University of Botswana (UB) in capital city of Republic of Botswana, Gaborone and from similar model online. The geographical positioning system location coordinates of Gaborone are 24° 39' 11.7252" S and 25° 54' 24.4512" E [18]. The parameters are detailed in Table 1.

Table 1: Photovoltaic Module Parameters

Name	Solaire Direct Technologies
Model	**SD ECO PLUS 150 W**
Electrical Ratings at STC: 1000 W/m^2; AM 1.5 spectrum; Temperature 25 ^oC	
Peak Power	150 W_p
ΔP_{max}	$\pm$ 2.5 W_p $with$ tolerance 1%
Warranted minimum P_{max}	147.5 (tolerance 1%)
Voltage (V_{mp})	18.0 V
Current (I_{mp})	8.05 A
Open Circuit Voltage (V_{OC})	22.6 V
Short Circuit Current (I_{SC})	8.72 A
Maximum System Voltage	1000 V

Ambient temperature and solar irradiance were obtained from the following weather stations. Sir Seretse Khama Airport (SSKI) Weather Station. The weather station is located in the Capital City Gaborone, in the south-eastern part of the country with geographical positioning System (GPS) coordinates of 24° 39' 11.7252" S and 25° 54' 24.4512" E. Kasane weather station, located in the town of Kasane at the northern part of Botswana. Its GPS location coordinates are 17° 48' 10.944" S and 25° 8' 56.0364" E. Tshane weather station, located in the town of Tshane in the southern part of Botswana. Its GPS location coordinates are 24° 1' 7.032" S; 21° 51' 33.084" E. Maun weather station, located in the town of Maun at the north-western part of Botswana. Its GPS location coordinates are 19° 59' 23.6004" S; 23° 25' 24.888" E. Shakawe weather station, located at the village of Shakawe with the following coordinates 18° 21' 46.728" S 21° 50' 51" E [18].

For each weather station, ambient temperature and solar irradiance for a five year period were considered. The temperature was considered because besides irradiance it is one of the factors affecting PV module electrical power output [19, 20]. Data provided from the weather stations was for daily recordings and using that data the monthly average values of temperature and irradiance were calculated. The monthly average values were used as input variables into the Simulink model to compute monthly values of P_{max} for each month. Figure 3 shows corresponding graphical representation of power output against the voltage obtained from the model. From each graph $P_{\max}$ obtained or indicated were recorded [1]. Similar procedure was followed to generate monthly P_{max} for the years 2009 to 2013.

To plot profiles, the lowest values of maximum power ($P_{\max\,(\min)}$) from each year were plotted. Same procedure was followed to plot a profile for highest values of maximum power ($P_{\max\,(\max)}$)from each year. The difference between two profiles (ΔP_{max}) was determined and used to formulate a factor referred to as performance adjustment factor. This factor gives a relationship between P_{max} in summer season and P_{max} in winter season in southern hemisphere.

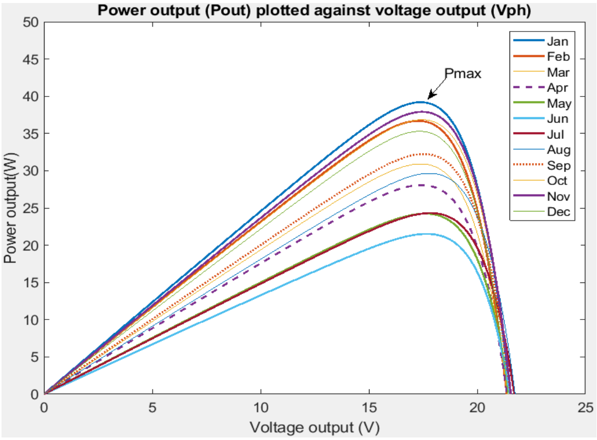


Figure 3: Photovoltaic module power output plotted against voltage for different months of the year

5. Results and Discussion

5.1. Sir Seretse Khama Airport (SSKI) Weather Station

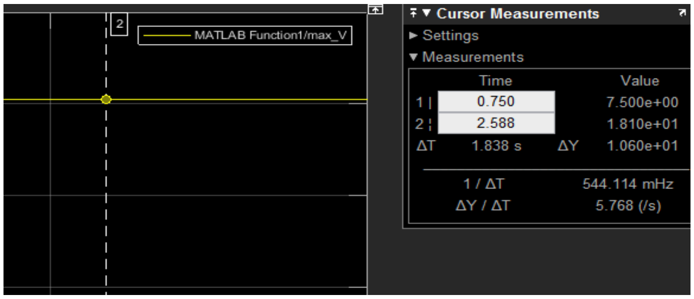
The irradiance and temperature data recorded at the station was used in the simulation model to get values of power output and voltage for the years 2009, 2010, 2011, 2012 and 2013. The Figures 4 and 5 show how output power and voltage values respectively were determined and recorded by Simulink model, with Tables 2, 3, 4, 5 and 6, of corresponding years respectively showing values obtained from Simulink model. From the tables the maximum and minimum values of output power recorded each year were obtained and recorded in Table 7. Also shown in Table 7 are corresponding values of voltages and months of the year at which such minimal and maximum values were recorded.

To establish five year-period profiles for highest maximum value of power ($P_{\max\,(\max)}$) and lowest maximum value of power ($P_{\max\,(\min)}$) as shown in Figure 6, data from Table 7 was used. Similarly voltage profile was also generated.

Table 2: 2009 Maximum power and voltage output determined with Simulink Model

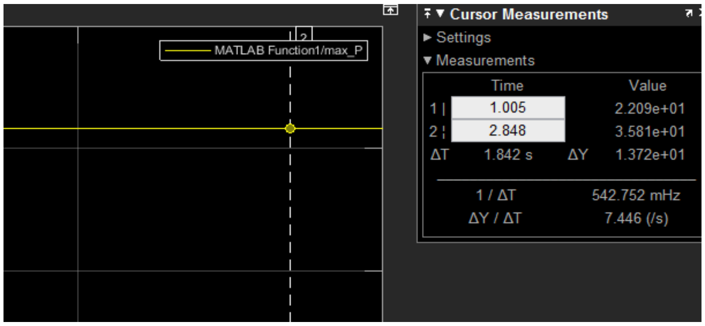
Month	Average Temperature	GHI	P_{max} (W)	V_{pv} (V)
Jan	24.2	264.0	34.6	17.6

Feb	24.9	259.0	33.8	17.5
Mar	20.8	234.0	30.8	17.7
Apr	19.4	235.0	31.1	17.8
May	18.1	188.0	24.6	17.6
Jun	14.0	150.0	19.5	17.7
Jul	10.0	187.0	25.2	18.2
Aug	14.7	226.0	30.4	18.1
Sep	19.8	254.0	33.8	17.8
Oct	22.9	252.0	33.1	17.6
Nov	22.4	271.0	35.8	17.7
Dec	25.2	316.0	41.7	17.7



	Time	Value
1	0.750	7.500e+00
2	2.588	1.810e+01
ΔT	1.838 s	ΔY 1.060e+01
1 / ΔT		544.114 mHz
ΔY / ΔT		5.768 (/s)

Figure 4: Simulink plotting output voltage



	Time	Value
1	1.005	2.209e+01
2	2.848	3.581e+01
ΔT	1.842 s	ΔY 1.372e+01
1 / ΔT		542.752 mHz
ΔY / ΔT		7.446 (/s)

Figure 5: Simulink plotting maximum power

Table 3: 2010 Maximum power and voltage output determined with Simulink Model

Month	Average Temperature	GHI	P_{max} (W)	V_{pv} (V)
Jan	25.3	261.0	34.0	17.5
Feb	25.4	279.0	36.5	17.5
Mar	24.0	235.0	30.5	17.5
Apr	20.6	163.0	20.8	17.3
May	16.6	173.0	22.6	17.7
Jun	12.1	181.0	24.2	18.0
Jul	13.9	182.0	24.1	17.9
Aug	15.6	229.0	30.7	18.0
Sep	21.5	275.0	36.5	17.8
Oct	25.5	304.0	40.0	17.6
Nov	25.4	266.0	34.7	17.5
Dec	24.6	290.0	38.2	17.6

Table 4: 2011 Maximum power and voltage output determined with Simulink Model

Month	Average Temperature	GHI	P_{max} (W)	V_{pv} (V)
Jan	24.0	256.0	33.5	17.5
Feb	25.2	287.0	37.6	17.6
Mar	25.9	257.0	33.3	17.4
Apr	17.8	176.0	22.9	17.6
May	12.0	191.0	25.6	18.1
Jun	13.8	184.0	24.4	17.9
Jul	10.4	187.0	25.2	18.2
Aug	12.7	224.0	30.4	18.2
Sep	20.7	275.0	36.6	17.8
Oct	22.2	287.0	38.1	17.8
Nov	24.0	287.0	37.8	17.7
Dec	27.0	240.0	30.8	17.3

Table 5: 2012 Maximum power and voltage output determined with Simulink Model

Month	Average Temperature	GHI	P_{max} (W)	V_{pv} (V)
Jan	25.3	320.0	42.3	17.7
Feb	26.6	297.0	38.8	17.5
Mar	23.9	266.0	34.9	17.6
Apr	19.8	230.0	30.3	17.7
May	17.1	206.0	27.2	17.8
Jun	14.0	180.0	23.8	17.9
Jul	14.0	195.0	26.0	18.0
Aug	17.0	225.0	30.0	17.9
Sep	21.0	267.0	35.5	17.8
Oct	21.9	260.0	34.3	17.7
Nov	25.9	316.0	41.6	17.6
Dec	23.3	290.0	38.4	17.7

Table 6: 2013 Maximum power and voltage output determined with Simulink Model

Month	Average Temperature	GHI	P_{max} (W)	V_{pv} (V)
Jan	26.9	293.0	38.2	17.5
Feb	26.8	302.0	39.5	17.5
Mar	24.5	275.0	36.1	17.6
Apr	21.7	220.0	28.7	17.6
May	16.6	203.0	26.9	17.8
Jun	14.5	182.0	24.1	17.9
Jul	16.0	192.0	25.4	17.8
Aug	17.5	223.0	29.6	17.9
Sep	23.9	255.0	33.3	17.5
Oct	24.0	292.0	38.5	17.7
Nov	25.7	312.0	41.1	17.6
Dec	25.5	247.0	32.0	17.4

Table 7: Yearly highest ($P_{max\,(max)}$) and lowest ($P_{max\,(min)}$) values of maximum power and average voltage output for SSKI

Year	$P_{\max(max)}$ (W)	Mth	$P_{\max(min)}$ (W)	Mth	Average V_{pv} (V)
2009	41.7	Dec	19.5	Jun	17.8
2010	40.0	Oct	20.8	Apr	17.7
2011	38.1	Oct	22.9	Apr	17.8
2012	42.3	Jan	23.8	Jun	17.7
2013	41.1	Nov	24.1	Jun	17.7

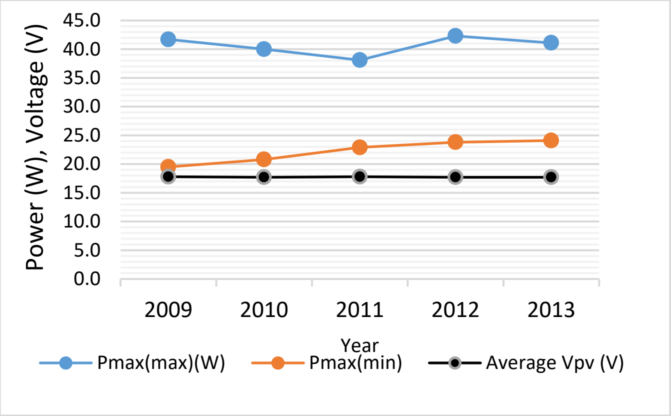


Figure 6: Profiles of highest and lowest values of maximum power and voltages for SSKI Weather station

5.2. Kasane, Maun, Shakawe and Tshane Weather Stations

To develop profiles for Kasane, Maun, Shakawe and Tshane weather stations the approached described in SSKI weather station at section 5.1 was followed. Simulated values of power output and voltages arising from irradiance and temperature values for different months and years are shown in Table 8. Corresponding power and voltage profiles of the same weather stations are shown in Figures 7 to 10 respectively.

Table 8: Highest and lowest values of maximum power and average voltage output

Place	Yr	Highest value of Maximum power, $P_{\max\,(max)}$ (W)	Mth	Lowest value of Maximum power, $P_{\max\,(min)}$ (W)	Mth	Average Output voltage, V_{pv} (V)
Tshane	2009	42.5	Dec	22.3	Jun	17.7
	2010	43.6	Dec	25.3	Jun	17.7
	2011	41.6	Nov	24.5	Jun	17.7
	2012	40.1	Nov	24.4	Jun	17.7
	2013	42.7	Nov	24.3	Jun	17.7
Maun	2009	37.5	Nov	25.4	Jun	17.6
	2010	39.2	Oct	25.6	Jul	17.5
	2011	40.6	Oct	27.8	Jun	17.6
	2012	38.5	Jan	27.0	Jun	17.6
	2013	39.9	Feb/Oct	27.2	Jun	17.6
Kasane	2009	36.9	Jan	27.2	Jun	17.6
	2010	38.9	Sept	25.0	Jul	17.6
	2011	38.9	Sept	28.8	Jun	17.6
	2012	38.1	Sept	28.4	Jun	17.6
	2013	38.9	Oct	28.8	Jun	17.6
SSKI	2009	41.7	Dec	19.5	Jun	17.8
	2010	40.0	Oct	20.8	Apr	17.7
	2011	38.1	Oct	22.9	Apr	17.8
	2012	42.3	Jan	23.8	Jun	17.7
	2013	41.1	Nov	24.1	Jun	17.7
Shakawe	2009	36.6	Nov	27.0	Jun	17.6
	2010	38.8	Oct	26.1	Jul	17.6
	2011	40.9	Oct	29.3	Jun	17.7
	2012	38.8	Sept	28.4	Jun	17.7
	2013	39.5	Oct	28.6	Jun	17.6

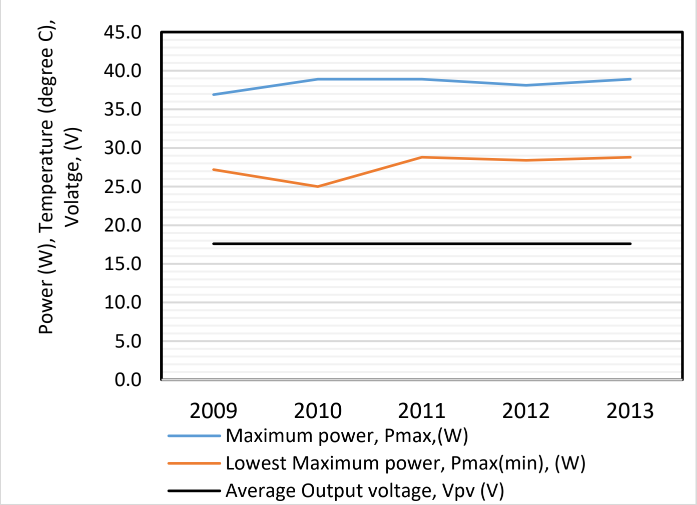


Figure 7: Profiles of highest and lowest values of maximum power and voltages for Kasane Weather station

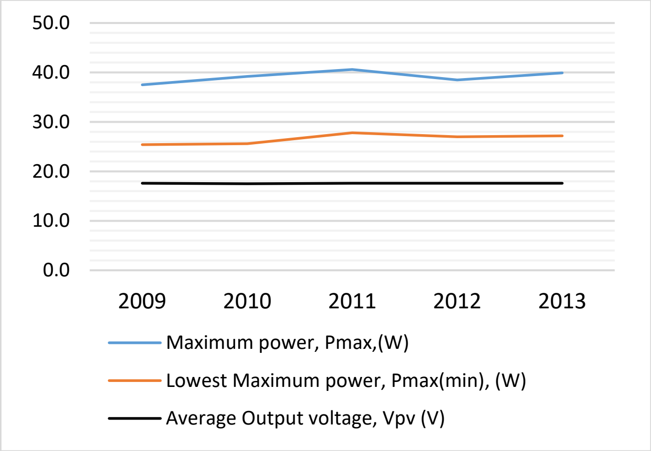


Figure 8: Profiles of highest and lowest values of maximum power and voltages for Maun Weather station

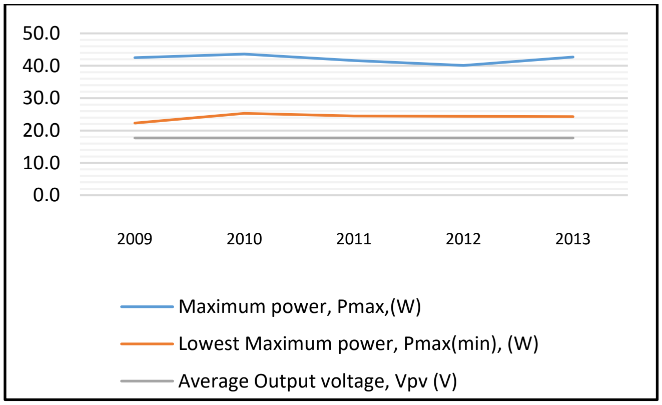


Figure 9: Profiles of highest and lowest values of maximum power and voltages for Tshane Weather station

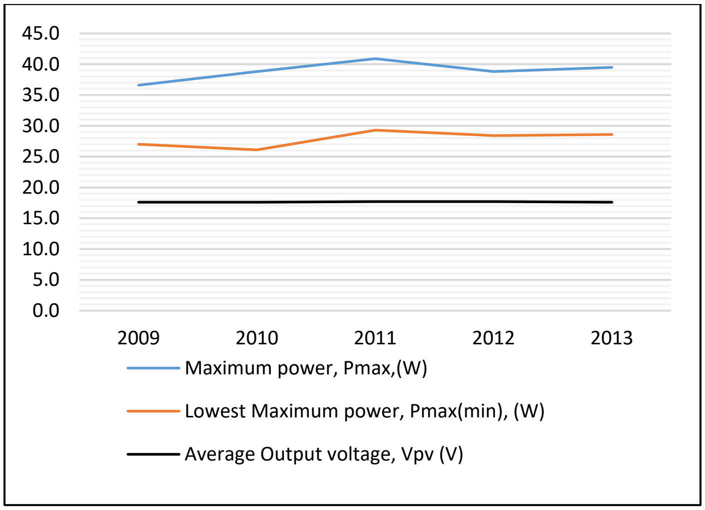


Figure 10: Profiles of highest and lowest values of maximum power and voltages for Shakawe Weather station

5.3. Five year national average profiles

Data from five weather stations shown in table 8 was used to come up with yearly averages shown in table 9, where also the yearly averages were computed to come up with a five year average value for $P_{\max(\max)}$ and $P_{\max(\min)}$ respectively. Table 9 was used to generate figure 11 where average values of highest value of maximum power $P_{\max(\max)}$, lowest value of maximum power $P_{\max(\min)}$ and voltage for each year were plotted to come up with national average profiles. Profiles in figure 11 were refined further as shown in figure 12 by plotting average values stretching on a five year long period.

Table 9: Yearly averages for maximum power and voltage output

Year	Average Highest Maximum power, $P_{\max(\max)}$ (W)	Average Lowest Maximum power, $P_{\max(\min)}$ (W)	Average Output voltage, V_{pv} (V)
2009	39.04	24.28	17.7
2010	40.10	24.56	17.6
2011	40.02	26.66	17.7
2012	39.56	26.40	17.7
2013	40.42	26.6	17.6
Average	39.8	25.7	17.7

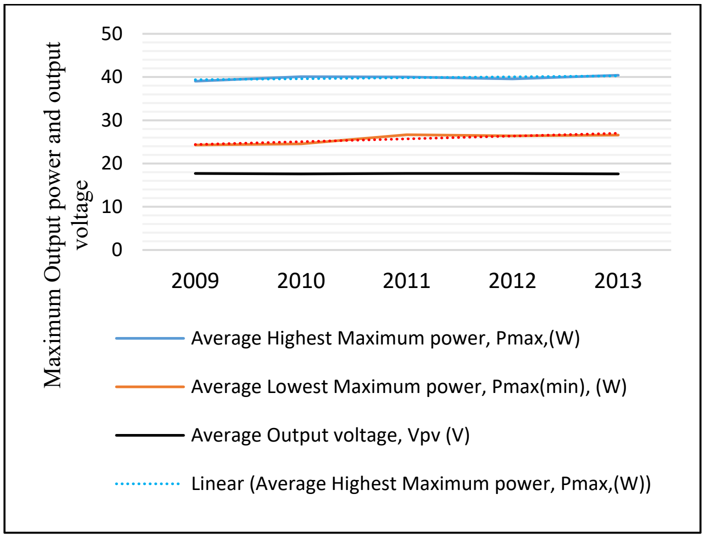


Figure 11: Five year period national averages of maximum power and voltage output

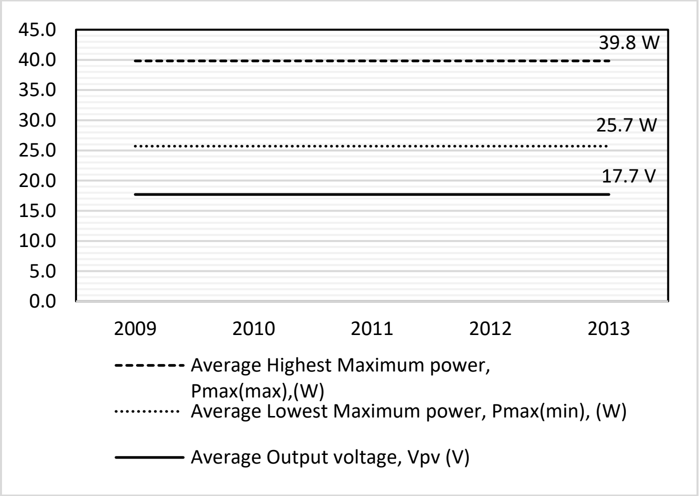


Figure 12: Refined profiles for five year period national averages of maximum power and voltage

5.4. Determining the value of performance adjustment factor, F_p

The relationship between two final average values of maximum power as shown in table 9 and figure 12 is represented by a factor herein referred to as performance adjustment factor, F_p. The purpose of this factor is to make adjustment on PV systems design output power which would have been determined or calculated during PV system design stage. The importance of adjustment is to ensure that the PV system or the PV module gives a required amount of power output throughout the year regardless of the season or month of the year at which such system was designed. F_p is obtained as shown in (5) to (8)

$$F_p = \frac{P_{\max\,(\min)}}{P_{\max\,(\max)}} = \frac{25.7}{39.8} = 0.645728643 \qquad (5)$$

$$F_p \approx 0.6457 \qquad (6)$$

The final value of the solar PV system or module output power obtained after adjustment is designated, $P_{\text{out (final)}}$, and is presented as in (3)

$$P_{\text{out}(final)} = \frac{P_{out(initial)}}{F_p} = \frac{1}{F_p}\left(P_{out(initial)}\right) \qquad (7)$$

where:

$$0.6457 \leq F_p \leq 1 \qquad (8)$$

$P_{\text{out(initial)}}$ is the output power, before adjustment, of the fixed solar PV system or module calculated using site specific parameters (temperature and irradiation) at any time of the year.

5.5. Conditions for F_p

According to (7) and (8) if $P_{out(initial)}$ is calculated using solar irradiance and temperature values of summer season, example October/November, then to get $P_{out(final)}$ the value of F_p at this season is maximum with a value of 1.0. For the same PV system or module, to adjust its calculated output for the purpose of winter performance the $P_{out(initial)}$ should be divided by 0.6457. The $P_{out(final)}$ obtained after adjustment can now be used to resize the system to ensure that it will deliver such adjusted power in winter.

If the $P_{out(initial)}$ is calculated using solar irradiance and temperature values of winter season, the value of $F_p$will remain maximum throughout, that is value of 1. In this case $P_{out(initial)}$ is taken as $P_{out(final)}$. This means that the system designed or sized using winter weather data (irradiance and temperature) is capable of delivering required output power throughout the year including during summer season and there is no need for any performance adjustment

6. Conclusion

According to the results obtained the module maximum output power in summer (around November/ December) is higher than the maximum power in winter (June / July). To ensure that PV system or module produce required amount of power throughout the year a performance adjustment factor was derived to be used in calculations for determining power output in winter. Further work will involve experimental setup to refine performance adjustment factor with other variables considered.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] K. Tsamaase, J. Sakala, E. Rakgati, I. Zibani, K. Motshidisi, "Solar PV Module Voltage Output and Maximum Power Yearly profile using Simulink-based Model," in 2021 10th International Conference on Renewable Energy Research and Application(ICRERA), 26-29, 2021, doi: 10.1109/ICRERA52334.2021.9598794

[2] C.B.Honsberg, S.G.Bowden, "Photovoltaics Educationte," www.pveducation.org, 2019.

[3] H. Ibrahim, H. Anani, "Variations of PV module parameters with irradiance and temperature," in 9th International Conference on Sustainability in Energy and Buildings, 5-7, 2017, doi:10.1016/j.egypro.2017.09.617

[4] T. M Silverman, M. G. Deceglie, I. Subedi, N. J. Podraza, I. M. Slauch, V. E. Ferry, I. M Slauch, "Reducing Operating Temperature in Photovoltaic Modules," Photovoltaics, **8**(2), 532-540, 2018, doi.org/10.1109/JPHOTOV.2017.2779842

[5] H. Singh, D. Kaur, P. S. Cheema, "Optimal design of Photovoltaic Power System for a residential load," in International Conference on Inventive Systems and Control (ICISC), 19-20, 2017, doi: 10.1109/ICISC.2017.8068595

[6] Y. Wang, Y. Liu, C. Wang, Z. Li, X. Sheng, H. Lee, N. Chang, H. Yang, "Storage-Less and Converter-Less Photovoltaic Energy Harvesting With Maximum Power Point Tracking for Internet of Things," Computer-aided design of integrated circuits and systems, **35**(2), 173 - 186, 2016, doi:10.1109/TCAD.2015.2446937

[7] T. Abu Dabbousa, I. Al-Reqeb, S.Mansour, M. Ammirrul Atiqi Mohd Zainuri, "The Effect of Tilting a PV Array by Monthly or Seasonal Optimal Tilt Angles on Energy Yield of a Solar PV System," in 2021 International Conference on Electric Power Engineering (ICEPE), 23-24, 2021, doi:10.1109/ICEPE-P51568.2021.9423475

[8] P.Hao, Y. Zhang, "An Improved Method for Parameter Identification and Performance Estimation of PV Modules From Manufacturer Datasheet Based On Temperature-Dependent Single-Diode Model," Photovoltaics, **11**(6), 1446 - 1457, 2021, doi:10.1109/JPHOTOV.2021.3114592

[9] P. Natarajan, and R. Muthu, "Performance Improvement of PV Module at Higher Temperature Operation," Engineering science and technology, 2(5), 2012, doi: ISSN 2250-3498

[10] S. T. Kim, S. Bae, Y. C. Kang, J. W. Park, " Energy Management Based on the Photovoltaic HPCS With an Energy Storage Device," Industrial electronics, **62**(7), 4608 - 4617, 2015, doi: 10.1109/TIE.2014.2370941

[11] V. Tamrakar, S.C. Gupta, Y. Sawle, "Single-diode PV cell modeling and study of characteristics of single and two-diode equivalent circuit," Electrical and electronics, **4**(3), 13-24, 2015, dio: 10.14810/elelij.2015.4302,

[12] H. Park, Y. Kim, H. Kim, "PV Cell Model by Single-diode Electrical Equivalent Circuit" Electrical engineering and technology, **11**(5), 1323-1331, **11**(5): doi:10.5370/JEET.2016.11.5.1323,

[13] T. Ahmed, T. Gonçalves, M. Tlemcani, "Single diode model parameters analysis of photovoltaic cell," in 2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA), 20-23, 2016, doi: 10.1109/ICRERA.2016.7884368,

[14] B. C. Babu, S. Gurjar, "A Novel Simplified Two-Diode Model of Photovoltaic (PV) Module," **4**(4), 1156-1161, 2014, doi: 10.1109/JPHOTOV.2014.2316371

[15] N. Islam Sarkar, "Effect of various model parameters on solar photovoltaic cell simulation: a SPICE analysis," Sarkar Renewables, **3**(13), 2016, doi: 10.1186/s40807-016-0035-3

[16] H. K. Mehta, H. Warke, K. Kukadiya, A. K. Panchal, "Accurate Expressions for Single-Diode-Model Solar Cell Parameterization," Photovoltaics, **9**(3), 803-810, 2019, doi: 10.1109/JPHOTOV.2019.2896264

[17] Department of Meteorological Services, Plot No. 54216, Corner of Metsimotlhaba/Maaloso Roads, meteo@gov.bw. Botswana. 2021.

[18] Longitude and latitude finder. 2021. www.latlong.net

[19] V. Herbort, R. von Schwerin, B. Compton, L. Brecht, H. te Heesen, "Insolation dependent solar module performance evaluation from PV monitoring data," in 38th IEEE Photovoltaic Specialists Conference, 3-8, 2012, doi:10.1109/PVSC.2012.6317834

[20] M. Islam, M. Mehedi Hasan Shawon, S. Akter, A. Chowdhury, S. Islam Khan, M. Rahman, "Performance Investigation of Poly Si and Mono Si PV Modules: A Comparative Study," in International Conference on Energy and Power Engineering (ICEPE), 14-16, 2019. doi: 10.1109/CEPE.2019.8726598

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 7, No. 4, 105-113 (2022)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

A Comparison of Cyber Security Reports for 2020 of Central European Countries

Kamil Halouzka*, Ladislav Burita, Aneta Coufalikova, Pavel Kozak, Petr Františ

Department of Informatics and Cyber Operations, University of Defence, Brno, 66210, Czech Republic

ARTICLE INFO

Article history:
Received: 11 May, 2022
Accepted: 23 July, 2022
Online: 16 August, 2022

Keywords:
Cyber security
Cyber threat
Cyber-attack
Annual reports
Strategic documents
Phishing email defense

ABSTRACT

The aim of the article is to analyze the annual reports on cyber security of Central European countries, i.e. the Czech Republic, Slovakia, Poland, Germany, and Austria. The article focuses on the development of the state of cyber security, actors of threats in cyberspace, cyber threats, and the most common types of attacks. The article evaluates the objectives of cyber-attacks from the point of view of state institutions, organizations, and state and private companies, and they have listed the follow-up measures here. The method used is a critical verbal evaluation with comments and comparative analysis to find the strengths and weaknesses of the evaluated cyber security strategies and learn from them. The experiment of the cyber defense against phishing attacks is mentioned as an example of the cyber defense of individuals. The rules in Microsoft Outlook were used by filtering incoming email messages. The result is promising by stopping 88 % of phishing emails. The discussion and conclusion state that COVID-19 played a big role in the cyber security situation in countries to the analyzed documents.

1. Introduction

This review paper is an extension of the work originally presented in the 2021 Communication and Information Technologies Conference Proceedings [1].

The goal of the article is to analyze the situation in the field of cyber security, using the latest strategic documents of Central European countries, reports of Computer Emergency Response Teams (CERT), and The National Security Agencies (NSA).

The result of the analysis is somewhat pessimistic because in cyberspace we are the object of an ever-increasing number of cyber-attacks and technologically we lag behind the attackers. On the other hand, the paper mentioned a positive example of an effective defense against phishing email attacks.

Cyberspace is not limited by geographical boundaries, and its actual state can be characterized by continuing enlargement of cyber threats and cyber-attacks, whereas any of them are very serious and the form of cyber warfare is approaching.

Private companies and public organizations have to face permanently against cyber-attacks and deal with their consequences. International cooperation in cyber security is very important, countries participate by exchanging information about cyber threats and attacks and organizing joint exercises.

Cyberspace was recognized by NATO as a new domain of warfare in 2016; comparable to other domains: land, air, sea, and outer space. Appropriate policies, action plans, committees, and agencies have been adopted to ensure Member States' cyber security. Cyber headquarters and operational centers with relevant troops have been established and intensive preparations are underway against cyber threats and cyber-attacks.

The article is interesting in existing threats in cyberspace with regard to their danger and with a focus on cyber-attacks, associated with real military activities (aggression), especially from Russia.

The core of the article is the analysis of new strategic documents in cyber security in selected countries in order to look for their identical and different areas and obtain the necessary lessons learned for the development of better endurance in cyber security and preparation of suitable sources for education.

The analyzed documents often deal with phishing attacks and defense against them, because these attacks are usually at the beginning of larger cyber-attacks.

Phishing is a form of attack with the help of social engineering techniques, in which an attacker pretends to be a trusted authority in order to obtain sensitive data from the victim. The attacker thus often tries to gain the trust of the victim, who then actually communicates the necessary information or data voluntarily.

*Corresponding Author: Kamil Halouzka, kamil.halouzka@unob.cz

https://dx.doi.org/10.25046/aj070414

An example of a possible effective defense of individuals against phishing attacks is mentioned in an experiment about the application of Microsoft Outlook email client functions.

2. The Literature Review

The section presents selected comparative studies of national and international documents in cyber security and analyzes the ability of cyber defense. The analyzed studies are prepared in the form of a verbal and tabular description of the comparison, but suitable analytical procedures and models with the form of graphical outputs are also used. The aim is to identify the strengths and weaknesses of the analyzed documents and learn to improve their own approaches.

An extensive study [2] prepared for the Italian Parliament compares NATO member countries' approaches to cyber defense and finds that some countries have more proactive approaches (US, UK, France), while others have a more defensive approach (Germany, Spain). Although there are differences in the approaches to cyber defense, it must be seen that all nations are affected by NATO's unified regulatory and doctrinal framework, so that despite existing differences, national elements of cyber security can be integrated with the Alliance's command structure. The 2019 NATO Summit declared that due to the geopolitical activities of China and Russia, preparation for cyber defense had become a top priority. From the technical point of view, the NATO Communications and Information Agency (NCIA) provide capabilities necessary to the Alliance’s structures in terms of cyber defense. The NCIA administrates some of the allied networks with the NATO Cyber Security Centre (NCSC) and the NATO Computer Incident Response Capability (NCIRC). Finally, outside the NATO military command structure, the Cooperative Cyber Defence Centre of Excellence (CCDCOE) in Estonia, created in 2008, prepares reports and other documents in the field of cyber defense and, since 2010, hosts regularly cyber security exercises.

A paper [2] further analyzes the procedures and results in the cyber defense of individual NATO nations. The US approach to cyber defense is qualitatively and quantitatively different from that of most European countries. The National Security Strategy from 2017 underlines the cyber domain as one of the main future battlegrounds, and the 2018 Strategy warns against adversarial capabilities damaging American armed forces, economy, and society in cyberspace. The US Department of Defense established a Cyber Command (USCYBERCOM) in 2009, within the Strategic Command, whose commander is at the same time the Director of the National Security Agency (NSA), to ensure seamless cooperation between cyber and intelligence operations. The USCYBERCOM strategy is focused to:

1. Achieve and sustain capabilities, by anticipating technological changes and exploiting them faster and more effectively than the adversaries.
2. Create cyberspace capabilities to support operations in other warfare domains.
3. Ensure information superiority to achieve strategic impact.
4. Operationalize the cyber battlespace for agile maneuvers.
5. Expand and deepen partnerships with agencies, the private sector, and academia.

The conclusion stated that cyber defense at the NATO level is not limited to the creation of command structures and the employment of dedicated personnel, but also involves broader partnerships. The necessity of equipping NATO with cutting-edge technology led in 2014 to the formation of specific cooperation with industries operating in the cyber sector. NATO-EU cooperation was already listing the cyber dimension among priority areas of collaboration in 2016.

Analysis of the Polish National Cyber Security Strategy (NCSS) in comparison to EU strategy and regulations and to other NCSS of the countries (US, UK, France, Lithuania, and Estonia) is the content of the paper [3]. The definition of cyberspace is not enough clear and often is in the documents missing. An example is a definition by the US Department of Defense “A global domain and the information environment including networks, information technology infrastructure, data sources, telecommunications, the Internet, computers, and embedded systems”, cited in [4]. The EU security strategy in cyberspace, issued in 2013, clarified goals, responsibility roles, and tasks as achieving cyber resilience, preparing an EU Cyber Defence Policy and sources, reducing cybercrime, and developing needed technologies. In 2017, the EU published a cyber security document including initiatives in resilience to cyber-attacks and cyber security capacity, effective criminal law, and complex stability in international cooperation. The 2019 Cyber Security Act has provided a consolidated cyber security certification framework. The sanctions system Cyber Diplomacy Toolbox, allows the EU to impose targeted restrictive measures to prevent and respond to cyber-attacks. The European Parliament adopted 2021 a decision on the EU’s Cyber Security Strategy for the next digital Decade to make developed tools and services secure from the start of development, resilient to cyber threats, and able to quickly react if vulnerabilities are discovered. Poland has agreed in 2017 to its NCSS for 2017-2022 which defines cyber security as “The resilience of information and communication systems, at a given level of confidence, to any activity that compromises the availability, integrity, authenticity, or confidentiality of data, or the related services oriented by or accessible via these information networks and systems”. In 2018, Poland adopted the National Cyber Security System Act establishing coordination measures of state policy in the area of cyber security, and in 2019, in accordance to the European Union Agency for Network and Information Security (ENISA) lifecycle approach, Poland adopted its improved NCSS for 2019-2024. The final evaluation of the Polish NCSS contains recommendations for its further development, based on comparisons with the approaches of other countries.

Complex comparative analyses [5] of national cyber security strategies (NCSS) combine areas of industry, economy, technology, and defense. The study characterizes the NCSSs of countries US, UK, Japan, and EU, and describes cyber security agendas for the revision of NCSS in South Korea, by applying topic modeling. Topic modeling involves statistical techniques to identify hidden structures from a set of documents. The result is 15 agendas in the areas of Infra Stability, Protection and Response Capability, Industry and Technology, and International Cooperation. The NCSS of the US emphasized improving incident response capabilities, especially cybercrime law enforcement and investigation capabilities, and establishing cyber security governance. Similar to the US, the UK prioritized protection and response capability. On the other hand, the NCSS of Japan establishes a relatively high proportion to the cyber security industry and technology sector. Finally, the NCSS of the EU picks up international cooperation.

The article [6] is interesting in terms of currently ongoing Russian aggression in Ukraine, it includes a comparative analysis of cyber security systems in Russia and Armenia. In the introduction, the close cooperation between both countries in the economic, political, and military fields is appreciated. Two levels (legal and practical) are used to analyze cyber strategies, policies, and institutions. Key theoretical concepts in information security, information warfare, etc. are described. The cyber security definition is mentioned as "A set of technical and non-technical (policies, security arrangements, actions, guidelines, risk management) measures allowing to provide social, ethnic and cultural evolutionary modernization of the critical cyber infrastructure, as well as protection of vital interests of human, society, and state." The experiences from the military operation of Russia in August 2008 in South Ossetia and Georgia changed the Russian Defense Ministry's intention to create informational troops, whose functions include all aspects of information warfare, from psychological operations and propaganda to security of computer networks and cyber-attacks on the enemy's information systems. Cyberspace in Armenia is rather liberal. The principle is "allowing everything that is not prohibited" when prohibited are direct and clear criminal acts. Similar to the Russian approach, the Armenian side uses a wider concept of information security without specifying the concept of cyber security.

3. Analysis of selected cyber threats

The aim of the section is not to list every possible cyber threat. A brief introduction into cyber threats was presented in [1]. The initial point of an intrusion into a system by attackers is a very often successful phishing campaign. Therefore, the aim of the section is to analyze phishing from several points of view, especially phishing associated with real military actions.

Phishing campaigns usually have the following aims:

1. They try to lure sensitive data from end-users and misuse obtained data. It can be personal data, corporate data, or governmental data.
2. The goal is to infect a computer for later abuse by adding malicious attachments to e-mails or creating links leading to fake malicious websites, thus computers may become part of a botnet, be infected with ransomware, contain a key logger, etc.

In addition to classical phishing attacks that target as many victims as possible, there are also other types of phishing [7]:

- Spear phishing attempts are targeted toward specific individuals or groups of individuals. They may include the recipient's name, position, or company. IT administrators can be great targets because of the level of access they usually have within the organization.
- Whaling targets high-level employees, like executives or directors. They typically have access to the most valuable information in a company, making them appealing targets for attackers.
- Clone phishing is typically targeted at a small group of people. Attackers copy a legitimate email that has previously been sent by a trusted organization but replace links to redirect the victim to a malicious site.

Phishing attacks usually take place over email but attacks using other mediums have also been observed. Smishing is the text message (SMS) version of phishing attacks. Vishing is phishing that is executed via telephone (Voice).

Phishers usually exploit three types of events. Firstly, phishing campaigns misuse long-term affairs such as humanitarian aid to countries affected by perpetual fighting, or an education for African children. Secondly, regularly occurring events are commonly exploited by phishers, e.g., holidays, summits (EU, NATO), or elections. Thirdly, phishers take advantage of current events, e.g., outbreaks of fighting, rapid changes in the financial market. A common characteristic of phishing is the feeling which is intended to evoke in people. This effect encourages users to react, and as a consequence, it increases the effectiveness of phishing. The feelings they usually evoke can be: sympathy, fear, joy from prize, urgency, stress, and patriotism.

Patriotic or nationalist hackers see themselves as irregular soldiers, or conscripts fighting a war for their country, a form of cyber militia. Their attacks are motivated by strong feelings of patriotism and nationalism, reflected in the language and rhetoric used. The actions of the patriotic hacker may result in serious damage to targeted systems [8].

In the case of intensified Russian military activities there are 2 possible vectors of cyber-attacks:

1. From Russia - attacking information systems (IS) of the enemy and enemy sympathizers.
2. Against Russia - attacking IS of Russia and Russian sympathizers.
3. From all around the world - attacking IS of adversaries and their sympathizers, depending on which side the attackers are inclined towards.

Phishing campaigns related to current events in the context of Russian activities may use the terms humanitarian aid, solidarity, support for the fight, signing petitions, and providing accommodation to refugees. The impact of such phishing campaigns tends to be personal or sensitive data leakage, payment card details leakage, subsequent misuse of the leaked data, and payments to the attackers' accounts instead of accounts for humanitarian purposes, computers infected with malicious codes, and their subsequent abuse for adversary activities.

Phishing can also contain fake actual news. It aims to manipulate people across the board and significantly influence their behavior. This method is often used in state information operations to weaken an adversary or, on the contrary, to strengthen the confidence of its own population in the government aggressive activities. Such phishing messages include fake news of invasion, troop movements, shortages of goods in shops, fuel shortages, shutdown of gas supplies by Russia, power cuts due to gas shortages, etc. On the other hand, fakes news about military successes, humane treatment of the enemy, liberation of the population from oppression and other justifications for fighting are then used to influence the confidence of the population in the governmental aggressive decisions.

In a widespread phishing campaign against an adversary, the consequences can be immeasurable - population panic, buying frenzies, and stockpiling of resources. With an unexpectedly height public reactions, we can expect repercussions in all sectors. Particularly for the communication and information area, we can expect disruptions or even unavailability of services or entire information systems, e.g., disruption of electronic banking, unavailability of telephone lines, unavailability of web information portals, and unavailability of bank card payments. Additionally, if the critical infrastructure is affected, the functioning of the state and human lives may be endangered.

Another frequent impact of successful phishing can be a ransomware infection or a data leakage. Both encrypted data as well as data leakage are common subject of a ransom. It is not recommended to pay the ransom due to uncertainty as to whether the attackers keep their promises. In case of encrypted data there is a risk the attackers will not send decrypting keys after payment or the keys will not work properly. In case of the data leakage there is a risk the attackers will sell stolen data to a third party after payment anyway. A new approach to ransomware has emerged. Anybody can buy ransomware as a service for a fee or a shared ransom. Effects of ransomware on information systems range from denial of service, and data loss, to loss of reputation, and bankruptcy. The consequences of an infection in the critical infrastructure are immense, as it was mentioned above.

A new and additional set of phishing related to Russian military activities and the current situation in Russia and Ukraine is emerging. This new set has specific keywords in conjunction with the words Russia or Ukraine, e.g., solidarity, refugees, war victims, aid, fundraising, petition, cohesion, but also shortages in connection with the words gas, wheat, energy, building materials, etc. Such keywords can be effectively used in individual defense against phishing attacks which is described in section 5.

There are two main approaches to reduce a rate of success phishing. From a technical point of view we can increase detection capabilities. From a non-technical point of view we can reduce a phishing risk by spreading security awareness.

Firstly, technical measures work reliably but they need to be updated and adapted regularly. It is very difficult to respond adequately to rapidly changing links in phishing emails. Links in emails are changed by attackers faster than it would be possible to manually respond to them in real time. Manually adding malicious links into blacklists is an inefficient human-consuming and time-consuming method. A more appropriate method is to automatically and regularly download updated blacklists from selected trusted sources that deal with this issue.

Blocking selected file types (.dll, .exe, .js, .msi, .reg...) is a very effective method. However, if we block files types inappropriately, users will not receive their attachments needed for their work. For example, .pdf files may contain links to malicious sites, but it is not possible to block all .pdf attachments in bulk. The solution is a technology that runs selected suspicious attachments in a sandbox. The technology automatically evaluates the behavior of the system after the attachment is launched and if everything works normally, the email is sent to the user's mailbox. If malicious behavior is found, the user is informed about the situation and the email is sent with a modified attachment, for example an attachment is converted from a .pdf file to an image (.png, .jpeg...). This prevents the user from clicking on the malicious link or executing the malicious code, but the information from the attachment is still delivered to the user.

Another technical measure is blocking emails based on sets of keywords. The sets must be chosen with caution, taking into consideration the possibility of a high number of false positives. In long-term tuning, this method achieves a very good level of reliability. The practical application of this technical measure is discussed in section 5.

Secondly, there are many methods of security awareness spreading. For example, users can attend specialized seminars and courses to learn how to recognize phishing. However, these trainings are usually only once a year due to financial and time constraints. In addition, this approach is not proving to be as effective as expected. In practice, short training sessions followed by testing employees with mock phishing is more effective method. Metrics such as the number of tricked users, the most clicked links, types of attachments launched, and, of course, leaked passwords are analyzed and final reports are published within the organization. Modified mock phishing emails are then repeated at random time intervals and repeat clickers are then invited to additional trainings. According to [9], it is advisable to increase the difficulty of mock phishing up to 3 levels:

- Tier 01 – generic type of mass phishing attacks. Emails include misspellings, poor graphic design, and well-known scams.
- Tier 02 – more professional or more personalized phishing attacks. Emails contain victim’s name and are work related.
- Tier 03 – targeted attack and customized for selected high-risk groups.

The best results for reducing the success rate of phishing attacks are achieved through a combination of technical measures and educated users. It depends on the availability of resources, how many security technical tools are used and to what extent users are trained and tested.

4. Cyber Security Status Report

The aim of the next section is to analyze the cyber security status reports of the Czech Republic, Slovakia, Austria, Germany, and Poland for the year 2020. Cyber security status reports are issued by these countries in the second half of the year. For this reason, the reports for 2020 were prepared. The reports for 2021 were not available at the time of writing. The aim is to assess the problems that each country has in the area of cyber security, to evaluate the frequency of cyber-attacks or incidents, and to evaluate the most common types of attacks.

4.1. Czech Republic

A report on the state of cyber security in the Czech Republic has been published by the National Cyber and Information Security Agency (NÚKIB) [10]. The year 2020 in the Czech Republic was characterized by an increase in the number of cyber-attacks against Czech institutions, organizations, and companies in all sectors. In 2020, the NÚKIB recorded a more than double increase in incidents compared to 2019. The most common types of attacks in the Czech Republic in 2020 were spam (59%), phishing (16%), and scanning (12%). Respondents ranked ransomware (19%), DoS / DDoS attacks (19%), spear-phishing emails (14%) and attempts to exploit vulnerabilities (13%) as the most serious attacks. Cybercrime has long been among the most serious threats to the country's cyber security. In 2020, cybercrime emerged in the form of ransomware attacks, which hit the Czech healthcare sector to a large extent.

The most serious threats to the Czech Republic's cyber security include state-sponsored actors in cyberspace and cybercrime. A new development is Ransomware as a Service (RaaS), which is a service provided by ransomware developers to other hackers, usually for a share of the ransom, and they do not care about the actual penetration of organizations' systems.

In terms of personnel and financial security, a large number of organizations in the country have been facing a lack of experts and insufficient budgets in the field of cyber security. This situation was more evident in the government sector than in private companies. Almost none of the interviewed organizations had all cyber security positions filled. More than half of the organizations cited inadequate salary conditions as the main factor.

In terms of training, the NÚKIB placed a strong emphasis on the training of state administration employees and trained more than 22,000 state administration employees, employees of the Army of the Czech Republic, and medical and prevention personnel from the education sector in e-learning courses during 2020.

In the report on the state of cyber security in the Czech Republic, much attention is paid to the security of 5G networks. In 2020, the Prague 5G Security Conference was organized jointly with the Office of the Government and the Ministry of Foreign Affairs, the main topic was the risks associated with building 5G infrastructure. The main outcome was the presentation and the launch of the Prague 5G Security Repository, a virtual library designed to share legislative, strategic and other tools that states adopted in the past year in the area of 5G network security.

4.2. Slovak Republic

A report on the state of cyber security in Slovakia has been published by the National Security Authority of Slovakia [11]. The report is divided into seven main parts. The focus is on the description of actors in cyberspace, namely inexperienced attackers (script-kiddies), cybercriminals motivated by financial gain, state-sponsored groups, hacktivist groups interested in obtaining sensitive and classified state information, and cyberterrorists, whose role is mainly to hit civilian and military targets with cyber-attacks.

Furthermore, the report focuses on the categories of cyber security incidents (monitored by SK-CERT) and threats. The most detected and reported incidents were in the "Unwanted Content" category and the most solved incidents were in the "Attempted Intrusion" category. The most frequent threats in Slovakia were phishing campaigns (Microsoft tech support scam, abuse of Slovak Post), malicious code distribution (mainly ransomware – the attack on Slovak TV Senzi, which refused to negotiate with the attackers and filed a criminal complaint), data leaks (leak of 130,000 personal data of patients tested on COVID-19) and vulnerability exploitation.

The Slovak cyber security status report also goes into detail on cyber threats targeting specific organizations, such as healthcare, public administration, banking, electronic communications, and digital infrastructure, and energy. The report pays great attention to, among other things, issues related to national and European legislation, the preparation of the National Cyber Security Strategy, national and international activities and cooperation, cyber security audits, and cyber defense exercises (Table-Top exercise BlueOLEx 2020 and Cyber Coalition 2020). Among other important security actions that were implemented in Slovakia was the establishment of the Competence and Certification Centre for Cyber Security. The aim of this center is to assist the National Security Authority in fulfilling its professional tasks in the field of cyber security, protection of classified information and cryptographic protection, and trust services in the public interest.

4.3. Austria

The Cyber Security Status Report was prepared by the Cyber Security Steering Group (CSS) in accordance with the Austrian Cyber Security Strategy (ÖSCS) [12]. As in the above-mentioned countries, the number of malware attacks on computer systems and networks in Austria increased over the past year. A large part of Austria's annual report was devoted to the impact of the SARS-CoV-2 pandemic on cyber security. At the beginning of the pandemic, many companies were forced to change to a home office for which they were unfortunately not prepared. Companies were often forced to reduce their own cyber security to allow their employees to work away from their offices.

There has been a sharp increase in the number of fraudulent sites, seemingly related to Covid-19, designed to phish or spread malware. The authors of these scams demanded from victims to pay them USD 4 000 in bitcoins. If they didn't pay, they were told their families would be infected with the coronavirus. Other frauds involved changing delivery times for shipments due to the pandemic. Opening the link and/or file contained in the message caused the installation of malware (including AZORuIt, Emotet, Nanocore RAT and Trick-Bot) on the target computer. There were major problems in Austria with data leaks from corporate computer networks, where attackers demanded a ransom for its return. Several waves of DDoS attacks occurred during the reporting period, mainly against banks, the financial sector and Internet Service Providers (ISPs). The aim of these attacks was not only to deny services but also to blackmail their victims.

A large part of the report is devoted to cyber security cooperation between Austria and European Union, United Nations, NATO, and other important committees and forums. Equal attention is paid to clarifying national cyber security actors such as Cyber Security Centre, Cybercrime Competence Centre, CIS and Cyber Security Centre, Austrian Armed Forces Security Agency, and many others.

The Austrian Cyber Security Status Report is the only one that does not provide any specific numbers of cyber security breaches, as all values were given only as percentages.

4.4. Poland

A report on the state of cyber security in Poland has been published by CERT (Computer Emergency Response Team) Poland [13]. In the period under review, 60.7% more cyber security attacks were registered than in the previous year. The most common type of incident was phishing, which represented 73% of all cyber-attacks. In March 2020, Poland released a list of dangerous websites (List of warnings), which had a significant impact on the number of phishing attacks recorded. These phishing attacks were targeted at obtaining e-banking authentication details, payment card details, email account access details, and social media accounts.

Cybercriminals used, for example, Facebook messages with sensationalist headlines, fake SMS messages, and WhatsApp messages for this purpose. There also was an increase in unwanted messages (spam) on mobile platforms (especially Android). CERT Polska focused on analyzing IP addresses localized in Poland that were used for Distributed Reflective Denial of Service (DRDoS) attacks. For that purpose, a list of poorly configured services that were the most frequently used for DRDoS attacks was published.

As in previous years, disinformation campaigns related to attacks on information portals and accounts of Polish politicians were recorded in Poland. Criminals used the accounts to publish fake information aimed at, for example, reducing the trust of public officials or spreading negative information about the US military in Poland.

The report on the state of cyber security in Poland also contains a large number of practical examples (in text and graphic form) of the most common form of malicious software distribution, fake job offers on Facebook, fake parcel post services, fake bills for the advertisement and others. The extensive chapter is completed with recommendations on how to avoid infection.

4.5. Germany

In Germany, the Federal Office for Information Security (BSI) monitors IT security threats [14]. According to the report on the state of cyber security in Germany, the trend of attackers using malware to launch mass cyber-attacks continued during the period. The malware was the most commonly used attack and was responsible for cyber-attacks on individuals, private companies, government offices, and other institutions. The malware was also used to launch targeted attacks against pre-selected victims.

A large amount of personal data was also leaked, which unfortunately also included data on patients and their clinical records (this data was unfortunately freely accessible online). The reporting period also showed the emergence of a number of vulnerabilities in software products that attackers were able to exploit to spread malware, attack or steal data. Some of these vulnerabilities were assessed as critical.

There were targeted attacks on financially powerful victims such as car manufacturers and their suppliers, attacks on airports and airlines, and on less known high-income companies. Attackers also used increasingly the "human factor" as a starting point for attacks that use social engineering to gain an entry point for further attacks. Also in Germany, the COVID-19 pandemic was often used for cyber-attacks.

One example was the large-scale waves of spam offering fake advice about the coronavirus. These emails urged company employees to post personal or company information on copies of official websites. Cybercriminals designed these sites similar to (government) websites.

One of the biggest threats mentioned in the annual report was Emotet. Emotet was used to create a cascade of other malware attacks that can culminate in targeted ransomware attacks on selected, usually wealthy victims. Critical BlueKeep and DejaBlue vulnerabilities in Windows Remote Desktop Protocol were also published. This vulnerability allowed attackers to execute malicious code, including malware, on unpatched systems.

4.6. Brief comparison of cyber security reports

As noted, each of these states issues an annual report on the state of cyber security. Each state publishes information in this report at its own discretion, and there is generally no rule specifying what the report must contain. The following table (Tab. 1) shows a basic comparison of the cyber security status reports of these states, with information from the United States and the United Kingdom reports added for comparison.

While for the UK the Annual Review 2020 report [15] issued by The National Cyber Security Centre was used for comparison, in the US two reports were used, namely the NSA cybersecurity year in review for 2020 [16] and the FISMA FY 2020 Annual Report containing The State of Federal Cybersecurity [17]. In the case of the United States, both reports are very general and contain little specific information compared to other reports.

Tab. 1 shows who is responsible for the policy/strategy in the countries listed, when the first National Cyber Security Strategy (NCSS) was issued, and when the last update was made or the period of validity. The table also shows important aspects of each cyber security status report, when the focus was on new IT security measures in organizations and the status of security-focused budgets, which were heavily affected by the COVID-19 pandemic and the resulting increase in security and budgetary measures. Cyber security Status Report lists the most frequent incidents during the reporting period, where the most popular type of incident was phishing. The second most reported incident type was ransomware. There was also mentioned a large increase in malware on mobile platforms in 2020. An important part of cyber preparedness is also participation in international cyber exercises, which have also been affected by the COVID-19 pandemic. Even though the Locked Shields 2020 and Cyber Europe 2020 international exercises were canceled in 2020 due to the pandemic, individual countries have organized several cyber security exercises that are important for training defense against cyber-attacks. The table also shows that national security authorities support cyber security education in the form of cooperation with schools (e.g. cyber security information cards for schools), business (e.g. various levels of consulting for business customers), and voluntary sector (e.g. guidance on cyber security).

Table 1: Basic comparison of the cyber security status reports

	USA	GBR	Germany	Polland	Austria	Slovakia	Czech Republic
Policy/strategy responsibility	CISA	NCSC	BMI	MDA	BMEIA	NBÚ	NÚKIB
First NCSS	2003	2011	2011	2009	2009	2008	2012
Actual NCSS	2018	2022 - 2030	2021	2019 - 2024	2021	2021 - 2025	2021 - 2025
New IT security measures in companies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Security budget in companies	Not mentioned	Increase	Not mentioned	Increase	Increase	Increase	Increase
The most frequent incident reported during the reporting period	Spear-phishing	Phishing	It's not specified. The report generally mentions malware and phishing.	Phishing	Phishing	Phishing	Spam
The second most frequent incident reported during the reporting period	Not mentioned	Ransomware		Ransomware	Not mentioned	Ransomware	Phishing
Participation in cyber exercises	Not mentioned	Not mentioned	Yes / Crossed Swords 2020, Common Roof	Yes / KSC-EXE 2020, Capture The Flag, Hack-A-Sat	Yes / Crossed Swords 2020, Common Roof	Yes / BlueOLEx, Cyber Coalition	Yes / Cyber Coalition
Cyber security education in companies	Yes / improving	Yes / improving	Yes / improving	Yes / improving	Yes / improving	Yes / improving	Yes / improving
Cyber security incidents in 2020	30819	723	419 (mentioned only critical infrastructure)	10420	Not mentioned	3793	1267

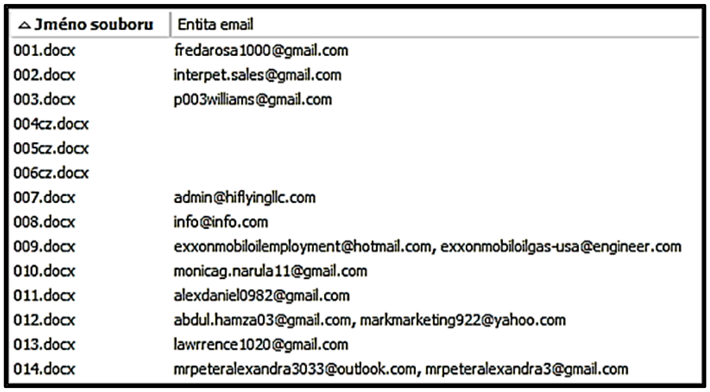
The latest comparison is on the number of cyber incidents in 2020. It is difficult to compare the number of reported cyber incidents from the cyber security status reports of these states because each state reported these cyber incidents in the form of the most frequent, most serious, reported or resolved cyber incidents. For this reason, a clear comparison of the number of cyber incidents is not possible. Austria, for example, did not mention the number of cyber incidents at all in its cyber security report.

5. Cyber Defense of the Individual

This part of the article focuses on the cyber defense of individuals against phishing attacks. The result of research on the phishing attacks in the previous two years has collected the set of almost 400 phishing emails that were sent to the email inbox of one of the authors. The emails were the subject of analysis, the necessary knowledge was obtained, and that was used for the cyber defense of individuals. The functions of the mail client MS Outlook were used for individual defense.

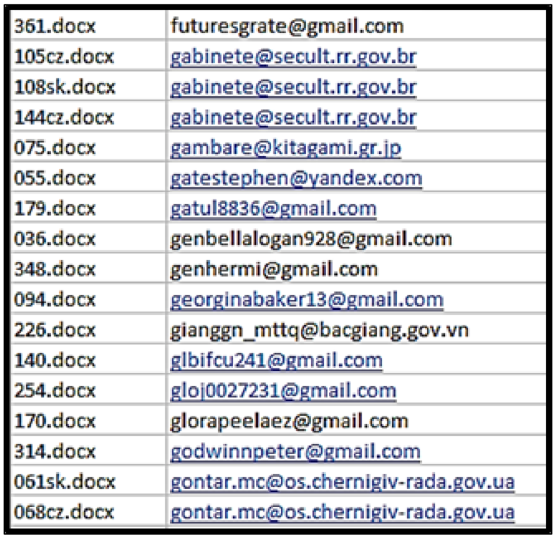
One of the options for individual cyber defense against phishing emails is the possibility to block all email addresses detected from the content of phishing email messages. Email addresses were obtained from phishing emails using the intelligent function entities extraction of the Tovek [18] software, using the Tovek Agent module, see Fig. 1. The first column in Fig. 1 is the file name (phishing email) and the second column includes email addresses.

A total of 365 emails were analyzed, from which 511 email addresses were extracted, of which only 17 were reused; i.e., 3%. It is obvious that blocking that volume of addresses with such low efficiency would not be effective in terms of protection against phishing attacks, see Fig. 2; normalized email addresses in the Excel table, and ordered.



△ Jméno souboru	Entita email
001.docx	fredarosa1000@gmail.com
002.docx	interpet.sales@gmail.com
003.docx	p003williams@gmail.com
004cz.docx	
005cz.docx	
006cz.docx	
007.docx	admin@hiflyingllc.com
008.docx	info@info.com
009.docx	exxonmobiloilemployment@hotmail.com, exxonmobiloilgas-usa@engineer.com
010.docx	monicag.narula11@gmail.com
011.docx	alexdaniel0982@gmail.com
012.docx	abdul.hamza03@gmail.com, markmarketing922@yahoo.com
013.docx	lawrrence1020@gmail.com
014.docx	mrpeteralexandra3033@outlook.com, mrpeteralexandra3@gmail.com

Figure 1: Entity extraction



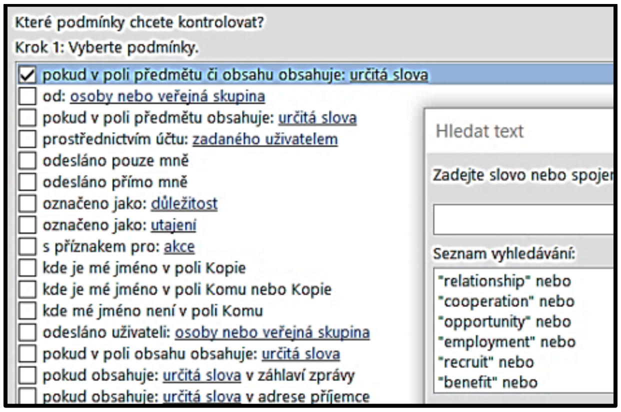
361.docx	futuresgrate@gmail.com
105cz.docx	gabinete@secult.rr.gov.br
108sk.docx	gabinete@secult.rr.gov.br
144cz.docx	gabinete@secult.rr.gov.br
075.docx	gambare@kitagami.gr.jp
055.docx	gatestephen@yandex.com
179.docx	gatul8836@gmail.com
036.docx	genbellalogan928@gmail.com
348.docx	genhermi@gmail.com
094.docx	georginabaker13@gmail.com
226.docx	gianggn_mttq@bacgiang.gov.vn
140.docx	glbifcu241@gmail.com
254.docx	gloj0027231@gmail.com
170.docx	glorapeelaez@gmail.com
314.docx	godwinnpeter@gmail.com
061sk.docx	gontar.mc@os.chernigiv-rada.gov.ua
068cz.docx	gontar.mc@os.chernigiv-rada.gov.ua

Figure 2: Extracted email addresses

Another possibility of individual cyber defense against phishing emails is the use of rules in the MS Outlook client to filter incoming emails. We have chosen rules for email detection based on the occurrence of keywords in the subject and content of the email message. The keywords were taken from the results of research on phishing emails, published in the article [19]. Keywords that characterize the particular email segment:

- BUSINESS (investment, project, contract, business, intention, export, service, product, partner, relationship, cooperation, employment, recruit, benefit, inquiry);
- FUND (deposit, fraud, scam, credit, compensation, inheritance, award, sum, price, prize, claim, winning);
- TRANSFER (transfer, property, gold, diamonds, box, packet, shipment, airport, bank, payment);
- CHARITY (charity, donate, Christ, God, sister, brother, promise, illness, disease, hospital, cancer, widow);
- OTHERS (offer, loan, communicate, message, response, contact, friendship, package, undelivered).

The rule is easily set up using the wizard in MS Outlook see Fig. 3. An overview of the rules is shown in Fig. 4.



Které podmínky chcete kontrolovat?

Krok 1: Vyberte podmínky.

☑ pokud v poli předmětu či obsahu obsahuje: určitá slova
☐ od: osoby nebo veřejná skupina
☐ pokud v poli předmětu obsahuje: určitá slova
☐ prostřednictvím účtu: zadaného uživatelem
☐ odesláno pouze mně
☐ odesláno přímo mně
☐ označeno jako: důležitost
☐ označeno jako: utajení
☐ s příznakem pro: akce
☐ kde je mé jméno v poli Kopie
☐ kde je mé jméno v poli Komu nebo Kopie
☐ kde mé jméno není v poli Komu
☐ odesláno uživateli: osoby nebo veřejná skupina
☐ pokud v poli obsahu obsahuje: určitá slova
☐ pokud obsahuje: určitá slova v záhlaví zprávy
☐ pokud obsahuje: určitá slova v adrese příjemce

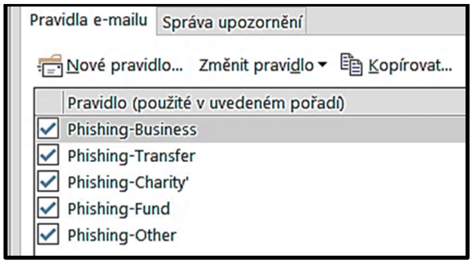
Hledat text

Zadejte slovo nebo spojen

Seznam vyhledávání:

"relationship" nebo
"cooperation" nebo
"opportunity" nebo
"employment" nebo
"recruit" nebo
"benefit" nebo

Figure 3: Setting email message filtering rules



Pravidlo (použité v uvedeném pořadí)

☑ Phishing-Business
☑ Phishing-Transfer
☑ Phishing-Charity'
☑ Phishing-Fund
☑ Phishing-Other

Figure 4: Overview of email message filtering rules

During the debugging of how the rules work successfully, conditions in the rules were continuously updated, on the basis of which emails were not excluded from phishing (manuscript, conference, journal, editor, publication, System NEWS, Smart Cities, Computerworld, Reuters, webcast, Deloitte, identity, Sophos). These are keywords associated with activities associated with publishing scientific articles and offering professional events. The phishing messages are stored in a dedicated folder; based on the set rules. The testing phase was realized for two weeks and after that followed the four weeks experiment.

The results of the experiment were evaluated, see Tab. 2. There were recognized two types of mistakes:

1. Uncaptured (undetected) phishing email (false negative).
2. Phishing erroneously detected, means the correct email that was included in the Phishing folder (false positive).

All seven false-negative messages were phishing, of which 2 in Czech. Content included in groups OTHERS-3, FOND-2, BUSINESS-1, TRANSFER-1, and CHARITY-0.

Table 2: Statistics of the experiment

Phish-email / Week	1	2	3	4	Σ
Total number	27	36	31	26	120
-of which undetected	1	2	2	2	7
Erroneously detected	1	3	1	2	7
Total detection errors	2	5	3	4	14

On the contrary, none of the false-positive messages was phish; they contained corporate information about the event or information about some facts. They contained some of the keywords for inclusion in the phishing folder and then included necessary changes in phish detection rules.

It is also worth mentioning that the number of phishing emails in the experiment increased, compared to the previous ones, by almost 50%. The result of the experiment is promising, correctly captured phishing emails account for 88%.

The described method of defense against phishing attacks can be recommended because it can be individually customized and is easily adjustable. It should be noted that phishing emails are an individual matter. Several years of experimentation have confirmed that the content and extent of phishing attacks against the same person vary a little.

6. Discussion and Conclusion

Cyber security is unfortunately a much-used term these days, and because of the current political situation (Russia-Ukraine conflict), it is not expected that there will be any return to the good old days. The information published in this article is not yet influenced by this conflict, but the main role in it is affected by the Covid-19 pandemic. This pandemic has greatly affected the behavior not only of attackers in the Internet world but also of ordinary users who have started to take the problem of cyber security at least a little bit seriously.

It is interesting to see how the countries mentioned in this article declare their cyber security challenges, and it is important to mention that a detailed comparison of the annual reports was almost impossible in terms of frequency of cyber incidents, structure, and content, as each national report had a different way of assessment. Comparable information was aggregated into Table 1, which shows information regarding the number of cyber incidents, where the United States clearly dominated, as well as the most frequent incidents encountered by those states, and information regarding support for training and cyber exercises.

Individual reports described each type of attack, with one of the most common being ransomware attacks, which caused hundreds of millions of euros in damage to various state and non-state institutions during the pandemic. Cyber-attacks have most often targeted critical infrastructure, the public sector, the financial sector, industry, healthcare, and, unfortunately, education.

As already mentioned, COVID-19 played a big role in cyber security in 2020, clearly showing how adaptable cybercriminals are. The most common methods of cyber-attacks that were recorded during COVID-19, according to the annual reports of the mentioned countries, were the following attacks:

- Fake news or links to fraudulent sites with disinformation such as there is a miracle cure for COVID-19.
- Fake messages or phone calls pretending to be from companies like Microsoft or Google Drive. Their goal was to extort a password from the user under the premise of offering help or threatening to cancel the account.
- Messages about non-existent packages being delivered.
- Fake appeals to donate money.
- Emails that appear to be from a medical organization.

Subsequently, the biggest threats according to the annual reports include:

- Ransomware - demanding a ransom to recover encrypted data.
- Threats associated with personal data - data breaches/leaks;
- Malware - malicious programs.
- Disinformation - spreading misleading or false information.
- Harmless threats - human error and system misconfiguration.
- Availability and integrity threats - attacks that prevent system users from accessing their information.
- Threats related to electronic mail - e-mail attacks.
- Supply chain threats - attacks (e.g. on service providers) to gain access to customer data.

The main contribution is analysis of cyber security documents, their comparison and evaluation in individual areas. There is a benefit for learning and developing a perspective on security requirements that follow from the findings of the analysis. An experiment was carried out on the subject of phishing, which, based on the use of MS Outlook rules, was able to significantly reduce phishing emails. This approach is original.

Acknowledgment

The article represents the results of the research in cyber security of the project [20] at the University of Defence in Brno, Faculty of Military Technology, Department of Informatics and Cyber Operations.

References

[1] K. Halouzka, L. Burita, P. Kozak, "Overview of Cyber Threats in Central European Countries," in 2021 Communication and Information Technologies Conference Proceedings, Liptovsky Mikulas, Armed Forces Academy of General Milan Rastislav Stefanik in Liptovsky Mikulas, 81–86, 2021.

[2] A. Marrone, E. Sabatino, "La difesa cibernetica nei Paesi NATO: modelli a confront, Cyber Defence in NATO Countries: Comparing Models," Rome, Senate, 2020, [Online]. Available: http://www.parlamento.it/documenti/repository/affariinternazionali/osservatorio/approfondimenti/PI0164.pdf.

[3] A. Jacuch, "Comparative analysis of cyber security strategies, European union strategy and policies. Polish and selected countries strategies", Online Journal Modelling the New Europe, **37**, 102–120, 2021.

[4] US Congressional Research Service, Defence Primer: Cyberspace Operations, https://sgp.fas.org/crs/natsec/IF10537.pdf.

[5] M. Song, D.H. Kim, S. Bae, S.-J. Kim, "Comparative Analysis of National Cyber Security Strategies using Topic Modelling," in International Journal of Advanced Computer Science and Applications, **12**, 62–69, 2021, doi:10.14569/IJACSA.2021.0121209.

[6] R. Elamiryan, R. Bolgov, "Comparative analysis of cyber security systems in Russia and Armenia: Legal and political frameworks," in Digital Transformation and Global Society, **858**, 195–209, 2018, doi:10.1007/978-3-030-02843-5_16.

[7] J. Fulmer, "Complete Guide to Phishing Attacks: What Are the Different Types and Defenses?," 2022, [Online]. Available: https://www.esecurityplanet.com/threats/phishing-attacks/.

[8] M. Dahan, "Hacking for the homeland: patriotic hackers versus hacktivists," in ICIW 2013 Proceedings of the 8th International Conference on Information Warfare and Security, 51–57, 2013.

[9] SANS, "Phishing Strategic Planning Document," 2022.

[10] NÚKIB, "Zpráva o stavu kybernetické bezpečnosti za rok 2020," 2021.

[11] NBÚ, SK CERT, "Správa o kybernetické bezpečnosti v Slovenskej Republice za rok 2020", 2021.

[12] Cybersecurity Report 2020, Austria, Vienna, 2021.

[13] NASK PIB/CERT Polska, "Security landscape of the Polish Internet, Annual report from the actions of CERT Polska 2020," 2021.

[14] Federal Office for Information Security, "The State of IT Security in Germany in 2020," 2021.

[15] National Cyber Security Centre, Annual Review 2020, UK, 2021.

[16] NSA United States of America, NSA cybersecurity year in review for 2020, 2021.

[17] FISMA FY 2020 Annual Report to Congres, The State of Federal Cybersecurity, Federal Information Security Modernization Act of 2014, 2021.

[18] Tovek, "The text analytical software TOVEK," 2022, [Online]. Available: https://www.tovek.cz.

[19] L. Burita, P. Matoulek, K. Halouzka, P. Kozak, "Analysis of phishing emails," in AIMS Electronics and Electrical Engineering, **5**(1), 93–116, 2021.

[20] DZRO FVT 2_KYBERSILY, Research project Cyber forces and resources, University of Defence, Faculty of Military Technology, Brno, Czech Republic, 2021.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 7, No. 4, 114-120 (2022)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Estimating a Minimum Embedding Dimension by False Nearest Neighbors Method without an Arbitrary Threshold

Kohki Nakane[1,*], Akihiro Sugiura[2], Hiroki Takada[1]

[1]*Graduate School of Engineering, University of Fukui, Fukui City, 910-8507, Japan*

[2]*Department of Radiological Technology, Gifu University of Medical Science, Seki City, 501-3892, Japan*

ARTICLE INFO

Article history:
Received: 30 May, 2022
Accepted: 10 August, 2022
Online: 24 August, 2022

Keywords:
False Nearest Neighbors
Attractor
Chaos
Van der Pol Equation
Lorenz Equation
Electrooculogram

ABSTRACT

The false nearest neighbors (FNN) method estimates the variables of a system by sequentially embedding a time series into a higher-dimensional delay coordinate system and finding an embedding dimension in which the neighborhood of the delay coordinate vector in the lower dimension does not extend into the higher, that is, a dimension in which no false neighbors or neighborhoods exist. However, the FNN method requires an arbitrary threshold value to distinguish false neighborhoods, which must be considered each time for each time series to be analyzed. In this study, we propose a robust method to estimate the minimum embedding dimension, which eliminates the arbitrariness of threshold selection. We applied the proposed approach to the van der Pol and Lorenz equations as representative examples of chaotic time series. The results verified the accuracy of the proposed variable estimation method, which showed a lower error rate compared to the minimum dimension estimates for most of the thresholding intervals set by the FNN method.

1. Introduction

In our previous studies, we have studied the development of mathematical models and feature extraction using artificial intelligence for biological signals. We simulated biological signals using GAN [1,2], a deep-learning generative model. Then, from the trained generative model, we have statistically analyzed a huge number of optimized parameters and investigated the possibility of extracting new features from the biological signals that have not been discovered so far. Then, since deep learning handles a large number of parameters, the generated model can be said to be a multivariable-dependent system. On the other hand, it has been pointed out that the biological signals to be generated are not multivariable in nature, but may be relatively simple systems that depend on a few variables at most [3,4]. To investigate the variables on which the system depends, several methods have been proposed to estimate the minimum embedding dimension of the time series using attractors.

The false nearest neighbor (FNN) method [5–7] estimates the variables of time series obtained from dynamical systems. The concept of false neighbors may be understood more clearly by reference to the Lorenz equation [8], as shown in Figures 1-3. A and B in the 1D attractor are no longer close to each other with increasing dimensionality of the attractor [9]. In the FNN method, the minimum embedding dimension is obtained using the following four steps. (i) Sequentially embed the time series into a higher-dimensional delay coordinate system. (ii) Compare the neighborhood distance of the lower-dimensional delay coordinate vector with that of the next higher-dimensional delay coordinate vector. (iii) If the rate of change of the neighborhood distance of the delay coordinate vector with increasing dimensionality of the delay coordinate system exceeds a given threshold value, the vector is considered a false neighborhood. (iv) The minimum embedding dimension is that in which all delay coordinate vectors are not false neighbors for the first time with increasing embedding dimensions.

The primary advantage of this method is that it uses all samples and thus is not an estimator, and is computationally fast owing to the simplicity of the necessary calculations. However, the FNN method requires an arbitrary threshold value to determine whether

[*]Corresponding Author: Kohki Nakane, kohki.pc@gmail.com

https://dx.doi.org/10.25046/aj070415

neighborhoods are false. Although observed time series are typically mixed with noise, the FNN method requires the selection of a threshold value depending on the level of noise, even for a single series. That is, the value of this threshold must be considered for each time series to be analyzed. In this study, we propose a robust minimum embedding dimension estimation method that eliminates this arbitrariness in threshold selection.

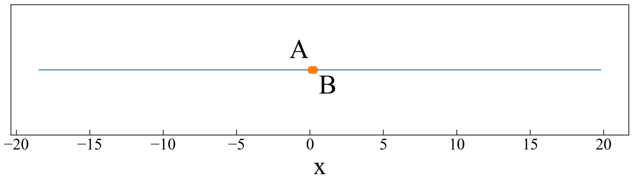


Figure 1: 1D attractor for the Lorenz equation

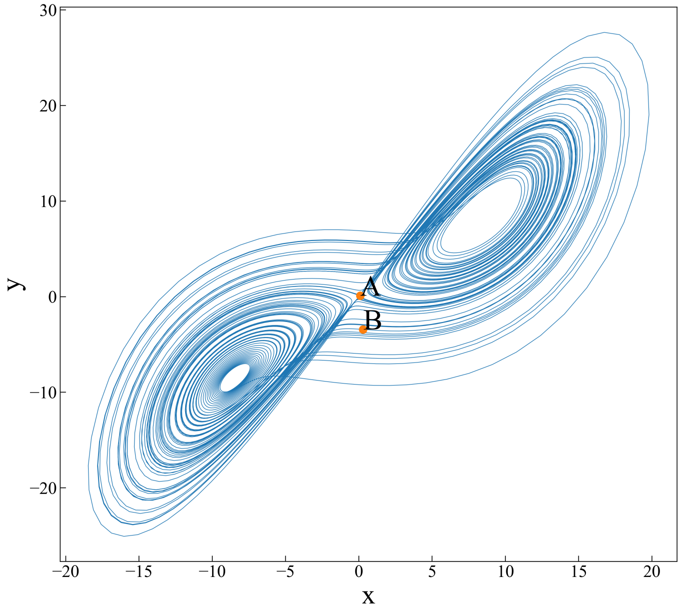


Figure 2: 2D attractor for Lorenz equation

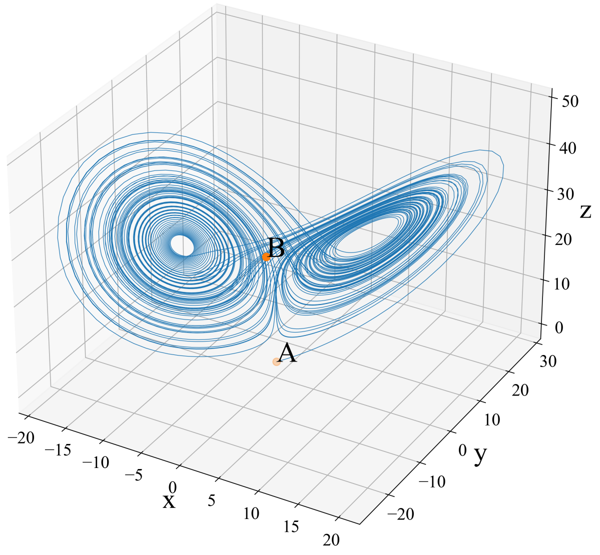


Figure 3: 3D attractor for Lorenz equation

2. FNN Methods

2.1. Conventional Methods

The authors proposed a method to estimate the minimum embedding dimension in 1992 [5]. Although this approach is simple in principle and works well for many nonlinear systems, it requires an appropriate threshold for every problem. The computational procedure of the FNN method is described in below.

Step 1: Reconstruct the attractor of the time series to be analyzed using Takens' embedding method [10]. Assuming that the target time series with N samples is x, the lag time is τ, and the embedding dimension is d, the d-dimensional attractor for x is reconstructed by $\{y_d(t)\}$ as follows.

$$\boldsymbol{y}_d(t) = \left(x(t), x(t+\tau), \cdots x(t+(d-1)\tau)\right). \tag{1}$$

Step 2: Calculate the time t' of the nearest neighbor vector of $\boldsymbol{y}_d(t)$, $\boldsymbol{y}_d^n(t)$, according to the distance criterion D.

$$D_d(t,t') = |\boldsymbol{y}_d(t) - \boldsymbol{y}_d(t')|\,,\ \ 0 \le t' \le N-(d-1)\tau\,,\ \ t \ne t', \tag{2}$$

$$t' = \underset{0\le i\le N-(d-1)\tau}{argmin}\ D_d(t,i)\,, \tag{3}$$

$$\boldsymbol{y}_d^n(t) = \boldsymbol{y}_d(t')\,. \tag{4}$$

However, the set of x values for which the function $f(x)$ on some set $\boldsymbol{A}$ is minimal is denoted as follows.

$$\underset{x\in\boldsymbol{A}}{argmin}\, f(x) = \{x \in \boldsymbol{A}\,|\, f(x) = \min_{y\in\boldsymbol{A}} f(y)\}\,. \tag{5}$$

Step 3: Determine the distance $D_d(t,t')$ between $\boldsymbol{y}_{d+1}(t)$ and $\boldsymbol{y}_{d+1}^n(t)$ when the embedding dimension is $d+1$.

$$\boldsymbol{D}_{d+1}(t,t') = |\boldsymbol{y}_{d+1}(t) - \boldsymbol{y}_{d+1}^n(t)|\,. \tag{6}$$

Step 4: Calculate the percentage of vectors in the attractor that are greater than or equal to the given threshold R_{tol} (false neighbor rate) relative to the rate of change in the neighbor point distance $D_{d+1}(t,t')/D_d(t,t')$ with increasing embedding dimensions.

$$\frac{\boldsymbol{D}_{d+1}(t,t')}{\boldsymbol{D}_d(t,t')} > R_{tol}\,. \tag{7}$$

Step 5: Repeat the operations in **Step 1** to **4** for the given maximum number of embedding dimensions. Then, count from the lowest dimension to the first dimension for which the false neighbor ratio is zero and set the latter as the minimum embedding dimension.

2.2. Proposed Method

According to this procedure, a neighborhood is considered false when the rate of increase exceeds the threshold value R_{tol} in **Step 4**. However, this value must be set for each time series to be analyzed. Therefore, in this study, we propose a false neighborhood reduction rate M_{tol} as a new criterion to replace the threshold R_{tol}. In **Step 4**, the rate of increase in the nearest neighbor distance with increasing dimensionality is calculated, and the percentage of vectors the rates of increase of which exceed the threshold value R_{tol} is calculated. In the proposed approach, the analysis is performed in the same manner up to the point at which the rate of increase of the nearest neighbor distance is obtained, except that the following calculation procedure is used for computable times: $\boldsymbol{T} = \{t | 0 \leq t \leq N - (d-1)\tau\}$.

Step 4': For the vector $\boldsymbol{y}_d^n(T) = \boldsymbol{y}_d(\boldsymbol{T'})$ that is the nearest neighbor to the vector $\boldsymbol{y}_d(\boldsymbol{T})$ of the attractor at time $\boldsymbol{T}$, find the rate of change of the nearest neighbor distance with increasing embedding dimensions. After applying the ordinary logarithm, the median value is M_{tol}, as given below.

$$M_{tol} = Med\left(\log_{10}\frac{\boldsymbol{D}_{d+1}(\boldsymbol{T},\boldsymbol{T'})}{\boldsymbol{D}_d(\boldsymbol{T},\boldsymbol{T'})}\right). \tag{8}$$

Step 5': With increasing embedding dimension d, the dimension in which M_{tol} first becomes less than 1 is considered the minimum embedding dimension, being that in which embedding without false neighbors is established.

Generally, time series may contain noise. Even for time series generated from systems of the same type, noise levels may vary considerably depending on the observation method. For example, consider the case in which the minimum embedding dimension is estimated using the FNN method for a time series with a higher noise level. Owing to the higher noise level, the trajectory of the reconstructed attractor becomes unstable. In this case, the variance of the rate of change of the distance to the nearest neighbor increases with increasing dimensionality. Therefore, the number of points that exceed the given threshold (false nearest neighbors) increases, and the estimated minimum embedding dimension is expected to be high compared to observed time series with less noise. However, because the proposed method refers to the median of the rate of change of the nearest neighbor distance with increasing dimensionality, the effect of increasing variance is expected to be small. Therefore, the proposed approach could serve as a more objective evaluation criterion by eliminating the arbitrariness of the threshold value.

3. Evaluation Method

To compare the FNN method with the proposed method, we used a time series of numerical solutions calculated from a mathematical model in which the variables of the system were known in advance. We considered the van der Pol equation, which depends on two variables, and the Lorenz equation, which depends on three, to investigate whether the estimated embedding dimension was affected by changing the nonlinearity of the time series by setting multiple decay coefficients for the van der Pol equation. In addition, Gaussian noise of 0%, 1%, 10%, and 100% of the standard deviation of the signal was added to the numerical solutions computed from both equations to comprehensively examine their robustness to noise. The minimum embedding dimension was estimated for the time series generated with each parameter setting using both the FNN method and the proposed method, and the difference from the expected minimum embedding dimension was evaluated using the error rate described below.

To demonstrate its applicability to general time series, we applied the proposed approach to time series data recorded from electrooculograms to evaluate how many factors or variables could be identified as underlying nystagmus movements in the system.

3.1. Van der Pol Equation

The van der Pol oscillator is among the first examples of deterministic chaos discovered [11]. Van der Pol found a stable oscillation in an electrical circuit using vacuum tubes, which he referred to as the relaxation oscillation. The van der Pol equation is a two-variable differential equation in *x* and *y* described by the following equations (9) and (10) with the damping coefficient μ.

$$\dot{x} = \mu\left(y + x - \frac{x^3}{3}\right), \tag{9}$$

$$\dot{y} = -\frac{x}{\mu}. \tag{10}$$

In this work, we evaluated whether the numerical solution of the van der Pol equation could be calculated as a second-order system with a minimum embedding dimension for the numerical solution of van der Pol equation. A total of 80 van der Pol time series were generated, including 20 with decay coefficients $\mu = 0.1$, 0.2, ... 2.1, and 4 with noise levels of 0%, 1%, 10%, and 100% of the standard deviation of the original time-series signal. For all generated time series, the initial values of *x* and *y* were set to 0.1, with a time-step width $dt = 0.01$, and a time-series length $l = 10000$ (Figure 4).

3.2. Lorenz Equation

The Lorenz equation [8] is a nonlinear ordinary differential equation of three variables *x*, *y*, and *z* that exhibits chaotic behavior, as given below. The equations depend on three variables, as shown in the following equations (11), (12), and (13). Lorenz presented this classical equation in 1963 while working on a model of atmospheric variability as a meteorologist at the Massachusetts Institute of Technology [8].

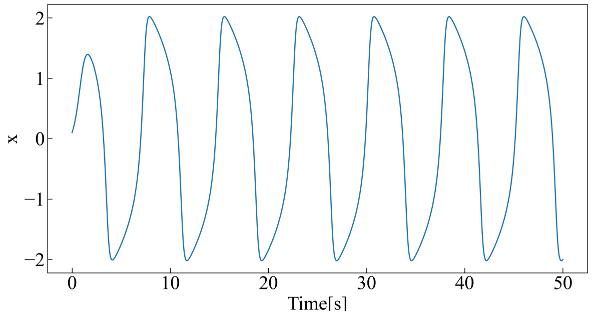


Figure 4: Typical example of van der Pol time series

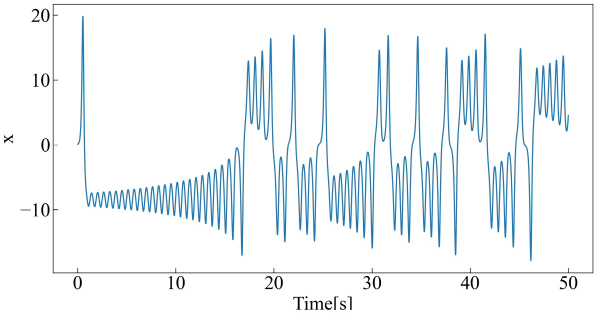


Figure 5: Typical example of Lorenz time series

$$\dot{x} = -px + py, \tag{11}$$

$$\dot{y} = -xz + rx - y, \tag{12}$$

$$\dot{z} = xy - bz, \tag{13}$$

where p, r, and b are constants that determine the behavior of the system.

Numerical solutions of the fourth-order Runge-Kutta method for the Lorenz equation were also obtained for reference, and the displacement x over time was plotted. For the parameter settings p=10, r=28, and b=8/3, the time interval was set to 0.01 second, and the initial values of x, y, and z were set to 0.1.

In this study, the numerical solution of the Lorenz equation was used to evaluate whether it could be calculated as a third-order system with a minimum embedding dimension. To evaluate the Lorenz time series, each parameter value was set to p=10, r=28, b=8/3, and four time series were generated with noise levels of 0%, 1%, 10%, and 100% of the standard deviation. For all generated time series, the initial values of x, y, and z were set to 0.1, the width of the time step dt = 0.01, and length of the time series l = 10000, respectively (Figure 5).

3.3. Electrooculogram data

An electrooculograms(EOG) is a time series recording changes in the distribution of electric potential around the eye caused by eye movements in two variables x and y. EOG are commonly used to examine the function of the retinal pigment epithelium and analyze eye movements. The electrical characteristics of the data involve a positive potential produced in the cornea closer to the environment and a negative potential produced closer to the retina. During measurement, single electrodes are affixed to a subject's inner and outer cornea, and their eye movements are recorded electrically via the potential difference between the electrodes [12].

Eye movements are divided into four types, including fixation, in which the eye concentrates on a single spot, pursuit, in which the eye slowly follows an object in the central fossa, saccadic and impulsive movements, in which the eye immediately detects anomalies, and vestibular movements, in which the eye responds to body movements when the body moves. The ocular muscles attached to the eyeballs move the eye and include the lateral rectus, medial rectus, superior rectus, inferior rectus, superior oblique, and inferior oblique muscles, as well as the upper eyelid elevator muscle, which is responsible for eye-opening action. Of the ocular muscles, the lateral rectus muscle is innervated by the abducens nerve, the superior oblique muscle by the pulmonic nerve, and all other muscles by the oculomotor nerve [13–15].

The EOG data used in this study were obtained from 11 healthy subjects (three men and eight women) aged 20-23 years. The subjects viewed experimental 3D images for 180 s each, and the angular velocity of their eye movement during this time was measured [16] (Figure 6).

3.4. Evaluation: definition of error value

We defined the error value as an index to evaluate the minimum embedding dimension obtained using the FNN algorithm and the proposed method. In this study, we used formula (14), where $\boldsymbol{D}$ is the minimum embedding dimension expected for the time series generated for each parameter setting, $\boldsymbol{D_{est}}$ is the minimum embedding dimension estimated from each method for those time series, and K is the total number of parameters when varying the attenuation coefficient and the amount of noise added using the error value E.

$$E = \frac{1}{K}\sum_{i=0}^{K-1}\left|\boldsymbol{D_i} - \boldsymbol{D_{est_i}}\right|. \tag{14}$$

4. Results

For Van der Pol Equation and Lorenz Equation, we report the estimated minimum embedding dimension and the error values of the expected minimum embedding dimension and the estimated values from the FNN method and the proposed method. However, for the Van der Pol Equation and Lorenz Equation, multiple time series were prepared under various conditions, so the estimated minimum embedding dimension is reported as an aggregate using the median value for each noise level. For the error between the expected minimum embedding dimension and the estimated value, we also report the error values aggregated by R_{tol} for the FNN method. The proposed method reports one error value per equation, since there is no threshold such as R_{tol}.

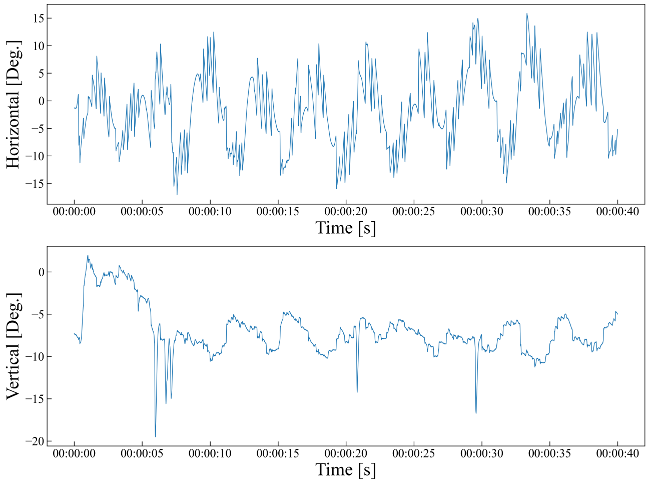


Figure 6: Typical example of an electrooculogram time series

Next, we report the results of estimating the minimum embedding dimension by each method for EOGs as an example of actual observed data.

4.1. Van der Pol Equation

First, we report on the estimated minimum embedding dimension for each method: the FNN method shows two dimensions when the noise level is 0%. The FNN method shows 2 dimensions when the noise level is 0%, 4 dimensions when the noise level is 1%, 5 dimensions when the noise level is 10%, and 6 dimensions when the noise level is 100%. On the other hand, the proposed method shows two dimensions when the noise level is 0% and 1%, and three dimensions when the noise level is 10% or higher.

Next, we report the error values for the estimated and expected minimum embedding dimension. When R_{tol} was set in the range of 1, 2, 4, and 1024 in the FNN method, the error value was minimized when R_{tol}=1024. In addition, the proposed method exhibited the same error value as when R_{tol}=1024 (Figure 7).

4.2. Lorenz Equation

First, we report on the estimated minimum embedding dimension for each method: the FNN method shows 6 dimensions for all noise levels from 0% to 100%, while the proposed method shows 2 dimensions for noise levels of 0% and 1%, and 3 dimensions for noise levels above 10%. On the other hand, the proposed method shows two dimensions for noise levels of 0% and 1%, and three dimensions for noise levels above 10%.

Next, we report the error values for the estimated and expected minimum embedding dimension. When R_{tol} was set in the range of 1, 2, 4, and 1024 in the FNN method, the error value was the smallest when R_{tol}=512,1024. In the proposed method, the error value was smaller than R_{tol}=256 and larger than R_{tol}=512, 1024 (Figure 8).

4.3. Electrooculogram data

For the variation in the horizontal axis, the minimum embedding dimensions were estimated using the FNN and proposed methods. With the FNN method, R_{tol} was estimated to be 1, 2, 4..., 1024. Assuming that the error value of the estimated minimum embedding dimension was smallest when R_{tol}=1024 in 4.1 and that 4.2 and can be applied to EOG data, the horizontal variation of EOG data by the FNN method was estimated and was considered as a system consisting of approximately 4 to 6 variables. In the vertical direction, the same can be considered as a system consisting of 8 to 11 variables. In contrast, using the proposed method, the minimum embedding dimension was estimated as two to three dimensions in the horizontal direction and two dimensions in the vertical direction, suggesting that the system is composed of approximately two to three variables (Figure 9-10).

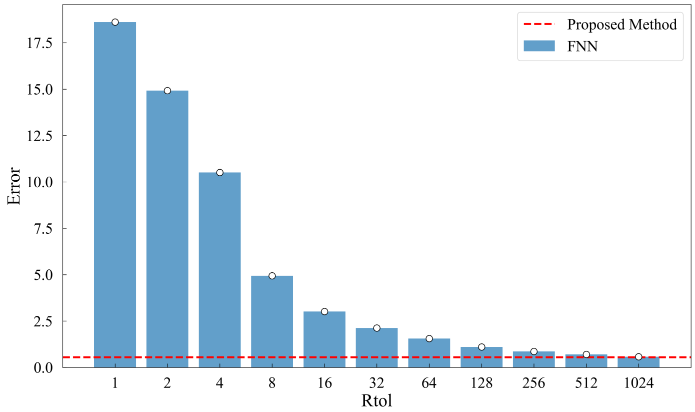


Figure 7: Error rate in estimating the minimum embedding dimension for van der Pol time series using the FNN method and the proposed method

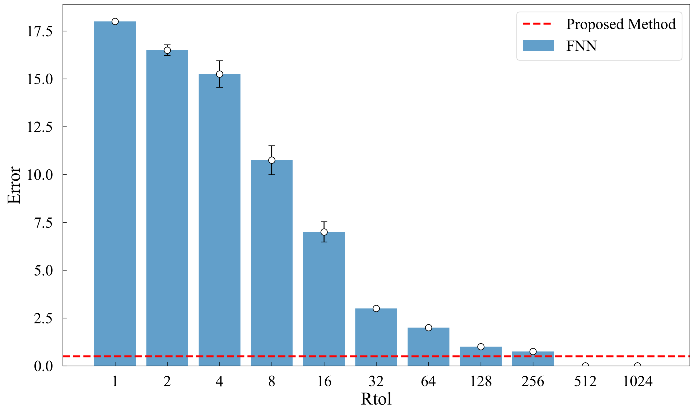


Figure 8: Error rate in estimating the minimum embedding dimension for Lorenz time series using the FNN method and the proposed method

5. Discussion

To compare the FNN method with the proposed approach, we considered a time series of numerical solutions calculated from a mathematical model in which the number of variables constituting the system is known in advance. In this study, the van der Pol equation consisting of two variables and the Lorenz equation consisting of three variables were used, and the difference between the expected minimum embedding dimension and the minimum embedding dimension calculated using each method was evaluated as an error value. First, we discuss the estimation results of the minimum embedding dimension by the proposed method and the FNN method. Next, we discuss the error values between the minimum embedding dimension estimated by each method and the actual minimum embedding dimension. Finally, we discuss the results of estimating the minimum embedding dimension by each method for EOGs as an example of actual observed data.

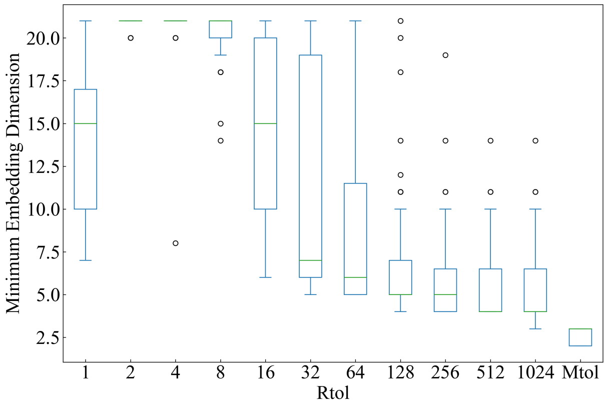


Figure 9: Estimation of the minimum embedding dimension for the horizontal-axis electrooculogram data

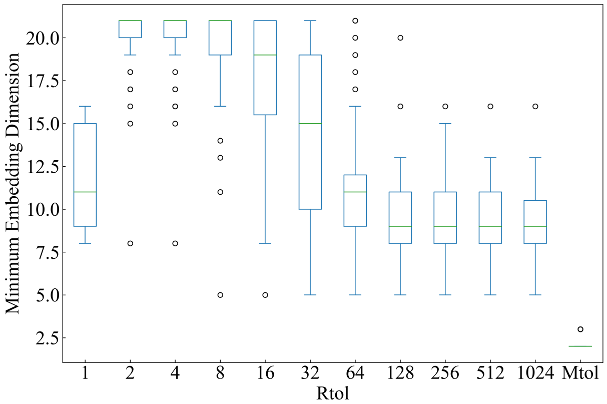


Figure 10: Estimation of the minimum embedding dimension for the vertical-axis electrooculogram data

5.1. Van der Pol Equation

The estimates of the minimum embedding dimension are discussed: both the FNN method and the proposed method show a minimum embedding dimension of 2 dimensions when the noise level is 0%. This is also the case for the FNN method in [17]. However, the FNN method suggested that the estimate may be vulnerable to noise, since an increase in the noise level to 1% resulted in a three-dimensional estimate. On the other hand, the proposed method did not affect the estimates until the noise level exceeded 10%, suggesting that it may be more robust against noise when compared to the FNN method.

Next, we discuss the error between the estimated first embedding dimension and the expected minimum embedding dimension for each method: the FNN method showed the smallest error when R_{tol}=1024. At this time, the proposed method also showed a comparable error. In other words, the proposed method does not require any threshold adjustment and shows the same error as the optimal value in the FNN method. Of course, since the error tends to decrease as R_{tol} increases, it is conceivable that the error would decrease more when the threshold is further increased. However, since the decrease in error stalls when R_{tol} exceeds 256, the decrease in threshold value is considered to be limited even if the threshold value is increased. In addition, it is difficult to determine the optimal value of R_{tol} because there are situations where the nature and data noise level are unknown in the actual time series to be analyzed. Even if the threshold is set in advance with a large value, it is believed that there may be cases where the minimum embedding dimension is underestimated.

5.2. Lorenz Equation

The estimates of the minimum embedding dimension are discussed: the FNN method estimated the minimum embedding dimension to be 6 under all noise levels of 0%, 1%, 10%, and 100%. It can be seen that even when the noise level is 0%, the results are very different from what one would expect. To reiterate, this estimate uses the median of the output values estimated for various R_{tol} settings, so it may output 3, the expected minimum embedding dimension, when the appropriate R_{tol} is set. Conversely, if an appropriate R_{tol} cannot be set, it is difficult to estimate the appropriate minimum embedding dimension. On the other hand, the proposed method is considered to be tolerant to noise because it could output the expected minimum embedding dimension even when relatively large noise levels of 10% and 100% were introduced. On the other hand, when the noise levels were relatively low (0% and 1%), the proposed method estimated one dimension less than the expected minimum embedding dimension. This may indicate that the minimum embedding dimension may be underestimated for time series with little disturbance.

Next, we discuss the error between the estimated embedding dimension and the expected minimum embedding dimension for each method: the FNN method showed the smallest error when R_{tol} = 512 and 1024. In this case, when R_{tol} is set to a value of 512 or 1024, the optimal minimum embedding dimension can be obtained even after considering the effects of various noises. On the other hand, if R_{tol} is set to a value higher than 1024, there is a possibility that the minimum embedding dimension will be underestimated. On the other hand, the proposed method had a smaller error value than the FNN method when R_{tol} was less than 512, but the FNN method had a smaller error value when R_{tol} was greater than 512. From these results, it can be said that if the optimal R_{tol} can be set, the FNN method has a smaller error value, i.e., a more accurate estimation of the minimum embedding dimension, but in most cases, the proposed method has a more accurate estimation value.

5.3. Electrooculogram data

To demonstrate an application to general time series, we applied both the FNN method and the proposed method to EOG data to determine the minimum embedding dimension and examined how many variables the eye movements consist of in the system. The results show that the FNN method found 4 to 6 dimensions in the horizontal direction and 8 to 11 dimensions in the vertical direction. In contrast, the proposed method suggested 2 to 3 dimensions in the direction of the horizontal axis and 2 dimensions in the vertical direction. Several mathematical models of eye movements have been devised, and the Westheimer model s[18–20] is a representative example. In this model, human eye movements are described by a third-order differential equation, which is also close to the minimum embedding dimension calculated by the method proposed in this work.

6. Conclusion

In this study, we have proposed a method to solve the arbitrariness of the threshold used in the FNN method for time series. We have also applied the proposed method to EOG data as an example, and compared the performance of the FNN method and the proposed method in estimating the number of factors that determine eye movements. Compared to the FNN method, the minimum embedding dimension calculated by the proposed method exhibited a lower error value than expected under various conditions. The proposed approach mitigates the difficulty of setting appropriate parameters for a time series of unknown nature and obviates the need for an arbitrary threshold value. In future research, we intend to verify the effectiveness of the proposed method by applying it to a wider range of general time series.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, "Generative adversarial nets," in Advances in Neural Information Processing Systems, 2014, doi:10.3156/jsoft.29.5_177_2.

[2] A. Radford, L. Metz, S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," in 4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings, 2016.

[3] F. Kinoshita, Y. Mori, Y. Matsuura, H. Takada, M. Miyao, "A study of numerical analysis and mathematical modeling of electrogastrograms in the young subjects," IEEJ Transactions on Electronics, Information and Systems, **136**(9), 2016, doi:10.1541/ieejeiss.136.1261.

[4] Y. Jono, T. Tanimura, F. Kinoshita, H. Takada, "Evaluation of Numerical Solution of Stochastic Differential Equations Describing Body Sway Using Translation Error," Forma, **35**(1), 2020, doi:10.5047/forma.2020.006.

[5] M.B. Kennel, R. Brown, H.D.I. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," Physical Review A, **45**(6), 1992, doi:10.1103/PhysRevA.45.3403.

[6] R. Hegger, H. Kantz, "Improved false nearest neighbor method to detect determinism in time series data," Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics, **60**(4 B), 1999, doi:10.1103/physreve.60.4970.

[7] S. Wallot, D. Mønster, "Calculation of Average Mutual Information (AMI) and false-nearest neighbors (FNN) for the estimation of embedding parameters of multidimensional time series in matlab," Frontiers in Psychology, **9**(SEP), 2018, doi:10.3389/fpsyg.2018.01679.

[8] E.N. Lorenz, "Deterministic Nonperiodic Flow," Journal of the Atmospheric Sciences, **20**(2), 1963, doi:10.1175/1520-0469(1963)020<0130:dnf>2.0.co;2.

[9] J. Milnor, "On the concept of attractor," Communications in Mathematical Physics, **99**(2), 1985, doi:10.1007/BF01212280.

[10] F. Takens, Detecting strange attractors in turbulence, 1981, doi:10.1007/bfb0091924.

[11] Balth. van der Pol, " LXXXVIII. On 'relaxation-oscillations' ," The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, **2**(11), 1926, doi:10.1080/14786442608564127.

[12] W.C. Lara, B.L. Jordan, G.M. Hope, W.W. Dawson, R.A. Foster, S. Kaushal, "Fast Oscillations of the Electro-oculogram in Cystic Fibrosis," Investigative Ophthalmology & Visual Science, **44**(13), 4957, 2003.

[13] Z.M. Hafed, J.J. Clark, "Microsaccades as an overt measure of covert attention shifts," Vision Research, **42**(22), 2002, doi:10.1016/S0042-6989(02)00263-8.

[14] H. Kobayashi, Y. Watanabe, N. Ohashi, K. Mizukoshi, "Pendular Optokinetic Nystagmus Test in Patients with Central Nervous System Disorders," Practica Otologica, Supplement, **1989**, 1989, doi:10.5631/jibirinsuppl1986.1989.Supplement36_133.

[15] R. Engbert, R. Kliegl, "Microsaccades uncover the orientation of covert attention," Vision Research, **43**(9), 2003, doi:10.1016/S0042-6989(03)00084-1.

[16] A. Sugiura, R. Ono, Y. Itazu, H. Sakakura, H. Takada, "Analysis of Characteristics of Eye Movement While Viewing Movies and Its Application," Nihon Eiseigaku Zasshi. Japanese Journal of Hygiene, **77**, 2022, doi:10.1265/jjh.21004.

[17] E. Conte, A. Federici, G. Pierri, L. Mendolicchio, J.P. Zbilut, A brief note on recurrence quantification analysis of bipolar disorder performed by using a van der Pol oscillator model, 2009.

[18] G. Westheimer, "Mechanism of saccadic eye movements," A.M.A. Archives of Ophthalmology, **52**(5), 1954, doi:10.1001/archopht.1954.00920050716006.

[19] K. Shimono, M. Kondo, K. Shibuta, S. Nakamizo, "Psychophysical and vergence responses of normal and stereoanomalous observers to pulse-disparity stimulus," The Japanese Journal of Psychology, **53**(3), 1982, doi:10.4992/jjpsy.53.136.

[20] P.D.S.H. Gunawardane, R.R. Macneil, L. Zhao, J.T. Enns, C.W. de Silva, M. Chiao, "A fusion algorithm for saccade eye movement enhancement with eog and lumped-element models," IEEE Transactions on Biomedical Engineering, **68**(10), 2021, doi:10.1109/TBME.2021.3062256.

Advances in Science, Technology and Engineering Systems Journal
Vol. 7, No. 4, 121-126 (2022)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

Regularity of Radon Transform on a Convex Shape

Pat Vatiwutipong[*]

Department of Mathematics and Computer Science, Kamnoetvidya Science Academy, Rayong, 21210, Thailand

ARTICLE INFO

Article history:
Received: 18 June, 2022
Accepted: 16 August, 2022
Online: 24 August, 2022

Keywords:
Radon transform
Regularity property
Convex shape

ABSTRACT

Radon transform is a mathematical tool widely applied in various domains, including biophysics and computer tomography. Previously, it was discovered that applying the Radon transform to a binary image comprising circle forms resulted in discontinuity. As a result, the line detection approach based on it became discontinued. The d-Radon transform is a modified version of the Radon transform that is presented as a solution to this problem. The properties of the circle cause the Radon transform to be discontinuous. This work extends this finding by looking into the Radon transform's regularity property and a proposed modification to a convex shape. We discovered that regularity in the Radon space is determined by the regularity of the shape's point. This leads to the continuity condition for the line detection method.

1 Introduction

This paper is an extension of work originally presented in the 2022 14th International Conference on Knowledge and Smart Technology (KST) [1].

Radon transform is a well-known integral transform used in many fields that was originally invented in 1917 by Johann Radon, [2]. The Radon is used to create an image from the projection data associated with cross-sectional scans, [3]. Many other integral transforms, such as the Hough transform, Penrose transform, and Fourier transform, are significantly associated with the Radon transform.

Intuitively, Radon transformation is to integrate a given object and project the value onto the lower-dimensional space in various directions. In computer tomography, for example, the Radon transform was applied to a 3D object to obtain a 2D image. In practice, we acquire the 2D projected picture and must use the inversion of the Radon transform to reconstruct the 3D object, [4].

Hough transform is one of the line detection techniques used in image processing. By applying the Hough transform to an image, the image's point accumulation in each direction is computed. In the Hough space, a direction that passes through more points will have a greater value if the image comprises multiple points. The line direction in the image can be used to detect the maximal argument on the Hough space, [5].

In 2016, R. Aramini et al. found a unifying perspective of the Radon and Hough transformations [6]. The Radon transform is considered a continuous version of the Hough transform, so it was also applicable as a line detection method by changing the accumulation function to the integration of images in each direction instead. Since the Hough transform is a discrete transformation, the line detection method's continuity cannot be expected. However, it is different for the Radon transform.

In [1], we explained the benefits of the continuity of line detection method. It has been shown that the continuity of the line detection method using Radon transform depends on the regularity of the Radon space, but that work focused only on the image containing points. In practice, an actual image contains not only points but also shapes. For a grey-scale image as a function from a 2D image space to the interval $[0, 1]$, the value 0 and 1 represent white and black colors, respectively. If the image is continuous, we have that the Radon transform is also continuous. Nevertheless, image discontinuity is expected for a binary image whose range is just $\{0, 1\}$. This paper aims to study the regularity properties of the Radon transform of a binary image containing shapes, which lead to the continuity of the line detection method.

2 Definition and Elementary Properties

This section presents the definition of *Radon transform* and its basic properties.

Definition 2.1. (Radon transform) For a given function $m : W \subset \mathbb{R}^2 \to \mathbb{R}$, the Radon transform of m is defined, for each pair of real numbers θ and r,

$$\mathcal{R}[m](\theta, r) = \int_{r=x\cos\theta+y\sin\theta} m(x, y)ds, \tag{1}$$

[*] Corresponding Author: Pat Vatiwutipong, Email: pat.v@kvis.ac.th

https://dx.doi.org/10.25046/aj070416

whenever the integration is defined.

Lemma 2.2. *We have $(x, y) \in \ell_{\theta,r}$ if and only if there exists $s \in \mathbb{R}$ satisfying $x = r\cos\theta - s\sin\theta$ and $y = r\sin\theta + s\cos\theta$.*

According to this lemma, we can rewrite the definition of Radon transform as

$$\mathcal{R}[m](\theta, r) = \int_{-\infty}^{\infty} m(r\cos\theta - s\sin\theta, r\sin\theta + s\cos\theta)ds. \tag{2}$$

It was proven in [7] that the Radon transform $\mathcal{R}[m]$ exists for all $m \in L^1(\mathbb{R}^2)$.

For a function $h : E \to \mathbb{R}$, define a *norm* as in [7] by,

$$\|h\|_{\infty,1} = \max_{\theta\in[-\frac{\pi}{2},\frac{\pi}{2})} \int_{-\infty}^{\infty} |h(\theta, r)|dr. \tag{3}$$

It was also proven in [7] that, for $m \in L^1(\mathbb{R}^2)$,

$$\|\mathcal{R}[m]\|_{\infty,1} \leq \int_{\mathbb{R}^2} |m(x, y)|dxdy. \tag{4}$$

Since the Radon transform is linear, we have

$$\|\mathcal{R}[m] - \mathcal{R}[\tilde{m}]\|_{\infty,1} \leq \int_{\mathbb{R}^2} |m(x, y) - \tilde{m}(x, y)|dxdy, \tag{5}$$

which means that if there is a small change on the function m, the change on its Radon transform will also be small.

3 Regularity of Radon Transform

3.1 Regularity on Circle

First, let B_1 be a single circle in image space W centered at $P_1 = (x_1, y_1)$ with radius $R_1 > 0$. The Radon transform of this single circle is equal to the length of the intersection between the circle and $\ell_{\theta,r}$, so if the circle is centered at the origin, we have that

$$\mathcal{R}[B_1](\theta, r) = \begin{cases} 2\sqrt{R_1^2 - r^2} & \text{if } |r| \leq R_1 \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

Linearity and shifting property of the Radon transform are proven in [8]. From the shifting property,

$$\mathcal{R}[B_1](\theta, r) = \begin{cases} 2\sqrt{R_1^2 - (r - x_1\cos\theta - y_1\sin\theta)^2} \\ \qquad\qquad \text{if } |r - x_1\cos\theta - y_1\sin\theta| \leq R_1 \\ 0 \qquad\quad \text{otherwise.} \end{cases} \tag{7}$$

Recall that we can represent $\mathcal{R}[P_1](\theta, r)$ by a sinusoidal curve $r_1(\theta) = x_1\cos\theta + y_1\sin\theta$, so we have that

$$\mathcal{R}[B_1](\theta, r) = \begin{cases} 2\sqrt{R_1^2 - (r - r_1(\theta))^2} & \text{if } |r - r_1(\theta)| \leq R_1 \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

In the Radon space E, the graph of $\mathcal{R}[B_1](\theta, r)$ can be seen as a "tunnel" centered at r_1 (see Figure 1).

When θ is fixed we can consider $\mathcal{R}[B_1](\theta, r)$ as a function of r which forms an upper-half of an ellipse centered at $r_1(\theta)$, with base length and height equal to $2R_1$ (see figure 2).

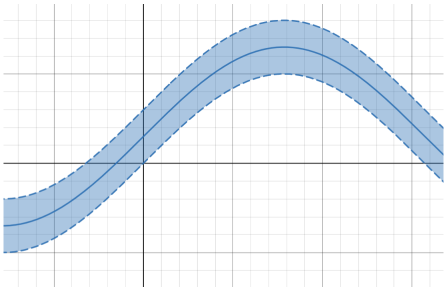


Figure 1: Graph of $\mathcal{R}[B_1](\theta, r)$ in E

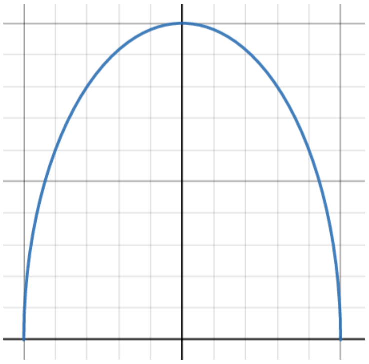


Figure 2: Cross section of $\mathcal{R}[B_1](\theta, r)$ for a fixed θ.

In this setting, the Radon transform $\mathcal{R}[B_1](\theta, r)$ is a continuous function on (θ, r) but it is not differentiable in r at $r = r_1(\theta) \pm R$.

To obtain more regularity, we modify the Radon transform by changing the integration domain from a line to a strip centered at that line instead.

Definition 3.1. (d-Radon transform) For $d > 0$ and $m \in L^1(\mathbb{R}^2)$, define the d−Radon transform of m, for each pair of real numbers θ and r, by

$$\mathcal{R}_d[m](\theta, r) = \int_{r-d}^{r+d} \mathcal{R}[m](\theta, \rho)d\rho. \tag{9}$$

We also have linearity and shifting properties for our new transformation. By the fundamental theorem of calculus and the fact that $\mathcal{R}[B_1](\theta, r)$ is continuous, the d-Radon transform $\mathcal{R}_d[B_1](\theta, r)$ is differentiable in r. Next, we will derive $\mathcal{R}_d[B_1](\theta, r)$ explicitly for each θ. For the sake of convenience, we shift the Radon space such that $r_1(\theta) = 0$. We have two cases.

Case I: If $0 < d \leq R_1$, we have that

$$\mathcal{R}_d[B_1](\theta, r) = \begin{cases} \int_{r-d}^{R_1} \mathcal{R}[B_1](\theta, \rho)d\rho & \text{if } R_1 \leq r \leq R_1 + d \\ \int_{r-d}^{r+d} \mathcal{R}[B_1](\theta, \rho)d\rho & \text{if } -R_1 \leq r \leq R_1 \\ \int_{-R_1}^{r+d} \mathcal{R}[B_1](\theta, \rho)d\rho & \text{if } -R_1 - d \leq r \leq -R_1. \end{cases}$$

That is $\mathcal{R}_d[B_1](\theta, r)$

$$= \begin{cases} \frac{\pi R_1^2}{2} - [(r-d)\sqrt{R_1^2-(r-d)^2} + R_1^2 \arctan(\frac{r-d}{\sqrt{R_1^2-(r-d)^2}})] \\ \qquad \text{if } R_1 \le r \le R_1 + d \\ [(r+d)\sqrt{R_1^2-(r+d)^2} + R_1^2 \arctan(\frac{r+d}{\sqrt{R_1^2-(r+d)^2}})] \\ \quad -[(r-d)\sqrt{R_1^2-(r-d)^2} + R_1^2 \arctan(\frac{r-d}{\sqrt{R_1^2-(r-d)^2}})] \\ \qquad \text{if } -R_1 \le r \le R_1 \\ \frac{\pi R_1^2}{2} + [(r+d)\sqrt{R_1^2-(r+d)^2} + R_1^2 \arctan(\frac{r+d}{\sqrt{R_1^2-(r+d)^2}})] \\ \qquad \text{if } -R_1 - d \le r \le -R_1. \end{cases} \tag{10}$$

The graph is shown in Figure 3.

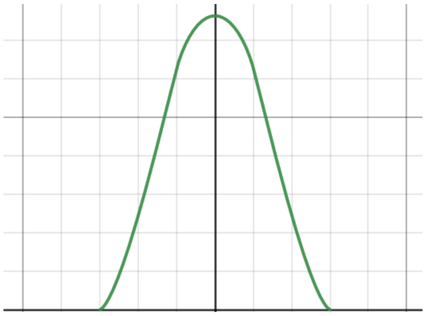


Figure 3: Cross section of $\mathcal{R}[B_1]_d(\theta, r)$ when $0 < d \le R_1$ for a fixed θ.

Case II: If $R_1 < d$, we have that

$$\mathcal{R}_d[B_1](\theta, r) = \begin{cases} \int_{r-d}^{R_1} \mathcal{R}[B_1](\theta, \rho)d\rho & \text{if } R_1 \le r \le R_1 + d \\ \int_{-R_1}^{R_1} \mathcal{R}[B_1](\theta, \rho)d\rho & \text{if } -R_1 \le r \le R_1 \\ \int_{-R_1}^{r+d} \mathcal{R}[B_1](\theta, \rho)d\rho & \text{if } -R_1 - d \le r \le -R_1. \end{cases}$$

That is $\mathcal{R}_d[B_1](\theta, r)$

$$= \begin{cases} \frac{\pi R_1^2}{2} - [(r-d)\sqrt{R_1^2-(r-d)^2} + R_1^2 \arctan(\frac{r-d}{\sqrt{R_1^2-(r-d)^2}})] \\ \qquad \text{if } R_1 \le r \le R_1 + d \\ \pi R_1^2 \qquad \text{if } -R_1 \le r \le R_1 \\ \frac{\pi R_1^2}{2} + [(r+d)\sqrt{R_1^2-(r+d)^2} + R_1^2 \arctan(\frac{r+d}{\sqrt{R_1^2-(r+d)^2}})] \\ \qquad \text{if } -R_1 - d \le r \le -R_1. \end{cases} \tag{11}$$

The graph is shown in Figure 4.

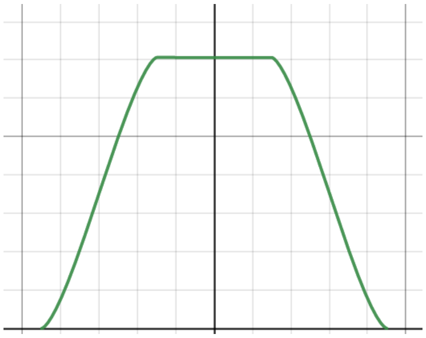


Figure 4: Cross section of $\mathcal{R}[B_1]_d(\theta, r)$ when $R_1 < d$ for a fixed θ.

For a fixed θ, $\mathcal{R}[B_1](\theta, r)$ reaches its maximum at $r = r_1(\theta)$. If $d \le R_1$, the maximum is $2d\sqrt{R_1^2 - d^2} + 2R_1^2 \arctan(\frac{d}{\sqrt{R_1^2-d^2}})$ and if $d > R_1$, the maximum is πR_1^2.

Lemma 3.2. *Let B_1 and B_2 be circles with same radius R, and fix $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2})$, we have that*

$$\max_{r \in \mathbb{R}} |\mathcal{R}[B_1](\theta, r) - \mathcal{R}[B_2](\theta, r)| = \min\{2R, 2\sqrt{2Ra_\theta - a_\theta^2}\} \tag{12}$$

and

$$\max_{r \in \mathbb{R}} |\mathcal{R}_d[B_1](\theta, r) - \mathcal{R}_d[B_2](\theta, r)| \le 2\sqrt{2}Ra_\theta \tag{13}$$

where $a_\theta = |r_1(\theta) - r_2(\theta)|$.

Proof. The proof of this lemma is straightforward. □

Now, we define another norm for the Radon space.
For a function $h : E \to \mathbb{R}$ define

$$||h||_\infty = \max_{(\theta,r) \in [-\pi,\pi) \times \mathbb{R}} |h(\theta, r)|. \tag{14}$$

Let B_1 be a circle and B_2 be another circle obtained by shifting B_1 with distance δ. We have that the value of $|\mathcal{R}[B_1](\theta, r) - \mathcal{R}[B_2](\theta, r)|$ will reach the maximum when $a_\theta = ||P_1 - P_2|| = \delta$. So, from Lemma 3.2, we get that $||\mathcal{R}[B_1] - \mathcal{R}[B_2]||_\infty = \min\{2R, 2\sqrt{2R\delta - \delta^2}\}$ which cannot be bounded by $C\delta$ for any constant C. Indeed, its derivative when δ is close to zero tends to infinity. So, in this case, a small change in the image will create a large change in its Radon transform.

For the d-Radon transform, the situation is totally different. Since we have a linear estimation in the previous lemma, we get that $||\mathcal{R}_d[B_1] - \mathcal{R}_d[B_2]||_\infty \le 2\sqrt{2}R\delta$. By linearity of d-Radon transform, we can extend this result to the case of images containing more than one circle. This leads to the next proposition.

Definition 3.3. Let $\mathfrak{B}$ be a set of all images containing a finite number of circles with the same radius $R > 0$. Define a distance function $\bar{D}$ on $\mathfrak{B}$ by $\bar{D}(m_1, m_2) = \bar{d}(m_1^P, m_2^P)$ where $\bar{d}$ is defined in Equation (4) and m^P is an image containing only the center of each circle in m.

Proposition 3.4. *Let $m_1, m_2 \in \mathfrak{B}$ be two images which contain N_1 and N_2 circles respectively, and $N_M = \max\{N_1, N_2\}$. We have that*

$$||\mathcal{R}_d[m_1] - \mathcal{R}_d[m_2]||_\infty \le 2\sqrt{2}RN_M\bar{D}(m_1, m_2). \tag{15}$$

3.2 Regularity on Convex Shape

Let m be an image that contains a single convex shape, that is, $m(x, y) = 1_A$, where A is a connected, compact and convex subset of W. Let curve C be the boundary of A. Suppose that C is continuous and piecewise differentiable.

Definition 3.5. Let P be a point on C. Let $\partial C(P)$ be the set of $\phi \in [-\pi, \pi)$ such that there exists r_ϕ which makes line $\ell_{\phi,r_\phi} : r_\phi = x\cos\phi + y\sin\phi$ intersects the set C only at point P or tangent of C in P.

Note that $\partial C(P)$ always exists for any P since C is convex, and it is a singleton when P is a differentiable point of C.

Remark 3.6. Since C is a closed curve, we get that the set $\{\phi : \phi \in \partial C(P), P \in C\}$ is equal to $[-\pi, \pi)$.

For all $\phi \in \partial C(P)$, let $\ell^{\perp}_{\phi,r_\phi}$ be the perpendicular line to ℓ_{ϕ,r_ϕ} passing through P. Denote the length of the segment of $\ell^{\perp}_{\phi,r_\phi}$ that intersected with A by D_ϕ. Parameterize $\ell^{\perp}_{\phi,r_\phi}$ by s. Set s to be zero at P and $s > 0$ in the direction that passes through A.

Definition 3.7. For a fixed P, define a function $f_\phi : \mathbb{R} \to \mathbb{R}$, for each ϕ, by letting $f_\phi(s)$ be the length of $\ell_{\phi,r_\phi+s}$ across A. Equivalently, $f_\phi(s) = \mathcal{R}[m](\phi, r_\phi + s)$. Note that $f_\phi(s) = 0$ for all $s \notin [0, D_\phi]$. When P is a differentiable point of C, since $\partial C(P)$ is a singleton, we may denote ℓ_{ϕ,r_ϕ} by ℓ_{θ_P,r_P} and f_ϕ by f_P.

All notations in definition 3.5 and 3.7 are shown in Figure 5.

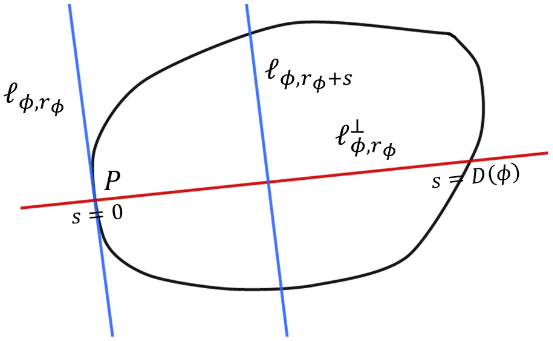


Figure 5: Point P on C and its ℓ_{ϕ,r_ϕ}, $\ell^{\perp}_{\phi,r_\phi}$ and D_ϕ

Definition 3.8. We say that point P is a *locally linear* point on C if there is a neighborhood of P where C is a line in it.

Lemma 3.9. *For all $s \in (0, D_\phi)$, the line $\ell_{\phi,r_\phi+s}$ is intersects with C in exactly two points.*

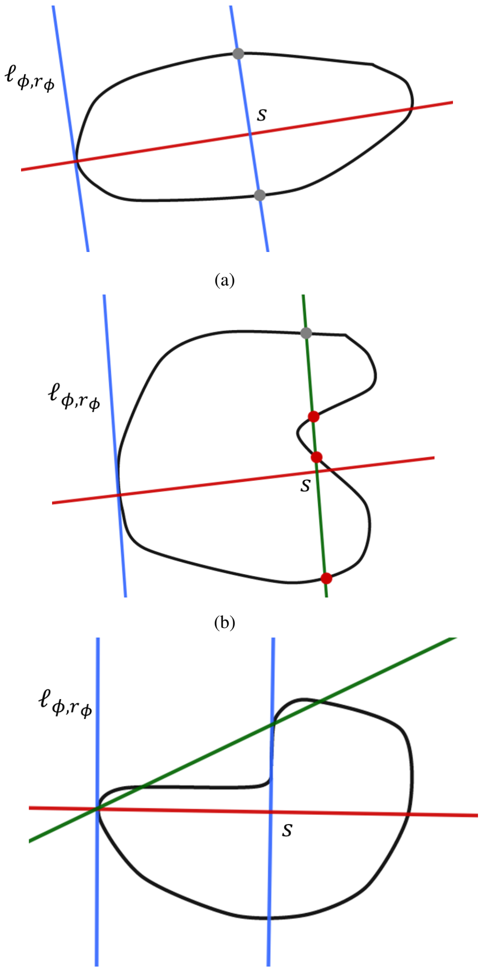
Proof. Suppose that there is $s \in (0, D_\phi)$ such that the line $\ell_{\phi,r_\phi+s}$ is intersects with C in more than two points. If one point is isolated, we can draw a line connecting that point to the others as in Figure 6 (b). It belongs to A, which contradicts the convexity of A. If there is no isolated point, there is a part of C that intersects with $\ell_{\phi,r_\phi+s}$ which is locally linear as in Figure 6 (c). Since C is the boundary of the shape, one side of it should not be in A. Without loss of generality, we can suppose that it is the left side. We pick the point on that part of C which is farthest from $\ell_{\phi,r_\phi+s}$ so that the line connecting that point and P will not be contained in A. Again, this contradicts the convexity of A. So the only case that can happen is that it crosses only two points as shown in Figure 6 (a). □

Proposition 3.10. *The function f_ϕ is continuous and piecewise differentiable on $(0, D_\phi)$.*

Proof. By previous lemma, we have that the value of $f_\phi(s)$ on $(0, D_\phi)$ is equal to the distance between the two intersection points of $\ell_{\phi,r_\phi+s}$ and C. Since C is continuous on $(0, D_\phi)$, so is f_ϕ. We can also get piecewise differentiability of f_ϕ by a similar argument. □

Proposition 3.11. *The curve C is locally linear at P if and only if f_P is not continuous at 0.*

Proof. First of all, note that $\lim_{s\to 0^-} f_P(s) = 0$. Suppose that C is locally linear at P, then C is differentiable at P. So, $f_P(0)$ is equal to the length of C that intersect with ℓ_{θ_P,r_P}, which is greater than zero. That is f_P is not continuous at 0. On the contrary, suppose that C is not locally linear at P. So, C must intersect ℓ_{θ_P,r_P} only at point P, that is $f_P(0) = 0$. Since C is continuous, we get $\lim_{s\to 0^+} f_P(s) = 0$. This proves continuity of f_P at $s = 0$. □



(a)

(b)

(c)

Figure 6: Non-differentiable point P on C and its ϕ^+, ϕ^-

Proposition 3.12. *If the curve C is differentiable at point P, then f_P is not differentiable at 0. Conversely, if f_ϕ is not differentiable at 0 for some $\phi \in \partial C(P)$, then C is differentiable at point P.*

Proof. Suppose that C is differentiable at P. Since ℓ_{θ_P,r_P} is tangent to C at P, we have $f_P'(0) = \infty$. On the contrary, suppose that C is not differentiable at P. There will be two distinct value $\phi^+ = \max\{\phi \in \partial C(P)\}$ and $\phi^- = \min\{\phi \in \partial C(P)\}$ as in Figure 7. Let $\phi \in \partial C(P)$. We get that $f_\phi'(s)$ will be bounded above and below by the slope of $\ell_{\phi^+,r_{\phi^+}}$ and $\ell_{\phi^-,r_{\phi^-}}$ evaluated with respect to ℓ_{ϕ,r_ϕ}. □

Hence, by Remark 3.6, we can separate $\theta \in [-\pi, \pi)$ into three types according to the property of its corresponding point P as: a locally linear, a differentiable but not locally linear and a non-differentiable points. The Radon transform is discontinuous for the locally linear point, and the d-Radon transform is continuous but not differentiable. For the differentiable point that is not locally linear, the Radon transform is continuous but not differentiable, and the d-Radon transform is differentiable. Lastly, at the non-differentiable point, both Radon and d-Radon transform are differentiable.

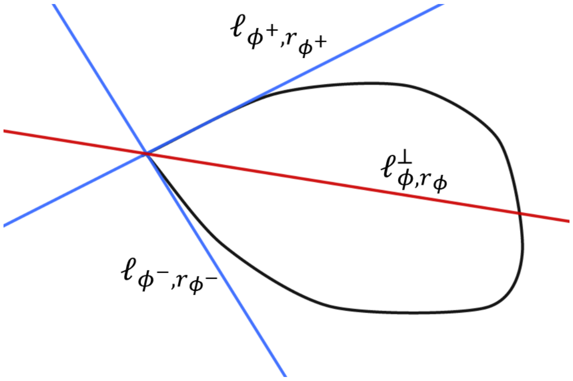

$\ell_{\phi^+,r_{\phi^+}}$
$\ell^\perp_{\phi,r_\phi}$
$\ell_{\phi^-,r_{\phi^-}}$

Figure 7: Non-differentiable point P on C and its ϕ^+, ϕ^-

To illustrated the result, take the following numerical examples.

Example 3.13. Consider a square

$$S = \{(x, y) : |x - y| + |x + y| \leq 2\}.$$

Only the corner points $(-1, -1), (-1, 1), (1, -1)$ and $(1, 1)$ are non-differentiable points of S, and the others are locally linear. So, the Radon transform of S is not continuous only in the horizontal and vertical directions, see Figure 8. For d-Radon transform, it is continuous in every direction.

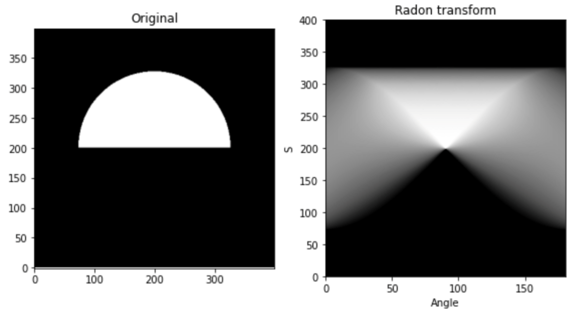


Figure 8: Image of the shape S in Example 3.13 and its Radon transform

Example 3.14. Consider a half circle

$$S = \{(x, y) : x^2 + y^2 \leq 1, y \geq 0\}.$$

Only the corner points $(-1, 0)$ and $(1, 0)$ are non-differentiable points of S. All point on a base of the half circle is locally linear and the other are non-differentiable. So, the Radon transform of S is not continuous only in the horizontal direction, $\theta = 0$. For the other directions, the Radon transform is continuous, see Figure 9. Again, for d-Radon transform, it is continuous in every direction.

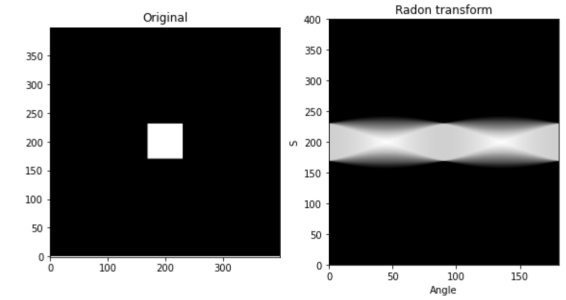


Figure 9: Image of the shape S in Example 3.14 and its Radon transform

4 Line Detection Methods

The line detection method is a procedure to indicate lines from a given image. In our setting, lines are parameterized by θ and r. So, line detection is a method to extract θ^{dt} and r^{dt} from each image. We can consider a line detection method as a set-valued function or correspondence that maps images to the set of detected parameters.

First, we consider the class $\mathfrak{B}_N$ of binary image consist of N circles on W. That is, $m \in \mathfrak{B}_N$ if $m : W \to \{0, 1\}$ in the form of

$$m(x, y) = \sum_{i=1}^{N} \delta_{B_{R_i}(P_i)} \tag{16}$$

where $B_R(P)$ denotes a circle centered at P with radius R.

We define a distance between images $m_1, m_2 \in \mathfrak{B}_N$ that consist of circles $B_{R_i}(P_i)$ for $i = 1, ..., N$ and $B_{S_i}(Q_i)$ for $i = 1, ..., N$ respectively by: $\bar{D} : \mathfrak{B}_N \times \mathfrak{B}_N \to [0, \infty)$ as $\bar{D}(m_1, m_2) =$

$$\min_{\sigma \in \Pi_N} \sum_{i=1}^{N} \left(2 \min\{R_i, S_{\sigma(i)}\} \|P_i - Q_{\sigma(i)}\| + \pi |R_i^2 - S_{\sigma(i)}^2| \right)$$

where Π_N denotes the set of permutations of $\{1, 2, ..., N\}$ and $\|\cdot\|$ is the Euclidean norm on $\mathbb{R}^2$. It was shown in [1] that the function $\bar{D}$ is a distance function on $\mathfrak{B}_N$. This distance is a specific case of the Wasserstein metric.

For the set of points, if all points are located on the given line $r^* = x \cos \theta^* + y \sin \theta^*$, we get that $(\theta^*, r^*) = \text{argmax}_{(\theta,r) \in E} \mathcal{R}[m](\theta, r)$ []. We can extend this result to the image consist of circles.

Theorem 4.1. *Let $m \in \mathfrak{B}_N$ and $d > 0$. If the center of all circles in m are on the line $r^* = x \cos \theta^* + y \sin \theta^*$, we get that*

1. $(\theta^*, r^*) = \text{argmax}_{\theta \in [-\frac{\pi}{2}, \frac{\pi}{2})} \mathcal{R}[m](\theta, r)$,
2. $(\theta^*, r^*) \in \text{argmax}_{\theta \in [-\frac{\pi}{2}, \frac{\pi}{2})} \mathcal{R}_d[m](\theta, r)$.

In particular, if $d < R_i$ for all radius R_i of circle consisted in m, we get that $(\theta^*, r^*) = \text{argmax}_{\theta \in [-\frac{\pi}{2}, \frac{\pi}{2})} \mathcal{R}_d[m](\theta, r)$.

The proof of this theorem can be found in [1]. The next step is to show the continuity of the d-Radon transform. Here, the maximum theorem is needed.

Lemma 4.2. *(Maximum Theorem, see Theorem 3.5 in [9]) Let X, Y to be topological spaces. For a continuous function $f : X \times Y \to \mathbb{R}$ and a compact value correspondence $C : Y \rightrightarrows X$ such that $C(y) \neq \emptyset$ for all $y \in Y$. Define a function $f^* : Y \to \mathbb{R}$ by $f^*(y) = \sup\{f(x, y) : x \in C(y)\}$ and a correspondence $C^* : Y \rightrightarrows X$ by $C^*(y) = \operatorname{argsup}\{f(x, y) : x \in C(y)\}$. If C is continuous at y, then f^* is also continuous at y and C^* is upper hemi-continuous at y.*

Theorem 4.3. *For $d > 0$ and $m \in \mathfrak{B}_N$. A correspondence $m \mapsto (\theta^{dt}, r^{dt})[m] = \operatorname{argmax}_{(\theta,r)\in E}\mathcal{R}_d[m](\theta, r)$ is upper hemi-continuous under the norm induced by metric $\bar{D}$.*

Proof. Let $d > 0$. Consider $\mathcal{R}_d$ as a function from $E \times \mathfrak{B}_N$ to $\mathbb{R}$ and a correspondence C defined by $C(m) = [-\frac{\pi}{2}, \frac{\pi}{2})$ for all $m \in \mathfrak{B}_N$. To show that $\mathcal{R}_d$ is continuous, let $m_1, m_2 \in \mathfrak{B}_N$ and $(\theta_1, r_1), (\theta_2, r_2) \in E$. By linearity of the d-Radon transform, we have that

$$
\begin{aligned}
&|\mathcal{R}_d[m_1](\theta_1, r_1) - \mathcal{R}_d[m_2](\theta_2, r_2)| \\
&= |\mathcal{R}_d[m_1](\theta, r) - \mathcal{R}_d[m_1](\theta_2, r_2) \\
&\qquad + \mathcal{R}_d[m_1](\theta_2, r_2) - \mathcal{R}_d[m_2](\theta_2, r_2)| \\
&\leq |\mathcal{R}_d[m_1](\theta, r) - \mathcal{R}_d[m_1](\theta_2, r_2)| \\
&\qquad + |\mathcal{R}_d[m_1](\theta_2, r_2) - \mathcal{R}_d[m_2](\theta_2, r_2)|
\end{aligned}
$$

□

The last step follows from the fact that $\mathcal{R}_d$ is continuously differentiable. As a consequence, $\mathcal{R}_d : E \times \mathfrak{B}_N \to \mathbb{R}$ is continuous. Then, by maximum theorem, the correspondence $m \mapsto (\theta^{dt}, r^{dt})[m] = \operatorname{argmax}_{(\theta,r)\in E}\mathcal{R}_d[m](\theta, r)$ is upper hemi-continuous. So, if $\mathcal{R}[m]$ has a unique maximum, then a correspondence $m \mapsto (\theta^{dt}, r^{dt})[m] = \operatorname{argmax}_{(\theta,r)\in E}\mathcal{R}_d[m](\theta, r)$ is a continuous function.

Hence, we can extend this fact from an image consisting of circles to convex shapes. The main theorem can be restated as follows:

Theorem 4.4. *For $d > 0$ and image m consist of convex shape, if every point of edge of every shape in m are not locally linear, then correspondence $m \mapsto (\theta^{dt}, r^{dt})[m] = \operatorname{argmax}_{(\theta,r)\in E}\mathcal{R}_d[m](\theta, r)$ is upper hemi-continuous.*

Here, the metric used in this theorem is defined in the same way as $\bar{D}$ but on the image consists of convex shapes instead of circles.

5 Conclusion

The regularity of the Radon transform, its modification called the d-Radon transform, and line detection methods based on these two transforms have all been investigated. The line detection method based on the Radon transform is not continuous for an image consisting of a circle. After more analysis, we determined that the circle has a differentiable edge at each point, which is the source of discontinuity. The locally linear point, differentiable point, and non-differentiable point are three types of shape edges. The Radon transform's regularity is determined by the sort of point on edge in that direction. A non-differentiable point is the only situation in which the Radon transformation can preserve its continuity. To achieve Radon transformation continuity in all directions, the shape must not be differentiable at all points, which is impossible.

To increase the regularity of the line detection method, we introduce the d-Radon transform by changing the domain of integration from line to strip. Equivalently, it integrates the Radon space in the r direction. The d-Radon transform of a convex shape was explored for regularity. We discovered that the differentiable requirement of the transformation is that the shape's edge is not locally linear. This is also a condition for the line detection method's continuity.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment The authors very much appreciate the support by my beloved native friend, Phatthamon Saenmuk, for editing my broken English.

References

[1] P. Vatiwutipong, "Continuity of line detection methods based on the Radon transform," in 2022 14th International Conference on Knowledge and Smart Technology (KST), 29–33, 2022, doi:10.1109/KST53302.2022.9729056.

[2] J. Radon, "On the determination of functions from their integral values along certain manifolds," in IEEE Transactions on Medical Imaging, 170–176, 1986, doi:10.1109/TMI.1986.4307775.

[3] S. R. Deans, The Radon Transform and Some of Its Applications, Springer, 2007.

[4] S. R. Deans, "Hough Transform from the Radon Transform," IEEE Transactions on Pattern Analysis and Machine Intelligence, **2**, 185–188, 1981, doi: 10.1109/TPAMI.1981.4767076.

[5] G. T. Herman, Fundamentals of Computerized Tomography, Springer, 2009.

[6] M. C. B. M. P. A. M. M. Riccardo Aramini, Fabrice Delbary, "Hough Transform from the Radon Transform," arXiv:1605.09201, 2016, doi:10.48550/arXiv.1605.09201.

[7] C. L. Epstein, Introduction to the Mathematics of Medical Imaging, Society for Industrial and Applied Mathematics, 2009.

[8] T. G. Feeman, The Mathematics of Medical Imaging: A Beginner's Guide, Springer, 2015.

[9] N. S. P. Shouchuan Hu, Handbook of Multivalued Analysis 1: Theory, Springer, 1997.

Advances in Science, Technology and Engineering Systems Journal Vol. 7, No. 4, 127-139 (2022)

www.astesj.com

ASTESJ

ISSN: 2415-6698

BER Performance Evaluation Using Deep Learning Algorithm for Joint Source Channel Coding in Wireless Networks

Nosiri Onyebuchi Chikezie[1,*], Umanah Cyril Femi[1], Okechukwu Olivia Ozioma[2], Ajayi Emmanuel Oluwatomisin[3], Akwiwu-Uzoma Chukwuebuka[1], Njoku Elvis Onyekachi[1], Gbenga Christopher Kalejaiye[3]

[1]*Department of Electrical and Electronic Engineering, Federal University of Technology, Owerri, 460114, Nigeria*

[2]*Department of Information System and Security Engineering, Concordia University Monstreal H3G 1M8, Canada*

[3]*Department of Electrical and Electronic Engineering, University of Lagos, 101017, Nigeria*

ARTICLE INFO

Article history:
Received: 21 May, 2022
Accepted: 18 August, 2022
Online: 29 August, 2022

Keywords:
JSCC
FCNN
Deep Learning
Bit Error Rate
Block Error Rate
Neural Network
BLSTM

ABSTRACT

In the time past, virtually all the contemporary communication systems depend on distinct source and channel encoding schemes for data transmission. Irrespective of the recorded success of the distinct schemes, the new developed scheme known as joint source channel coding technique has proven to have technically outperformed the conventional schemes. The aim of the study is centered in developing an enhanced joint source-channel coding scheme that could mitigate some of the limitations observed in the contemporary joint source channel coding schemes. The study tends to leverage on recent developments in machine learning known as deep learning techniques for robust and enhanced scheme, devoid of explicit code dependence for the signal compression and as well in error correction but learn automatically on end-to-end mapping structure for the source signals. It primarily aimed at providing an improved channel performance approach for wireless communication network. A deep learning algorithm was implemented in the study, the scheme focused on improving the Bit Error Rate (BER) performance while reducing latency and the processing complexity in Joint Source Channel Coding systems. The deep learning autoencoder model was deployed to compare with the hamming code, convolution code, and uncoded systems. JSCC using neural networks were simulated based on BER performance over a range of energy per symbol to noise ratio (E_b/No). Training and test error for the fully connected neural network autoencoder models on channels with 0.0dB and 8.0dB were carried out. The results obtained showed that the autoencoder model had a better BER performance when compared with the convolution code and uncoded systems, it also outperformed the uncoded BFSK with an approximately equal BER performance when compared with the hamming code (soft decision) decoding system.

1. Introduction

The world has recently witnessed a great revolution in the way information is transmitted from one place to another. Wireless communication has advanced from mere point-to-point communication to becoming a viable tool to facilitate economic development, security enhancement and reliable public service delivery. The basic task for a communication system is to reliably deliver information from the source to the destination, using a transmitter and a receiver across a channel. The performance of conventional communication techniques is seen limited in operation and are sub-optimal due to the challenges which present themselves in the form of latency, reliability, energy efficiency, flexibility etc [1]. The fast emergence of many unprecedented services such as artificial intelligence, smart homes, factories and cities, wearable devices for physical challenged, robots, autonomous vehicles, big data, internet-of-things etc. are challenging the conventional approaches and mechanisms to communication. Recent research and technology advancements have contributed to an enviable progress in developing novel and enhanced mechanisms in the layers of communication system.

[*] Corresponding Author: Nosiri Onyebuchi Chikezie, buchinosiri@gmail.com

https://dx.doi.org/10.25046/aj070417

Despite that, more research is on the progress in providing optimal performance for wireless communication networks.

It is important to note that emerging wireless communication systems typically transmit high data rates to provide wide range of services for better voice quality, improved data, images and other multimedia applications. Conversely, during wireless signal propagation, the systems usually encounter channel impairments, resulting in data errors at the receiver end. To correct this, requires adequate error correction codes to detect and correct symbol errors during transmission.

The introduction of Joint Source-Channel Coding (JSCC) technique in wireless communication has been able to address most of the challenges that are inherent in the separation-based schemes, (i.e., the conventional two-step encoding process for the image/video data transmission, source coding and channel coding) [2]. Recently, it became a considerable research topic in communication systems and information technology, with the application in areas like audio/video and satellite transmission, as well as in space exploration. Despite the successes recorded by JSCC techniques, it still encounters some performance flaws inherent in its fundamental assumptions that could prove very costly for modern communication systems. This flawed assumption ripples through the design of systems based on conventional JSCC techniques in the form of increased processing and algorithmic complexity to combat noise in its various forms and also cater for additive information. This complexity can introduce a certain level of latency which is detrimental to the actualization of low latency systems. Furthermore, other inherent limitations include inability to fulfil bulky data and very high-rate communication requirements in multifaceted conditions as seen in most complex channel models. Others include; in low latency communication systems, in rapid and reliable signal processing application and in limited and sub-optimal block structures, due to the fixed block configuration of the communication system etc [1]. However, the recent introduction of Deep Learning (DL) technique and its fundamental based autoencoder concept, characterized with its simplicity in implementation, flexibility and ability in adapting to complex channel models, has been able to handle most of the complexities due to the stated advantages it possesses [1]. It has recently been successfully applied in solving various real-life applications such as in pattern recognition, speech and language processing, media entertainment, medicine, biology and security systems. DL is quite robust and scalable in application.

Deep learning is a subset of Machine Learning (ML) that exhibits greater potentials in building complex concept from simpler concepts. It has useful tools to process ultra-high data and shows high performance accuracy in recognition and prediction. Deep learning algorithm is seen to outperform machine learning algorithm, especially in handling difficult and complex tasks such as in image and voice recognition, it is considered to be more valuable in cases where needful reduction in computational complexities and overhead processing are preferred. Deep learning tends to rely on its intelligence to define its own finest features, it does not require humans to perform any feature-creation activity. Among the existing DL models, Deep Neural Networks (DNNs) are considered to be the most known model, other deep architectures such as Neural Processes(NPs), Deep Gaussian Processes (DGPs) and Deep Random Forests (DRFs) could be categorized as deep models made up of multiple layered structures [1,3].

Based on the research motivations, the study focuses on implementing JSCC using DL approach without the need for explicit codes. The study aims to develop a Deep Neural Network (DNN) symbol models for JSCC in an end-to-end pattern. Python/Keras and TensorFlow backend are simulation tools used to evaluate the error correction performance and data reconstruction. Bit Error Rate (BER) and Block Error Rate (BLER) are the selected parameters for the system analysis. Simulations were carried out to perform the BER/BLER and its reliability compared to the conventional communication approach with preference in reducing the processing complexity and latency. The performance analysis of the developed deep joint source-coding algorithm with different Signal-to-Noise Ratio (SNR) values were also evaluated.

In our study, the motive is to extend the preceding study on autoencoder-based end-to-end learning of communications system, evaluate its characteristic performance in varying system configurations and also realize the potentials of autoencoder-based end-to-end learning mechanisms for communications systems.

1.1 Model of A Simple Communication System

A typical communication system in its most fundamental form consists of a transmitter, channel, and receiver as illustrated in figure1. The communication system facilitates the transfer of information signal from one point to another through a process that involves three basic stages; coding, mapping and decoding. Firstly, the information signal is encoded into a message x of block length k. Each message x can be represented in the block length $k = log_2(M)$ number of bits. The transmitter can transmit any of the acceptable messages out of M possible messages $x \in M = \{1, 2, ..., M\}$ of block length k through n discrete users of the allocated communication channel. The transmitter on the other hand, performs the mapping $f_\theta: C^k \rightarrow C^n$ [4]. A vector g of n complex symbols is transmitted across the channel to facilitate the sending of the message x to the receiver[1]. The presence of noise in the channel causes signal distortions to the transmitted symbols. The receiver stage is concerned with mapping the transmitted signal to the receiver. The mapping of the transmitted signal is actualized using the transformation $g_\theta: C^n \rightarrow C^k$, portraying the fact that the decoding function inverts the operation of the encoding function. At the receiving end, as the signal is intercepted by the receiver, that is the signal $i \in C^n$, the receiver generates the estimate $\tilde{x}$ of the originally transmitted message x. Figure 1 shows the structure of a simple communication system which could be modelled using an autoencoder. A communication system can in its simplest term be described as an autoencoder that tries to reconstruct the transmitted message at the receiver as accurately as possible with the least possible errors[1]. For clarity sake, we can further describe the encoder function of the autoencoder as the transmitter block while the decoder function as the receiver block of the system. A block diagram of an autoencoder is represented in figure 2.

From figure 2, the encoder attempts to transform the input value x into a low dimensional latent vector $z = f(x)$. The latent vector is usually characterized of low dimension with a

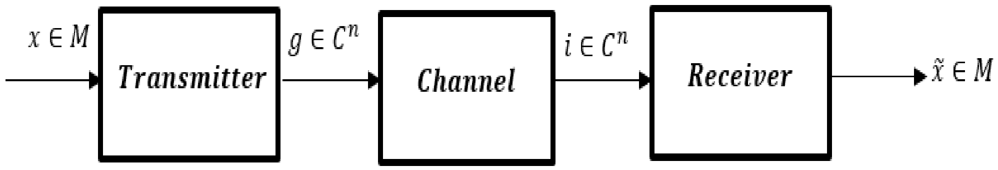


Figure 1: Block diagram of a Communication System

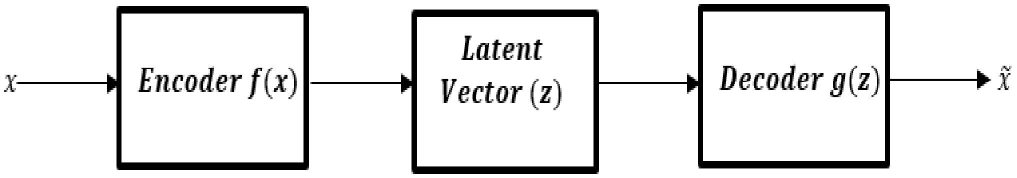


Figure 2: A block diagram representing an Autoencoder

compressed representation of the input distribution. The decoder in contrast, tries to recover the input signal from the latent vector, g\left(z\right)=\widetilde{x} . The expectation should be that the output recovered by the decoder could only approximate the input (i.e. making \widetilde{x}\ as close as possible to x)[5]. The variations between the input and the output is measured as a loss function. It is necessary to note that both the encoder and the decoder are non-linear functions.

2. Related Works

In recent years, DL-based techniques were introduced for different processing blocks of the wireless communication systems as substitutes to conventional applications such as modulation recognition[6], channel encoding and decoding[7,8], and channel estimation and detection [9–11], owing to the development of DL algorithms and system architectures.

Authors of [1], investigated the DL-based end-to-end communications performance models when deployed in a single user communication network under an Additive White Gaussian Noise (AWGN) channel. An autoencoder-based end-to-end communications system was implemented in the system validation. In[12], considered the challenges of JSCC of text /structured data using deep learning approach from natural language processing over noise channel. Their proposed technique is said to have an edge over the existing distinct source and channel coding, particularly in scenarios when a smaller number of bits were used in describing each sentence. Their scheme achieved lower word error rates from the developed deep learning-based encoder and decoder system. The developed system uses a fixed bit length for enconding sentences of different length. This was observed to be a major drawback of their algorithm.

The authors of[13], proposed the use of neural networks to address the design of systems with block length when k =1. In [14], used simple neural network architecture in encoding and decoding of Gauss-Markov sources over additive white Gaussian noise channel. Authors of [15,16], proposed neural network for signal compression devoid of a noisy channel (i.e. only source coding), where image compression algorithms were developed using RNNs. In[17], used neural networks, in particular, Variational Autoencoders (VAEs) to design neural network based Joint Source Channel Coding and extended the system design to where k ≠1. However, their performance was reasonable but a lot was required to improve upon their performance in order to meet up with the benchmark set by[18] . The authors of,[18] developed a new scheme for JSCC of Gaussian sources over AWGN channels. VAEs was implemented in their design but with a novel encoder architecture for the VAE specifically developed for zero-delay Gaussian JSCC over AWGN channels, a situation where the source dimension (*m*) is greater than the channel dimension (*k*). Their proposed scheme was able to improve on works of [17] with about 1dB.

Our study therefore seeks to evaluate the performance of Bit Error Rate in wireless networks using Deep Neural Network (DNN) system model for joint source channel coding in an end-to-end manner without the need for explicit codes to provide error correction. The approach is envisaged to minimize the block length of transmitted data with maximal utilization of bandwidth, increased data rate and power efficiency. Simulation models such as Python/Keras and TensorFlow backend will be implemented to oversee the process of error correction improvement and data reconstruction.

3. Method

The proposed deep learning approach for JSCC is implemented by simulation in Keras using TensorFlow as its backend. TensorFlow provides a robust environment, creating a relatively easy-to-use package. A model is trained to mimic the conventional end-to-end communication system under certain conditions and constraints. The trained model is then tested against random data under varying conditions to determine its performance in practical scenarios.

3.1. Autoencoder Implementation for the Proposed Scheme

An autoencoder’s main objective is to actualize a compressed representation of a given input data. An autoencoder is a neural network architecture, comprised of two distinct units; Encoder and Decoder functional units. The encoder unit primarily converts the input data into a different representation while the decoder unit

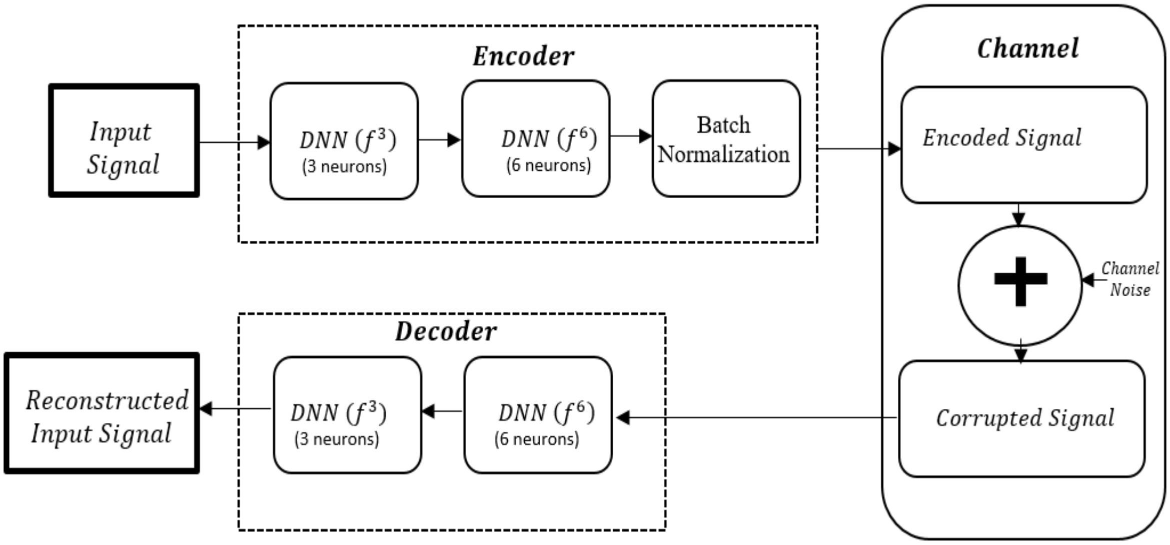


Figure 3: The FCNN Autoencoder Block Diagram

converts the new representation back into the original format, trying to recover the input data. The input data could be in different configurations such as in speech, text, image or video format. Figure 3 represents the functional block diagram of a Fully Connected Neural Network (FCNN) autoencoder.

To illustrate the performance characteristics of the deep JSCC scheme, the functional block diagram of the Fully Connected Neural Network in Fig. 3 is implemented in our simulation. The autoencoder in this context was implemented as a fully connected feedforward Neural Network, enabled to propagate the information realized from the input, through a sequence of non-linear transformations to get to the output.

The autoencoder models is assumed to have undergone training for a predetermined message size (*M*) with its accompanying communication rate. We have two hidden feedforward DNN layers situated at the encoder end as shown in Fig.3. The first layer is f^3 having 3 neuros while the second layer is f^6which constitute 6 neurons. It is designed in such a way that the output of the first layer feeds into the input of the second layer etc. The two-layer blocks are connected to the batch normalization layer as represented. The batch normalization layer is introduced in the representation to satisfy the average power constraint. An activation function known as Rectified Linear Activation Unit (ReLU) was employed by the convolutional layers (each dense layer) in order to apply nonlinearities to the model. A SoftMax activation function was implemented at the output layer in order to output the probability distributions for each of the output category. We used the Gaussian noise layer to simulate an additive white Gaussian noise channel which in this case is represented as the noise layer.

The autoencoder is trained at full length over the stochastic channel model. The Stochastic Gradient Descent (SGD) method of optimization is used and the Adam optimizer is the preferred choice for the optimizer. The Adam optimizer's learning rate is set at 0.001. The steps taken to select the energy per symbol to noise ratio *(E_b/No)* values for the AWGN channel during training are shown thus:

i. Training was done at a fixed E_b/N_0 value, 0 dB and 8 dB in this case.
ii. Testing of the trained model using random E_b/N_0 values picked from a predetermined E_b/N_0 range for each training epoch. This is done to determine the BER performance during varying channel conditions.
iii. The testing is initiated using a higher E_b/N_0 value which decreases gradually along training epochs. In the case of the 8 dB training value, the test starts from 8 dB and is reduced by 2 dB after every 10 epochs.

We applied autoencoder model for our training and testing analysis in Keras using the TensorFlow application as its default tensor backend engine. The model was trained for fifty (50) epochs using sixty thousand (60,000) images, generated randomly with E_b/N_0 values for Additive White Gaussain Noise (AWGN) channel in the model training. The BER performance for the 0dB and 8dB were then compared with the Hamming code utilizing a BPSK modulation scheme.

The Bidirectional Long Short-Term Memory (BLSTM) autoencoder is also trained and tested in this simulation. It exhibits a parallel architecture to FCNN autoencoder, though, it has dissimilar structural components as shown in figure 4.

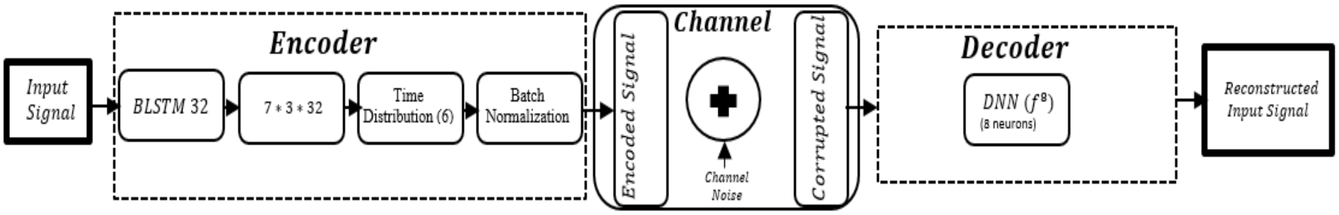
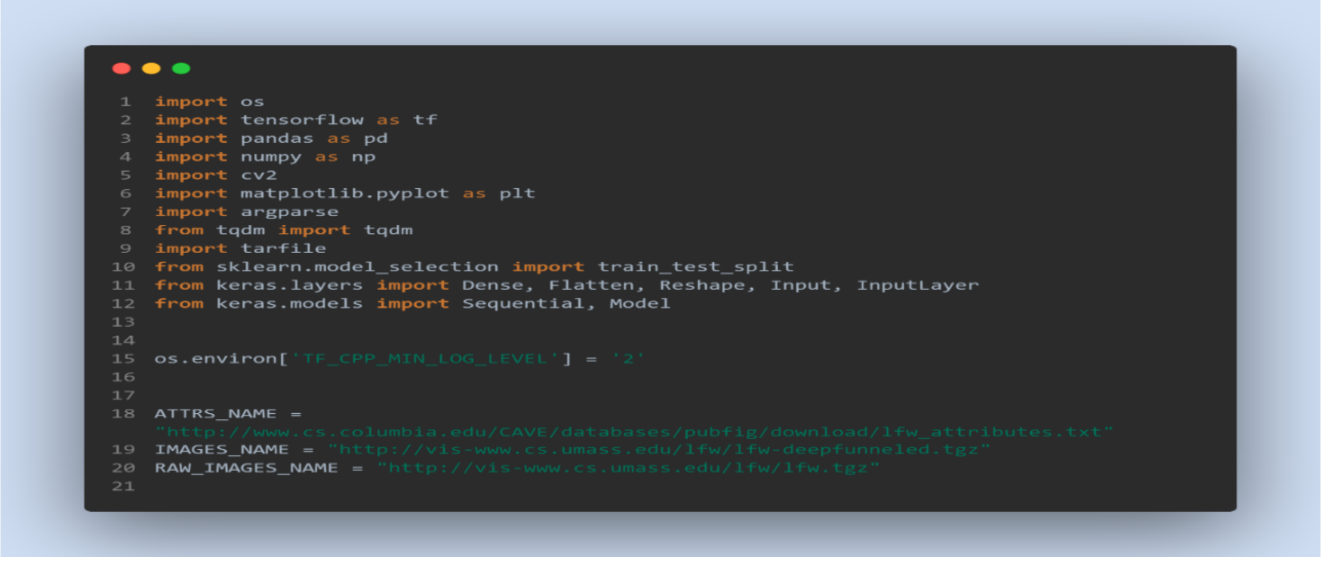


Figure 4:The BLSTM autoencoder Block diagram.



```
import os
import tensorflow as tf
import pandas as pd
import numpy as np
import cv2
import matplotlib.pyplot as plt
import argparse
from tqdm import tqdm
import tarfile
from sklearn.model_selection import train_test_split
from keras.layers import Dense, Flatten, Reshape, Input, InputLayer
from keras.models import Sequential, Model


os.environ['TF_CPP_MIN_LOG_LEVEL'] = '2'


ATTRS_NAME =
"http://www.cs.columbia.edu/CAVE/databases/pubfig/download/lfw_attributes.txt"
IMAGES_NAME = "http://vis-www.cs.umass.edu/lfw/lfw-deepfunneled.tgz"
RAW_IMAGES_NAME = "http://vis-www.cs.umass.edu/lfw/lfw.tgz"

```

Figure 5: Snapshot of FCNN generated codes showing imports

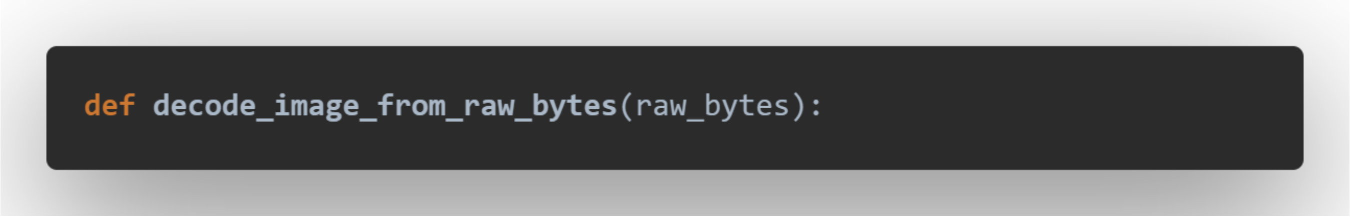
In figure 4, the encoder with a BLSTM has its dimension of the hidden units set to 32, while 32 refers to the number of features in each input sample. The BLSTM cell comprise of seven (7) hidden states units. In the context, each input character or element is linked to each neuron in the hidden layer. The product of the input feature and the size of the hidden layer is evaluated as the total number of connections established. The time distributed layer at the encoder section is introduced to help in flattening the output from the previous layer. At the decoder unit is implemented with a DNN layer constituting 8 neurons.

3.2 Deep Learning Algorithm for Fully Connected Neural Network (FCNN) Model

The basic steps followed to train and test the designed model in section 3.2 are highlighted as follows:

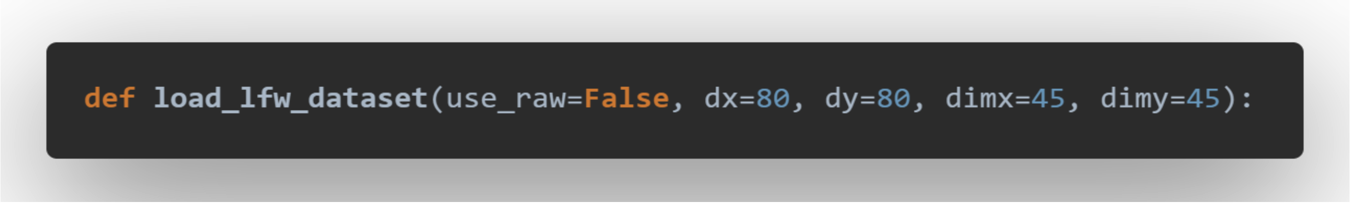
A. The imports for the model are specified. A snapshot of the code in figure 5 shows that lots of TensorFlow modules were imported. The snapshot also showed that the Labelled Faces in the Wild (LFW) dataset was utilized to train the model.

B. The dataset was loaded from its location on the internet and normalized. This entails converting the raw matrix into an image and changing the color system to RGB. The screenshots of the system functions are shown in figures 6 and 7.



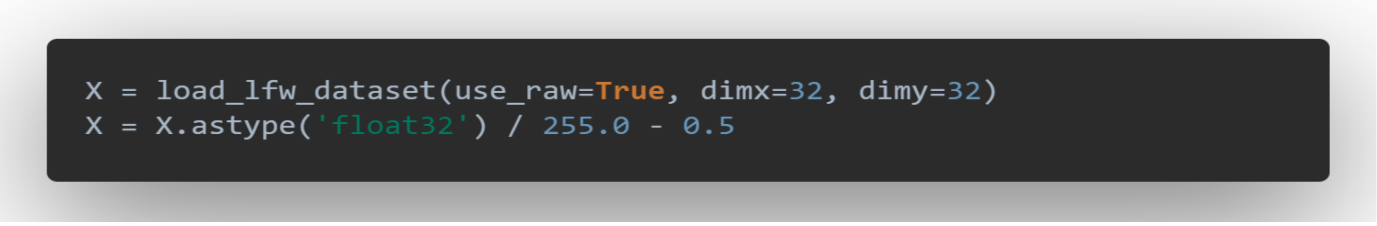
```
def decode_image_from_raw_bytes(raw_bytes):
```

Figure 6: Snapshot of function that converts raw matrix to image



```
def load_lfw_dataset(use_raw=False, dx=80, dy=80, dimx=45, dimy=45):
```

Figure 7: Snapshot of function that loads LFW dataset.



```
X = load_lfw_dataset(use_raw=True, dimx=32, dimy=32)
X = X.astype('float32') / 255.0 - 0.5
```

Figure 8: Snapshot of dataset normalization

The images could have large values for every pixel from 0 to 255 range. Usually, in ML, the focus is to make sure the values are small and concerted around 0. This concept adopted enabled our model to train faster with optimal results. This task is achieved through normalization of the dataset as shown in figure 8.

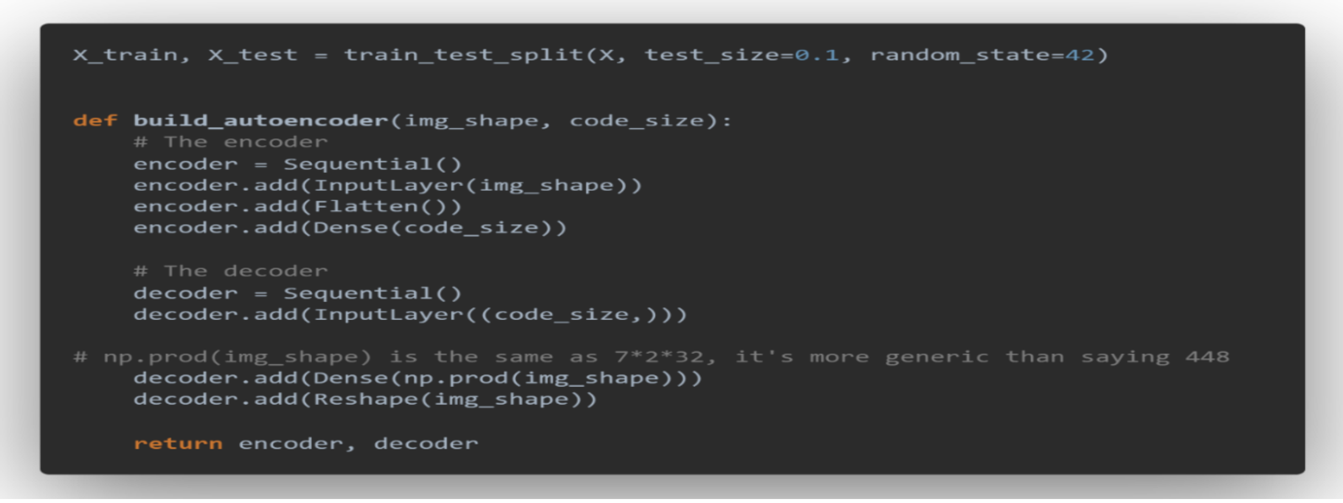
The dataset was split into training and test data sets. The training data is used in building the autoencoder. The algorithm for this is shown in figure 9.

C. The model was compiled in order to enable us train the model. The optimizer and loss function are specified in this stage. Figure 10 shows a snapshot of the algorithm used to accomplish the task

D. A summary of the model was generated to inspect the model in greater detail. The generated model summary is shown in figure 11.

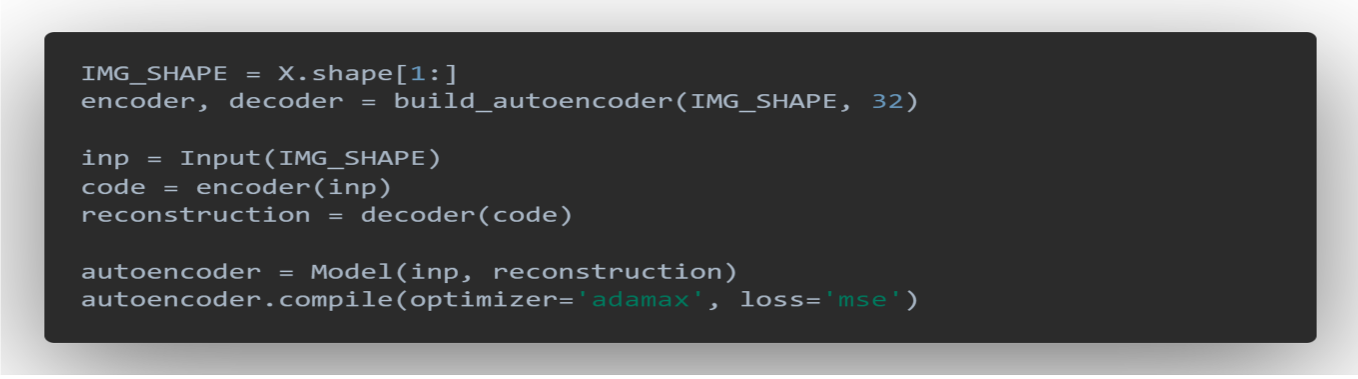
E. Finally, the model was trained and tested by simulating practical channel conditions. Noise was introduced into the model prior to testing the model. Figure 12, shows the function used to introduce noise into the model while figure 13, shows the algorithm for training of the model at a set SNR.

The result for the simulation over 50 epochs is captured in figure 14.



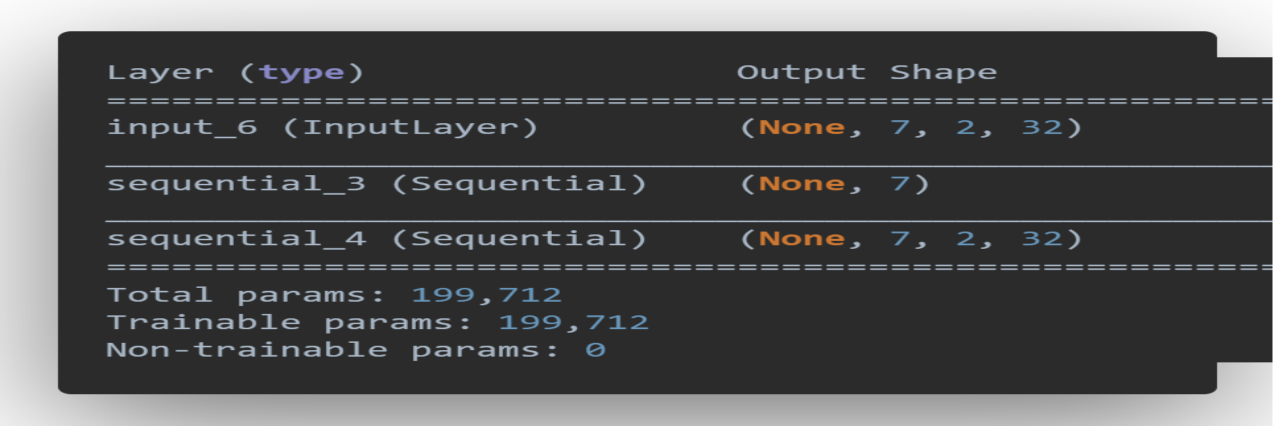
```
X_train, X_test = train_test_split(X, test_size=0.1, random_state=42)


def build_autoencoder(img_shape, code_size):
    # The encoder
    encoder = Sequential()
    encoder.add(InputLayer(img_shape))
    encoder.add(Flatten())
    encoder.add(Dense(code_size))

    # The decoder
    decoder = Sequential()
    decoder.add(InputLayer((code_size,)))

# np.prod(img_shape) is the same as 7*2*32, it's more generic than saying 448
    decoder.add(Dense(np.prod(img_shape)))
    decoder.add(Reshape(img_shape))

    return encoder, decoder
```

Figure 9: Snapshot of algorithm for building the FCNN autoencoder



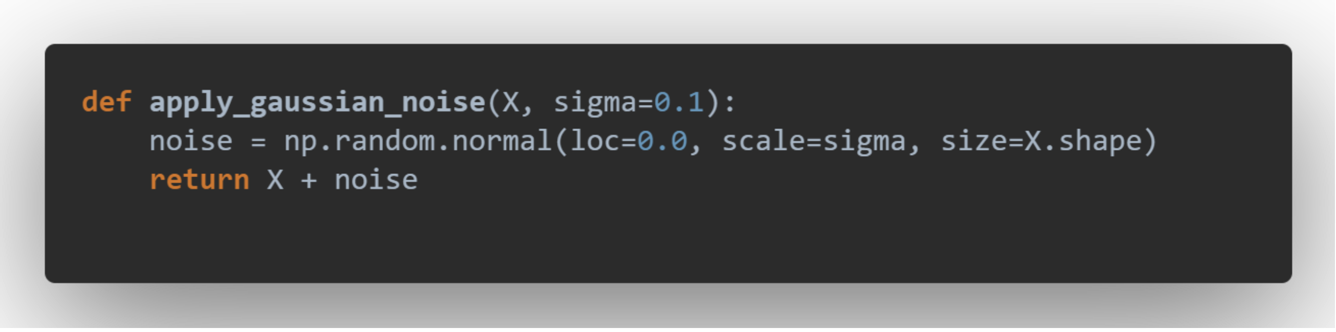
```
IMG_SHAPE = X.shape[1:]
encoder, decoder = build_autoencoder(IMG_SHAPE, 32)

inp = Input(IMG_SHAPE)
code = encoder(inp)
reconstruction = decoder(code)

autoencoder = Model(inp, reconstruction)
autoencoder.compile(optimizer='adamax', loss='mse')
```

Figure 10: Snapshot of algorithm for compiling model



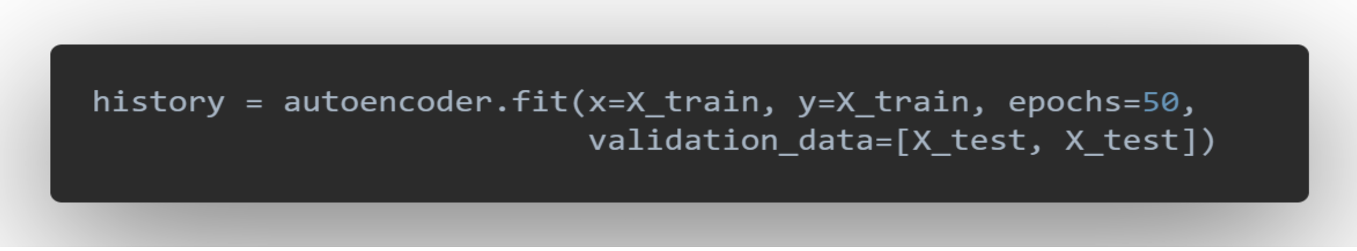
```
Layer (type)                 Output Shape
==============================================
input_6 (InputLayer)         (None, 7, 2, 32)
______________________________________________
sequential_3 (Sequential)    (None, 7)
______________________________________________
sequential_4 (Sequential)    (None, 7, 2, 32)
==============================================
Total params: 199,712
Trainable params: 199,712
Non-trainable params: 0
```

Figure 11: Snapshot of FCNN model summary



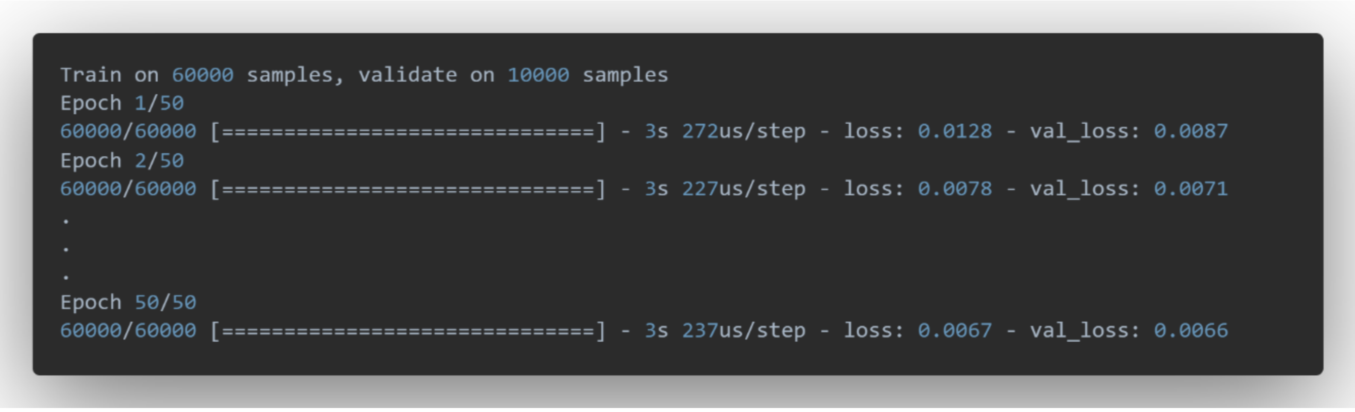
```
def apply_gaussian_noise(X, sigma=0.1):
    noise = np.random.normal(loc=0.0, scale=sigma, size=X.shape)
    return X + noise
```

Figure 12: Snapshot of algorithm to introduce noise into the FCNN model



```
history = autoencoder.fit(x=X_train, y=X_train, epochs=50,
                          validation_data=[X_test, X_test])
```

Figure 13: Snapshot of algorithm to train the FCNN model



```
Train on 60000 samples, validate on 10000 samples
Epoch 1/50
60000/60000 [==============================] - 3s 272us/step - loss: 0.0128 - val_loss: 0.0087
Epoch 2/50
60000/60000 [==============================] - 3s 227us/step - loss: 0.0078 - val_loss: 0.0071
.
.
.
Epoch 50/50
60000/60000 [==============================] - 3s 237us/step - loss: 0.0067 - val_loss: 0.0066
```

Figure 14. Snapshot of training results for the FCNN model

The steps to implement the algorithm for the LSTM and BLSTM autoencoder follow the same process and pattern that had earlier been elaborated. The only key difference is the introduction of state in the LSTM and BLSTM.

3.3 Performance Measure

The model was trained using the mean squared error (MSE) and categorical cross-entropy. The MSE together with two other error metrics were used to give an insight into the performance of the model. The average mean squared-error between the original input image x and reconstruction $\hat{x}$ at the output of the decoder is taken to be the loss function[2]. The loss function is given as[2]:

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N} d(x_i, \hat{x}) \tag{2}$$

Where $d(x,\hat{x}) = \frac{1}{n}||\mathrm{x} - \hat{x}||^2$ is the mean squared-error distortion and N = 7 in the simulation. This represents the distance apart the approximation is far from the original input. The other two error metrics factored into the simulation are the Bit Error Rate (BER) and the Block Error Rate (BLER) on channels with 0.0 dB and 8.0 dB Eb/No respectively. The BER is the number of bit errors divided by the total number of transferred bits during a studied time interval (https://en.wikipedia.org/wiki/Bit_error_rate). Bit errors in this context is simply the number of bits that are incorrectly reconstructed. Block Error Rate (BLER) refers to as the ratio of the number of blocks with error to the total number of blocks transmitted on a digital circuit. BER is affected by several factors including noise in the channel, code rate and the transmitter power level. The BLSTM autoencoder model was simulated and compared with the Hamming codes for soft and hard decoder.

The code rate R is given as[19]:

$$R = \frac{k}{n} \tag{3}$$

where k refers to the number of bits at the encoder input and n is the number of bits at the encoder output. The variance of additive white Gaussian noise is given as[19]:

$$\beta = (2RE_b/N_o)^{-1} \tag{4}$$

The Mean Square Error (MSE) is the variance around the fitted regression line at the decoder. It could also be referred as the Euclidean distance between the reconstructed vector $\hat{v}_i$ and the input vector v_i and indicates the distance apart the approximation is from the original input. The MSE which could be refered to as an example of a loss function is represented in equation 5[4].

$$MSE = \frac{1}{N}\sum(v_i - \hat{v}_i) \qquad (5)$$

To evaluate the reconstruction accuracy of the deep JSCC algorithm in a noisy channel, an additive white Gaussian noise is modelled in the system. The average power constraint, P, is set to one (i.e. P = 1), and vary the channel SNR by varying the noise variance N_0. The channel SNR is computed as[20]:

$$SNR = 10log_{10}\frac{P}{N_0}dB \qquad (6)$$

The performance of the deep JSCC algorithm is measured in terms of the Peak Signal-to-Noise Ratio (PSNR) of the reconstructed images at the output of the decoder, defined as follows[20]:

$$PSNR = 20log_{10}\left(\frac{255}{\sqrt{MSE}}\right)dB \qquad (7)$$

All simulations were conducted on 24-bit depth RGB images (8 bits per pixel per color channel), thus, maximum power signal is given by $2^8 - 1 = 255$.

4. Simulation Results

The performance of the JSCC for wireless image transmission were evaluated using computer simulations. Simulation results of the JSCC using neural networks were based on the bit error rate (BER) performance over a range of signal-to-noise ratios. The results of the simulations were captured and displayed using graphical plots of errors versus epoch units (the number of passes through the complete dataset) for a given SNR value.

Figures 15 and 16 showed the training and test error results for the fully connected neural network autoencoder models trained on channels with 0.0dB and 8.0dB respectively. We trained the models using 60,000 samples of data, tested on 10,000 samples. Checkpointing, a fault tolerance technique for long running processes was used to retain the model state that yielded the best loss. It is an approach where a snapshot of the state of the system is taken in case of system failure.

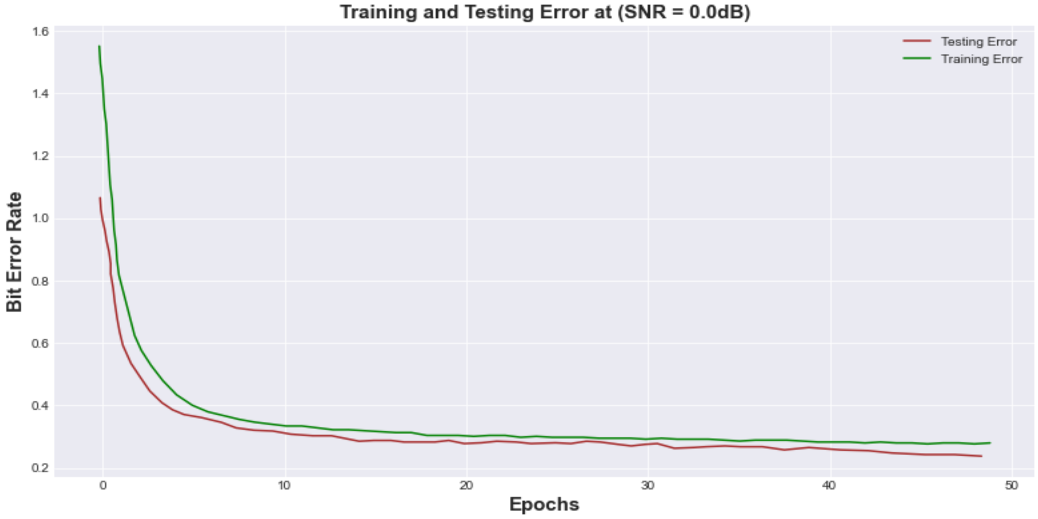


Figure 15: Bit Error Rate vs Training Epochs for FCNN Autoencoder at (SNR= 0.0 dB) channel.

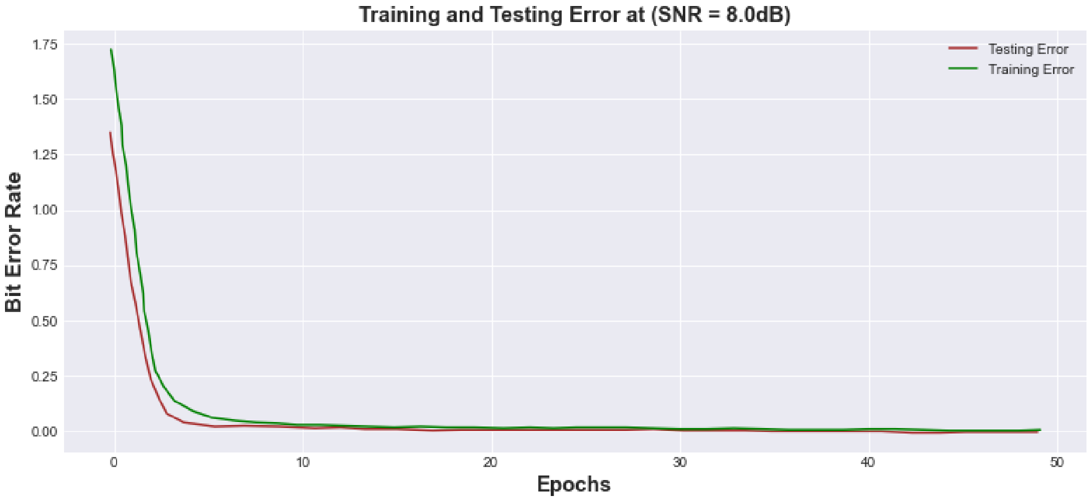


Figure 16: Bit Error Rate vs Training Epochs for FCNN Autoencoder at (SNR= 8.0 dB) channel

Figures 17 and 18 represent the error results from the training/test for the BLSTM Autoencoder model for SNR of 0dB and 8 dB respectively. From the graphical representation, it was observed that BLSTM autoencoder converges at a bit faster rate compare with the FCNN and appears to have a better solution to the problem of effective input reconstruction.

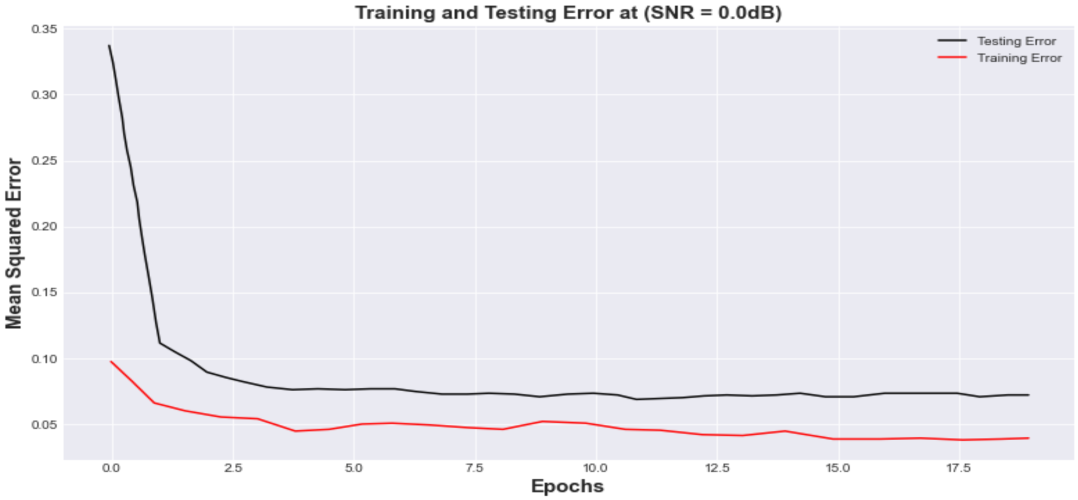


Figure 17: Mean Squared Error vs Training Epochs for BLSTM Autoencoder at (SNR= 0.0 dB) channel

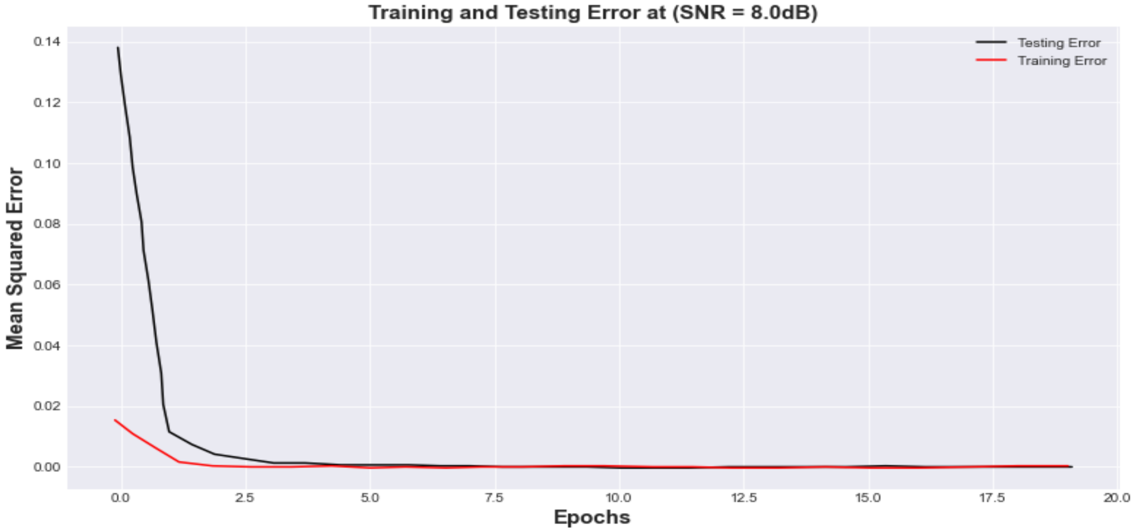


Figure 18: Mean Squared Error vs Training Epochs for BLSTM Autoencoder at (SNR= 8.0 dB) channel

4.1 Discussion

The simulated results shown in section 4, highlighted the robustness of the proposed coding scheme to various channel conditions. Figures 15-18 illustrate the number of errors in the reconstructed images versus the number of Epoch units for two different SNR values of the AWGN channel. Each curve in the figure is obtained by training the end-to-end system using a specific channel signal to noise ratio value. The performance of the learned encoder/decoder parameters on the 10,000 test images for slightly varying SNR value due to varying channel conditions were also evaluated. When the SNR_{test} is less than the $\text{SNR}_{\text{train}}$, our deep Joint Source Channel Coding algorithm is seen to demonstrate a robust scheme performance over channel deterioration and failed to experience recurrent cliff effect observed in digital systems, where the quality of the decoded signals experience nongraceful degradation whenever the SNR_{test} drops below a critical value close to the $\text{SNR}_{\text{train.}}$ The deep JSCC design performance is more tolerable with a better stable state in the presence of channel fluctuations and exhibits a graceful degradation as the channel deteriorates. The performance is due to the autoencoder's potential to map analogous images to nearby points in the channel's input signal space. On the other hand, when SNR_{test} is said to increase above $\text{SNR}_{\text{train}}$, a gradual improvement is observed in the quality of the reconstructed images. It is important to note that the performance in the saturation region is obtained majorly by the level of compression realized during the training phase for a given value of $\text{SNR}_{\text{train}}$.

4.1.1 Deep Learning JSCC Versus Hamming Code

The deep JSCC algorithm is compared with the channel uncoded BPSK and Hamming code. A version of the (7,4) hamming code was implemented in the study as a measure for the experimental simulation comparison as shown in figure 19.

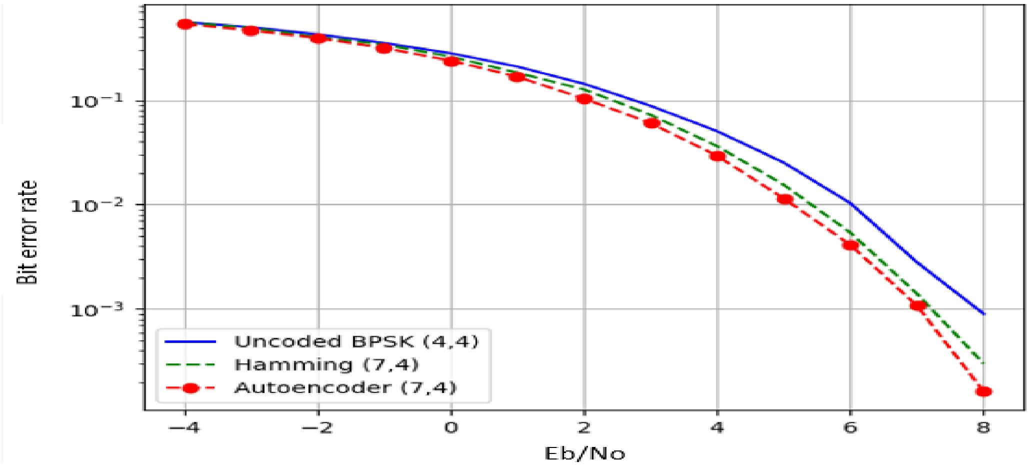


Figure 19: Bit Error Rate for BPSK in AWGN Channel Hamming Code and Autoencoder

Figures 20 and 21 represent the comparison of the BER and BLER respectively with Hamming code and uncoded BPSK. It can be observed that, the binary long short-term memory autoencoder (BLSTM) exhibited a slight match with the Hamming code in terms of performance. The BLSTM autoencoder was able to learn the proposed joint-coding scheme by leveraging its non-linearity property. This is in direct contrast with the Hamming code which relies on a linear transformation.

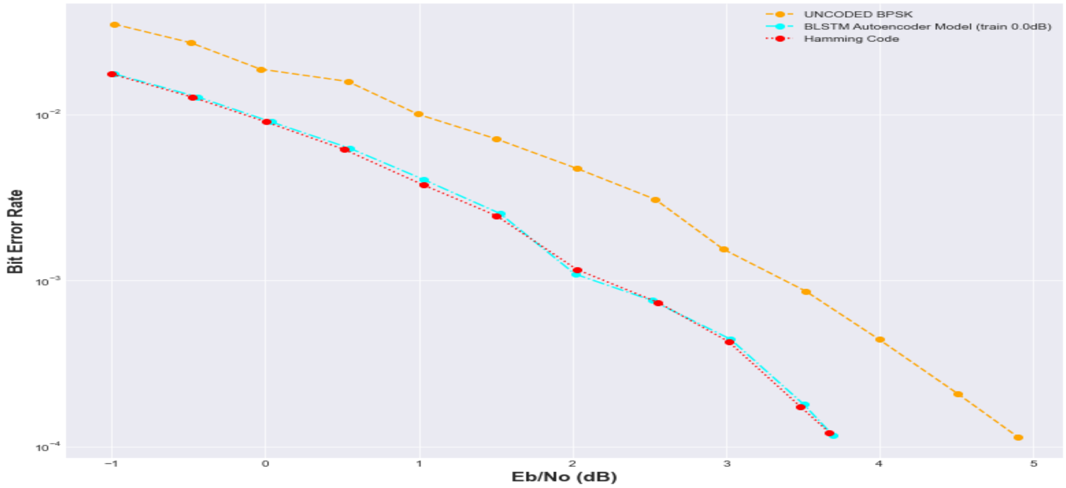


Figure 20: BER Performance comparison with Hamming code and Uncoded BPSK

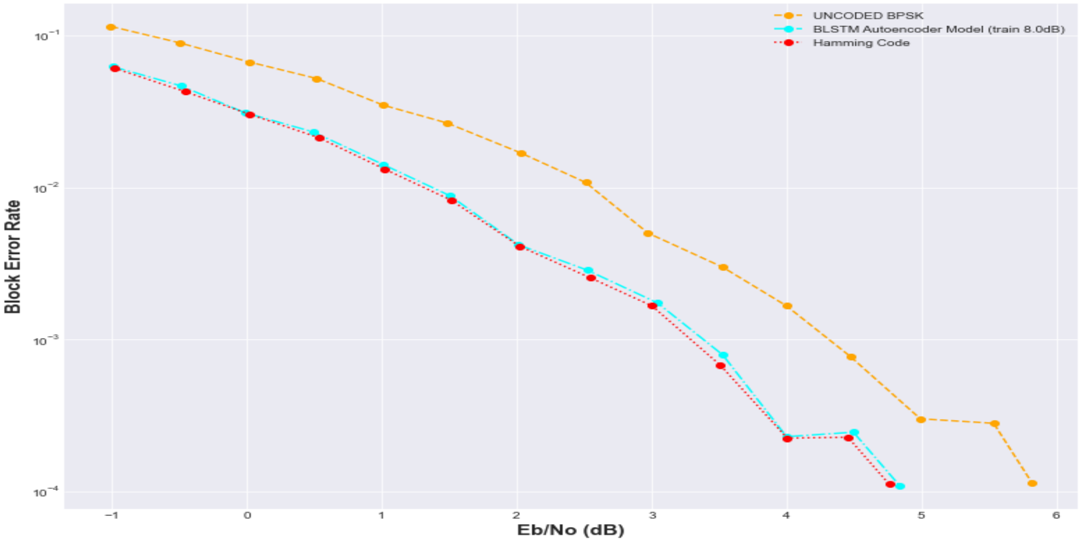


Figure 21: BLER performance comparison with Hamming code and Uncoded BPSK

From figure 20, it could be seen that autoencoder BER performance demonstrated better performance than the uncoded BPSK over the full Eb/No deployed. Hamming code implementation was closely matched by the autoencoder BER performance. It can also be observed from figures 20 and 21 that, the BLSTM configuration have almost a matched BER performance across the full E_b/N_0 range. The performance is very phenomenal owing to the fact that Hamming code is just a channel coding technique. It can also be observed from figure 20 that as the SNR improves, the BLSTM gets closer to matching the Hamming code in terms of BER performance.

4.1.2 Comparison of Bit Error Rate Performance

Figures 22-24, show the BER performance of BPSK coding schemes and the implementation of convolution code at 1/2 and 1/3 code rate and the uncoded systems respectively. When comparing the results with the results obtained from simulations of deep learning JSCC algorithm for various models, it can be observed that the convolutional codes have a similar BER performance to the deep learning JSCC based systems implemented in the study. However, as can be deduced from figure 20, the deep learning JSCC algorithm for the optimized BLSTM model was slightly better than the convolution code for R=1/3 in figure 22 and also with an improved performance over the convolution code for R=1/2 in figure 23. Under adverse channel conditions (SNR = 0), the BLSTM has a BER far lower than the convolutional codes for $R = 1/3$, $R = 1/2$ and the uncoded system. This shows that the autoencoder outperforms the the convolution codes.

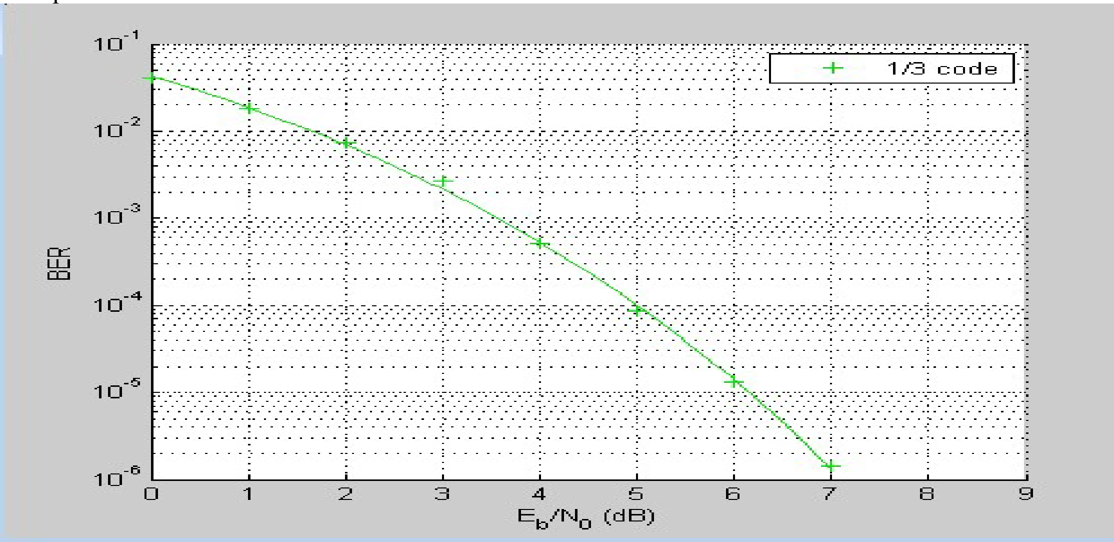


Figure 22: Result showing the Convolution Code for 1/3

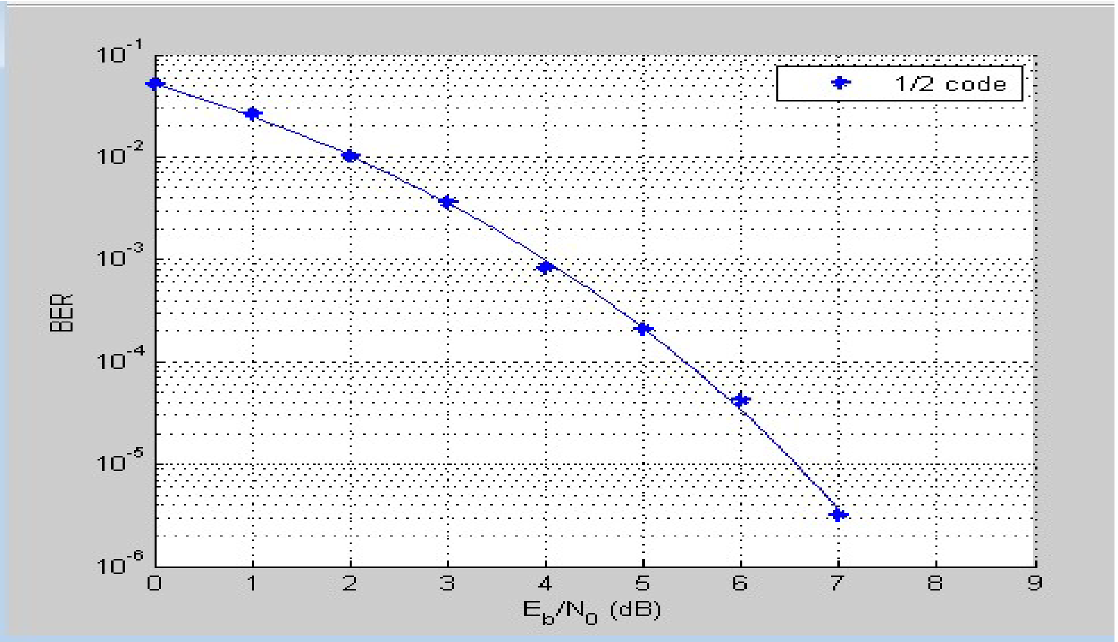


Figure 23: Result showing the Convolution Code for 1/2

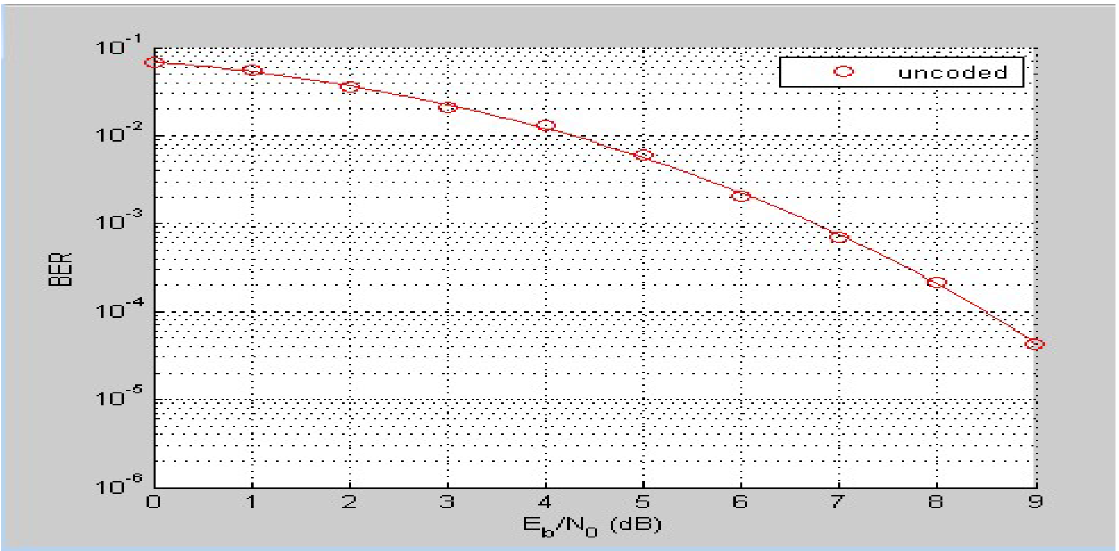


Figure 24: Result showing the uncoded system

4.2 Significance of Results

The results illustrated in figures 15-18, demonstrate the robust performance of the proposed deep learning JSCC algorithm. These results showed that the proposed algorithm has an acceptable BER performance without the need for explicitly specified codes. The comparative analyses results in figures 20 and 21 illustrated that, the BLSTM autoencoder exhibits robustness and performs favorably when compared with the Hamming code with different values on channel SNR. Owing to the fact that the autoencoder-based system does not require input block sizes of a larger dimension to operate as the input size to the model[1], it still has the capacity to achieve an acceptable BER performance with only k bits at k = 8, which exhibits quite a small block size length in comparison with the existing systems that usually operate within the range of 100s to 1000s bits long block sizes[19]. A system that exhibits such performance would be preferable, and stand a good characteristic advantage for low latency and low throughput communication systems. In such scenario, it is believed that short message transmission is possible to achieve even at a very low error rate, with minimal computational and processing complexities and delay response compare with the existing technique(s)[19]. This could also reduce the transmitter or antenna cost, with improved data rates for the same transmitter power and antenna size.

5. Conclusion

The study was primarily aimed at providing an improved channel performance approach for wireless communication network. The study sought to achieve this aim by focusing on improving the BER performance, reducing latency and the processing complexity in Joint Source Channel Coding systems. The study implemented a deep learning algorithm to enhance on the limiting performance of the conventional systems.

The Deep learning autoencoder system models were applied as an equivalent to existing models. The hamming and convolution codes in addition to the uncoded system were carefully analyzed with deep learning autoencoder models. The deep learning autoencoder model demonstrated a performance that compared favorably with the hamming code and better than the convolution codes, and uncoded systems. The results obtained showed that the autoencoder model exhibits better and or approximately equal BER performance even when hamming code (soft decoding) was utilized.

Further studies could be focused towards exploiting more advanced deep learning architectures and models in the autoencoder that could further enhance the compression performance of data with minimal BER.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors extend their appreciation to Federal University of Technology, Owerri for providing the enabling environment for learning and to all the colleagues that contributed in various ways towards the success of the research study.

References

[1] R.N.S. Rajapaksha, “Master’s Thesis: Potential Deep Learning Approaches for the Physical,” (July), 1–59, 2019.

[2] E. Bourtsoulatze, D. Burth Kurka, D. Gunduz, “Deep Joint Source-Channel Coding for Wireless Image Transmission,” IEEE Transactions on Cognitive Communications and Networking, **5**(3), 567–579, 2019, doi:10.1109/tccn.2019.2919300.

[3] I. Goodfellow, Y. Bengio, A. Courville, Deep learning, MIT press, 2016.

[4] I.I. Akpabio, “Joint Source-Channel Coding Using Machine Learning,” (May), 2019.

[5] R. Atienza, Advanced Deep Learning with Keras: Apply deep learning techniques, autoencoders, GANs, variational autoencoders, deep reinforcement learning, policy gradients, and more, 2018.

[6] T. O’Shea, J. Hoydis, “An Introduction to Deep Learning for the Physical Layer,” IEEE Transactions on Cognitive Communications and Networking, **3**(4), 563–575, 2017, doi:10.1109/TCCN.2017.2758370.

[7] E. Nachmani, Y. Be’Ery, D. Burshtein, “Learning to decode linear codes using deep learning,” 54th Annual Allerton Conference on Communication, Control, and Computing, Allerton 2016, 341–346, 2017, doi:10.1109/ALLERTON.2016.7852251.

[8] E. Nachmani, E. Marciano, D. Burshtein, Y. Be’ery, “RNN Decoding of Linear Block Codes,” 2017.

[9] N. Samuel, T. Diskin, A. Wiesel, “Deep MIMO detection,” in 2017 IEEE

18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), 1–5, 2017, doi:10.1109/SPAWC.2017.8227772.

[10] N. Farsad, A. Goldsmith, “Detection Algorithms for Communication Systems Using Deep Learning,” 2017.

[11] H. Ye, G.Y. Li, B.H. Juang, “Power of Deep Learning for Channel Estimation and Signal Detection in OFDM Systems,” IEEE Wireless Communications Letters, **7**(1), 114–117, 2018, doi:10.1109/LWC.2017.2757490.

[12] N. Farsad, M. Rao, A. Goldsmith, “Deep Learning for Joint Source-Channel Coding of Text,” ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, **2018-April**, 2326–2330, 2018, doi:10.1109/ICASSP.2018.8461983.

[13] Y.M. Saidutta, A. Abdi, F. Fekri, “M to 1 Joint Source-Channel Coding of Gaussian Sources via Dichotomy of the Input Space Based on Deep Learning,” in 2019 Data Compression Conference (DCC), 488–497, 2019, doi:10.1109/DCC.2019.00057.

[14] L. Rongwei, W. Lenan, G. Dongliang, “JOINT SOURCE CHANNEL CODING MODULATION BASED ON BP,” 156–159, 2003.

[15] G. Toderici, S.M. O’Malley, S.J. Hwang, D. Vincent, D. Minnen, S. Baluja, M. Covell, R. Sukthankar, “Variable rate image compression with recurrent neural networks,” 4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings, 1–12, 2016.

[16] G. Toderici, D. Vincent, N. Johnston, S.J. Hwang, D. Minnen, J. Shor, M. Covell, “Full resolution image compression with recurrent neural networks,” Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, **2017-Janua**, 5435–5443, 2017, doi:10.1109/CVPR.2017.577.

[17] D.P. Kingma, M. Welling, “Auto-encoding variational bayes,” 2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings, (Ml), 1–14, 2014.

[18] Y.M. Saidutta, A. Abdi, F. Fekri, “Joint Source-Channel Coding of Gaussian sources over AWGN channels via Manifold Variational Autoencoders,” 2019 57th Annual Allerton Conference on Communication, Control, and Computing, Allerton 2019, 514–520, 2019, doi:10.1109/ALLERTON.2019.8919888.

[19] N. Rajapaksha, N. Rajatheva, M. Latva-Aho, “Low Complexity Autoencoder based End-to-End Learning of Coded Communications Systems,” IEEE Vehicular Technology Conference, **2020-May**, 2020, doi:10.1109/VTC2020-Spring48590.2020.9128456.

[20] A.D. Setiawan, T.L.R. Mengko, A.B. Suksmono, H. Gunawan, “Low-bitrate medical image compression,” Proceedings of the 12th IAPR Conference on Machine Vision Applications, MVA 2011, 544–547, 2011.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 7, No. 4, 140-146 (2022)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Scalability of Multi-Stage Nested Mach-Zehnder Interferometer Optical Switch with Phase Generating Couplers

Masayuki Kawasako, Toshio Watanabe*, Tsutomu Nagayama, Seiji Fukushima

Kagoshima University, 1-21-40 Korimoto, Kagoshima, 890-0065, Japan

ARTICLE INFO

Article history:
Received: 31 May, 2022
Accepted: 07 August, 2022
Online: 29 August, 2022

Keywords:
Optical Switch
Optical Waveguide
Mach-Zehnder Interferometer

ABSTRACT

A nested Mach-Zehnder interferometer (MZI) configuration whose phase shifters are placed in parallel is suitable for silicon-silica hybrid structure to realize a high-speed optical switch. Even when the signal wavelength deviates from an optimal wavelength, the crosstalk of the nested MZI optical switch can be suppressed by employing phase generating couplers (PGCs) in place of directional couplers. We calculate the characteristics of a 4-stage nested MZI switch with PGCs, and show that crosstalk is lower than −40 dB over a wavelength range of as wide as 200 nm from 1450 to 1650 nm in six output ports. We also examine the scalability of the multi-stage nested MZI switch, and deduce the required number of switch stages for given output port counts with low crosstalk.

1. Introduction

In optical fiber communication systems, optical switches are employed to route the optical signals passing through a node without any conversion into electrical signals. Among various types of optical switches, those based on silica waveguides have many advantages [1-5]. Wafer-level fabrication process enables large-scale integration and mass productivity. Polarization and temperature insensitive operation eliminates the need of polarization diversity optics and thermo-electric cooler. Physical and chemical stability of silica as well as the absence of moving mechanical parts offer high reliability. The silica-based optical switches with thermo-optic phase shifters have been investigated since as early as 1980s [1], and their excellent performance as well as high feasibility has been well qualified [2]. They have been deployed in reconfigurable optical add/drop multiplexing (ROADM) systems in 2000s [3], where their millisecond response time is sufficient for this application. Their versatility has been successfully demonstrated thanks to the fact that they can integrate with various kinds of optical components [4]. They have been also deployed in multi-degree ROADM systems as a multicast switch in 2010s [5].

Recently, an application of optical switches to data center network is intensively considered [6-8], where switching speed is a critical issue and microsecond response time is required. A new optical switch architecture based on wavelength routing was proposed, which is suitable for data center network [6]. A fast-optical switch employing digital micromirror device (DMD)-based wavelength selective switch was demonstrated [7], and a review on optical switches for data centers was reported [8]. To realize a waveguide-based optical switch with high speed, silicon-silica hybrid structure [9] is promising since silicon provides excellent thermal properties, although it was reported the investigation to enhance the switching speed of the silica-based optical switch by optimizing the waveguide structure and using an over-driving technique [10].

In the waveguide-based optical switches, a Mach-Zehnder interferometer (MZI) is commonly employed as a basic element [1]. An MZI consists of two directional couplers (DCs) connected by arm waveguides with phase shifters. It acts as 1 × 2 or 2 × 2 optical switches, and we can construct 1 × N or M × N optical switch by cascading them [11-13]. However, conventional cascaded MZI optical switches are not suitable for the hybrid structure because their phase shifters are placed in series. In contrast, a nested MZI configuration [14, 15], whose phase shifters are placed in parallel, is suitable for the hybrid structure.

Crosstalk of the nested MZI optical switch as well as the conventional MZI optical switch becomes larger when the input signal wavelength deviates from an optimal wavelength because both the coupling ratio of the DC and the phase given by the arm waveguides vary as wavelength. To mitigate this wavelength dependence, a previous study [16] proposed a nested MZI optical switch employing phase generating couplers (PGCs) [17] in place of DCs. PGCs create the phase to cancel the wavelength dependence of the arm waveguides. Calculation results showed

*Corresponding Author: Toshio Watanabe, wata104@eee.kagoshima-u.ac.jp

https://dx.doi.org/10.25046/aj070418

that a single-stage MZI switch with PGCs has low crosstalk less than −30 dB over a wavelength band of 200 nm ranging from 1450 to 1650 nm, whereas a conventional one with DCs has crosstalk lower than −30 dB within 60 nm wavelength range from 1520 to 1580 nm [16]. It was also found that crosstalks in a 2-stage nested MZI switch are further suppressed in two output ports chosen among four output ports. A 3-stage nested MZI optical switch has three output ports with low crosstalk among eight output ports [16]. In addition, we have reported that a 4-stage nested MZI switch has low crosstalk less than −40 dB over a wavelength range of as wide as 200 nm from 1450 to 1650 nm by choosing six ports among 16 output ports [18].

In this paper, we further examine the crosstalk characteristics of the multi-stage nested MZI optical switch and reveal how crosstalk is suppressed in selected output ports. Section 2 describes the difference between a conventional cascaded MZI switch and a nested MZI switch. Section 3 shows the simulation method for calculating crosstalk characteristics of this optical switch, including the PGC circuit configuration and their parameters used in this study. We discuss the calculation results and exhibit the scalability of the multi-stage nested MZI switch in Section 4. Finally, we conclude this paper in Section 5.

2. Conventional MZI Switch and Nested MZI Switch

Figure 1 shows a conventional cascaded MZI optical switch [11-13]. It has two output ports (E3, E5) with low crosstalk among five output ports, because crosstalks to these two ports are blocked by the MZI in through state twice, respectively. This is a common way to suppress crosstalks to unrouted output ports in the conventional cascaded MZI switch. The through and cross ports in the MZI are never symmetric in terms of their optical transmittance when the coupling ratio of DC deviates from 50 % and the phase given by the arm waveguide does from π due to their wavelength dependence. The transfer matrix of DC is expressed by

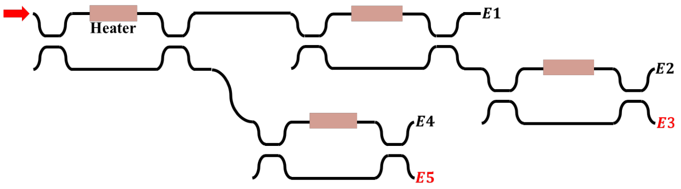


Figure 1: Conventional cascaded MZI optical switch

$$T_{\mathrm{DC}}=\begin{bmatrix}\kappa' & -j\kappa\\ -j\kappa & \kappa'\end{bmatrix}, \tag{1}$$

where κ' and κ are the amplitudes of optical field coupled to the through and cross ports, respectively, and $\kappa'^2+\kappa^2=1$ for a lossless DC. Then the optical output powers in the through and cross ports are given as

$$|E_1|^2 = 1 - 4K(1-K)\cos^2\left(\frac{\psi}{2}\right) \tag{2}$$

and

$$|E_2|^2 = 4K(1-K)\cos^2\left(\frac{\psi}{2}\right), \tag{3}$$

respectively, where $K=\kappa^2=1-\kappa'^2$ is the power coupling ratio of DC and ψ is the phase given by the arm waveguide [1]. Equations (2) and (3) indicate that crosstalk to through port becomes zero only if $K=0.5$ in cross state (ψ=0), while crosstalk to cross port vanishes regardless of K in through state (ψ=π). In general, the wavelength dependence of K is greater than that of ψ, cross port (in through state) offers lower crosstalk than through port (in cross state). In a 1 × N switch composed of cascaded MZIs shown in Figure 1, the number of output ports with low crosstalk is less than half of the total optical output counts.

The conventional cascaded MZI optical switch, however, is not suitable for the hybrid structure because its phase shifters are placed in series. This makes it difficult to fabricate the phase shifter and other waveguide regions with different material. As the output port counts and the number of MZIs with the phase shifter increase, the transmission loss becomes larger since the coupling loss between the waveguides with different materials is accumulated every time passing through the MZI. In contrast, a nested MZI switch [14-16, 18] is suitable for the hybrid structure because its phase shifters are placed in parallel. It is easy to fabricate the phase shifter and the other waveguide regions separately with different materials, and transmission loss does not increase with the number of output ports.

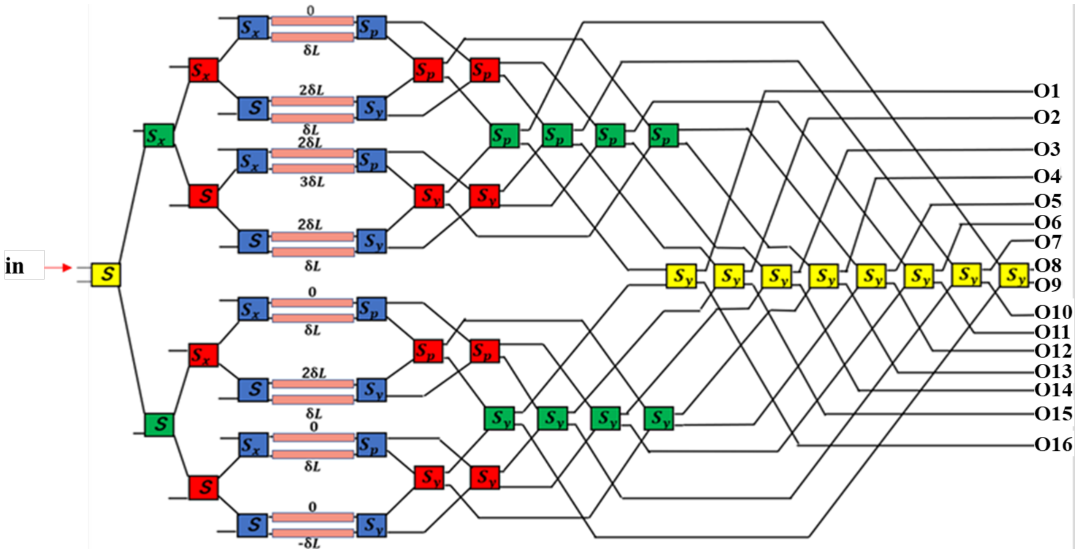


Figure 2: Circuit configuration of 4-stage nested MZI switch with PGCs

Figure 2 shows a circuit configuration of the optical switch we examined in this paper. This circuit corresponds to an MZI switch arranged in a nested architecture with four stages [18]. The PGCs indicated in different colors (blue, red, green and yellow) construct the first, second, third and fourth stages, respectively. An inherent symmetry of this circuit configuration makes it easy to increase the number of stages and output port counts [16]. However, the crosstalk characteristics of the nested MZI optical switch with a large number of stages has not been clarified yet.

3. Simulation Method

We employ PGCs instead of the DCs in all nested MZIs. Figure 3 shows the circuit configurations of the PGCs: a basic element as well as those having the line symmetries with respect to X- or Y-axes and point symmetry. The PGCs in each MZI stages are arranged in mirror symmetry as shown in Figure 2. In order to give an optical phase shift of π, appropriate optical path length differences are set between the arm waveguides connecting the PGCs, as also shown in Figure 2. Table 1 shows the MZI state of each stage for given optical output ports. In Table 1, C denotes that the corresponding MZI stage is in cross state while T does that it is in through state. Table 2 shows the phase settings to route each optical output port, where ϕ_i $(i = 1, 2, \ldots, 16)$ is that of i-th phase shifter, and $\phi_i = 0$ and 1 indicate the phase shift of 0 and π, respectively.

We calculated the optical transmittance of the nested MZI to each output port by the same method as the previous study [16]. The circuit parameters used in this study was also the same as those given previously. The power coupling ratio of the PGCs was designed to be 50 % (3 dB) at 1550 nm, the center wavelength of

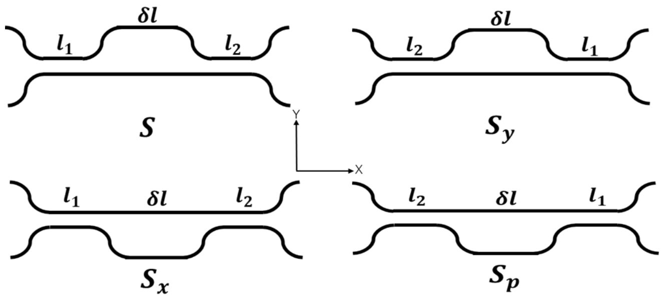


Figure 3: Basic PGC element, line-symmetric (with X- or Y-axes) and point-symmetric PGCs

Table 1: MZI states

Stage Output	1st stage	2nd stage	3rd stage	4th stage
O1	T	T	T	T
O2	C	T	T	T
O3	C	C	T	T
O4	T	C	T	T
O5	T	C	C	T
O6	C	C	C	T
O7	C	T	C	T
O8	T	T	C	T
O9	T	T	C	C
O10	C	T	C	C
O11	C	C	C	C
O12	T	C	C	C
O13	T	C	T	C
O14	C	C	T	C
O15	C	T	T	C
O16	T	T	T	C

MZI state: C = cross, T = through

Table 2: Phase settings

	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	ϕ_7	ϕ_8	ϕ_9	ϕ_{10}	ϕ_{11}	ϕ_{12}	ϕ_{13}	ϕ_{14}	ϕ_{15}	ϕ_{16}
O1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O2	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0
O3	0	1	0	1	1	0	1	0	0	1	0	1	1	0	1	0
O4	0	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0
O5	0	0	1	1	0	0	1	1	1	1	0	0	1	1	0	0
O6	0	1	0	1	0	1	0	1	1	0	1	0	1	0	1	0
O7	0	1	1	0	1	0	0	1	1	0	0	1	0	1	1	0
O8	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
O9	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
O10	0	1	1	0	1	0	0	1	0	1	1	0	1	0	0	1
O11	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
O12	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
O13	0	0	1	1	1	1	0	0	1	1	0	0	0	0	1	1
O14	0	1	0	1	1	0	1	0	1	0	1	0	0	1	0	1
O15	0	1	1	0	0	1	1	0	1	0	0	1	1	0	0	1
O16	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1

operating range. The parameters of the PGC are as follows: coupling lengths l_1 = 20 μm and l_2 = 10 μm, delay line length δl = 0.086 μm, and optical path length difference δL = 0.4 μm.

The optical output field of a single-stage MZI switch with PGCs (configuration S and S_y) is given as

$$\begin{pmatrix} E_1 \\ E_2 \end{pmatrix} = T_{\mathrm{latter}} T_{\mathrm{phase}} T_{\mathrm{former}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad (4)$$

where E_1 and E_2 indicate output fields in through and cross ports, respectively. In equation (4),

$$T_{\mathrm{latter}} = \begin{bmatrix} H & F \\ -F^* & H^* \end{bmatrix}, \qquad (5)$$

$$T_{\mathrm{phase}} = \begin{bmatrix} \exp(\text{-j}\psi(\lambda)/2) & 0 \\ 0 & \exp(\mathrm{j}\psi(\lambda)/2) \end{bmatrix}, \qquad (6)$$

$$T_{\mathrm{former}} = \begin{bmatrix} H & -F^* \\ F & H^* \end{bmatrix} \qquad (7)$$

indicate transfer matrix of S_y, the arm waveguides, and S, respectively [16]. Here, H and F indicate transfer functions of through and cross ports in the PGC, respectively. H^* is a complex conjugate of the transfer function H, and $\psi(\lambda)$ is a phase difference between the arm waveguides connecting the PGCs. We simulated the optical output power of a multi-stage MZI switch with PGCs by multiplying these transfer matrices in appropriate order.

4. Results and Discussion

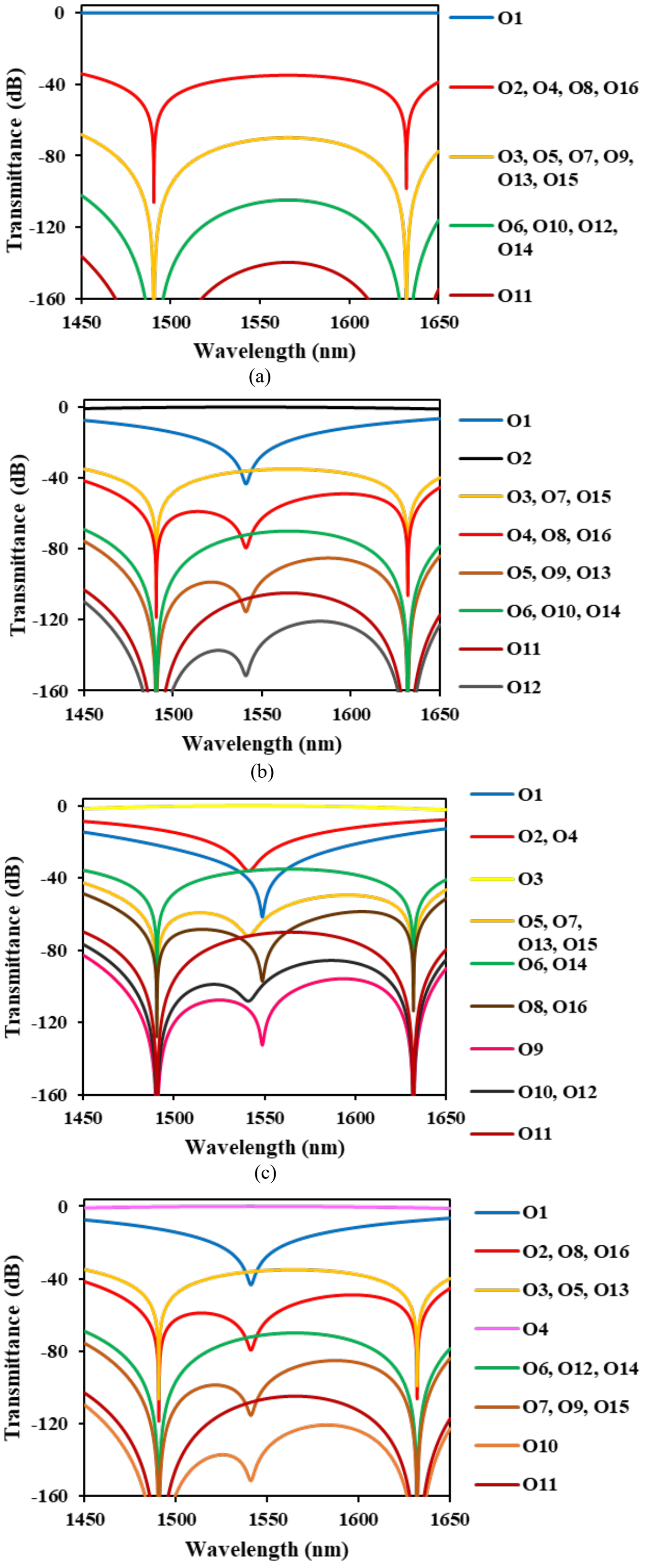
4.1. Results

Figures 4(a)–4(p) show the calculated optical transmittance of the switch we propose when the optical signal is routed to output ports O1–O16, respectively. Here, the transmittance is defined as the ratio of the optical output power to the optical input power. Wavelength dependence was computed in increments of 0.1 nm.

In Figure 4(a) where the optical input signal is routed to O1, crosstalks to unrouted ports of O2, O4, O8 and O16 are the same and less than −30 dB in 200 nm wavelength range from 1450 to 1650 nm. Crosstalks to O3, O5, O7, O9, O13 and O15 are lower than −60 dB, while those to O6, O10, O12 and O14 are below −100 dB. In particular, O11 has quite low crosstalk below −130 dB. In all O2–O16, crosstalks are dropped at two wavelengths of 1490 and 1630 nm, where the PGCs create the phase that exactly cancels the wavelength dependence of the arm waveguides.

However, in Figure 4(b) where the optical signal is directed to O2, crosstalk to O1 increases and becomes higher than −20 dB as the signal wavelength is away from a center wavelength. This is because the power coupling ratio of the PGC deviates from 50 % (3 dB) due to its wavelength dependence. Crosstalks to O3, O7 and O15 are identical and less than −30 dB within the 200 nm wavelength range around 1550 nm, while those to O4, O8 and O16 are lower than −40 dB. Other ports have much lower crosstalks below −60 dB.

When the optical path is set toward O4, O8 and O16, respectively, O1 has high crosstalk as same as the output is set at O2. Crosstalk to O1 is a little bit low but higher than −20 dB when the optical input signal is directed to O3, O5, O7, O9, O13 and O15, respectively. When the optical signal is routed to O6, O10, O12 and O14, respectively, crosstalk to O1 is identical and also exceeds −20 dB. In the case that the routed port is O11, it remains higher than −20 dB within the wavelength range of 1550 ± 100 nm. These results indicate that we should avoid to use O1 as the output port of the nested MZI optical switch.

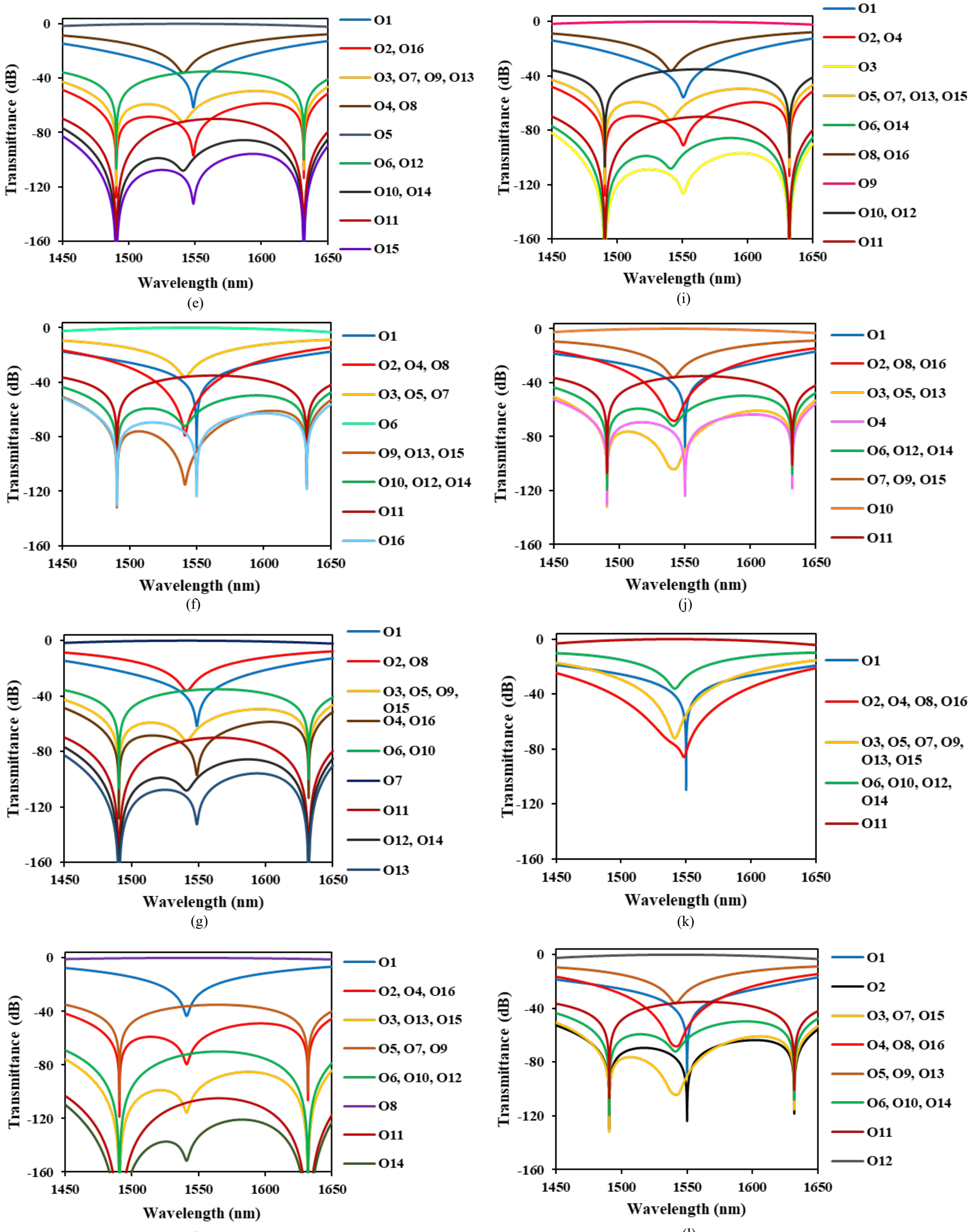


(a)

(b)

(c)

(d)


Transmittance (dB)
Wavelength (nm)
0
-40
-80
-120
-160
1450
1500
1550
1600
1650
O1
O2, O16
O3, O7, O9, O13
O4, O8
O5
O6, O12
O10, O14
O11
O15
O1
O2, O4
O3
O5, O7, O13, O15
O6, O14
O8, O16
O9
O10, O12
O11
O1
O2, O4, O8
O3, O5, O7
O6
O9, O13, O15
O10, O12, O14
O11
O16
O1
O2, O8, O16
O3, O5, O13
O4
O6, O12, O14
O7, O9, O15
O10
O11
O1
O2, O8
O3, O5, O9, O15
O4, O16
O6, O10
O7
O11
O12, O14
O13
O1
O2, O4, O8, O16
O3, O5, O7, O9, O13, O15
O6, O10, O12, O14
O11
O1
O2, O4, O16
O3, O13, O15
O5, O7, O9
O6, O10, O12
O8
O11
O14
O1
O2
O3, O7, O15
O4, O8, O16
O5, O9, O13
O6, O10, O14
O11
O12

(e)

(i)

(f)

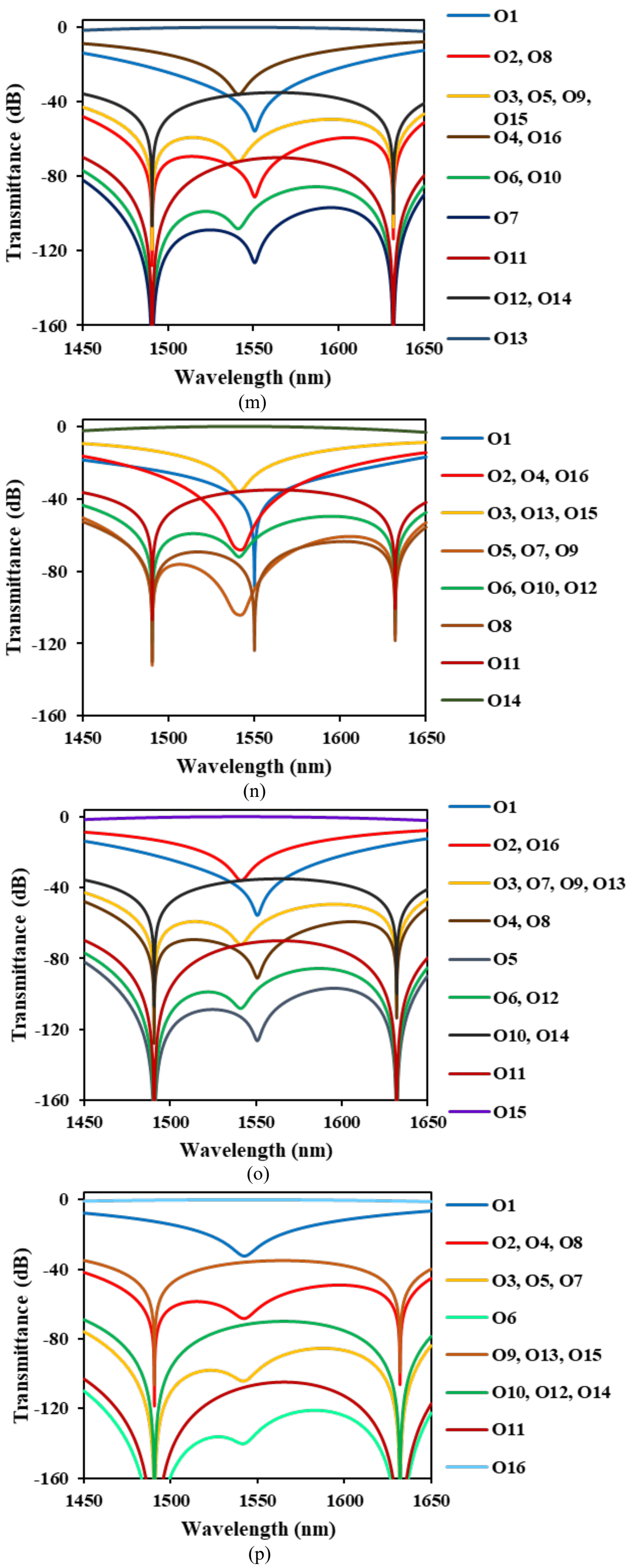
(j)

(g)

(k)

(h)

(l)



(m)

(n)

(o)

(p)

Figure 4: Transmittance of 4-stage nested MZI switch with PGCs when the optical path is set toward O1−O16 (a) O1 (b) O2 (c) O3 (d) O4 (e) O5 (f) O6 (g) O7 (h) O8 (i) O9 (j) O10 (k) O11 (l) O12 (m) O13 (n) O14 (o) O15 (p) O16

We examined the crosstalk characteristics shown in Figures 4(a)−4(p), and found that crosstalks to unrouted ports are suppressed to −40 dB or less over the wavelength band of 200 nm ranging from 1450 to 1650 nm when we choose only O3, O5, O7, O9, O13 and O15 as the output port of the nested MZI optical switch. In Figure 4(c) where the optical input signal is routed to O3, crosstalks to O5, O7, O13 and O15 are identical and less than −40 dB. O9 has much lower crosstalk below −80 dB. When we avoid to use other ports as output, crosstalks to unrouted ports (O5, O7, O9, O13 and O15) never exceed −40 dB. In Figures 4(e), 4(g), 4(i), 4(m) and 4(o), crosstalks remain lower than −40 dB when the optical signal is routed to O5, O7, O9, O13 and O15, respectively.

4.2. Discussion

The crosstalk characteristics of 16 ports shown in Figures 4(a)−4(p) is classified into 5 groups. One is O1, second is O2, O4, O8 and O16, third is O3, O5, O7, O9, O13 and O15, forth is O6, O10, O12 and O14, and the last is O11. Each group has the same number of the MZI states (cross or through) as shown in Table 1. This means that the port settings in the same group have two or more different MZI states among four stages each other. For example, the port settings in the third group (O3, O5, O7, O9, O13 and O15) have two cross states and two through states among four MZI stages. O3 has the MZI states of CCTT, where C or T denotes each MZI state of four stages as shown in Table 1. This state differs in two stages to those of O5 (TCCT), O7 (CTCT), O13 (TCTC) and O15 (CTTC), while in all four stages to O9 (TTCC). Thus, when the optical path is set toward O3, crosstalks to O5, O7, O13 and O15 are blocked in two MZI stages where one is in cross state and the other is in through state, respectively, and that to O9 in four MZI stages where two are in cross state and the other two are in through state. Crosstalks to unrouted ports are blocked similarly when the optical signal is routed to O5, O7, O9, O13 and O15, respectively. This is why the lower crosstalk is given when we choose only O3, O5, O7, O9, O13 and O15 as the output ports of the nested MZI optical switch.

Following the analysis above, we can deduce how many output ports in multi-stage nested MZI switch has low crosstalk. We should choose the port settings which have the same number of the MZI states (cross or through) and equalize the cross and through states. Then the number of output ports N with low crosstalk in an n-stage switch ($n \geq 2$) is given as

$$N = {}_{n}C_{[n/2]} = \binom{n}{[n/2]}, \tag{8}$$

where $[x]$ is a floor function. In the previous study [16], it has been shown that a 2-stage nested MZI switch has two ports with low crosstalk among four output ports. A 3-stage nested MZI optical switch gives three ports with low crosstalk in eight output ports. This study shows that 4-stage nested MZI switch has six ports with low crosstalk in 16 output ports. Equation (8) agrees with these results, and indicates that 5-stage nested MZI switch must have ten ports with low crosstalk in 32 output ports. Figure 5 shows the required number of stages n for given output port counts N with low crosstalk. The ratio of N to 2^n is less than 50 % and decreases as n increases. However, the conventional MZI optical switch commonly employ two cascaded MZIs per output port to suppress crosstalk sufficiently [16-18]. In that case the ratio $N/2^n$ is around 50 % for the conventional cascaded MZI

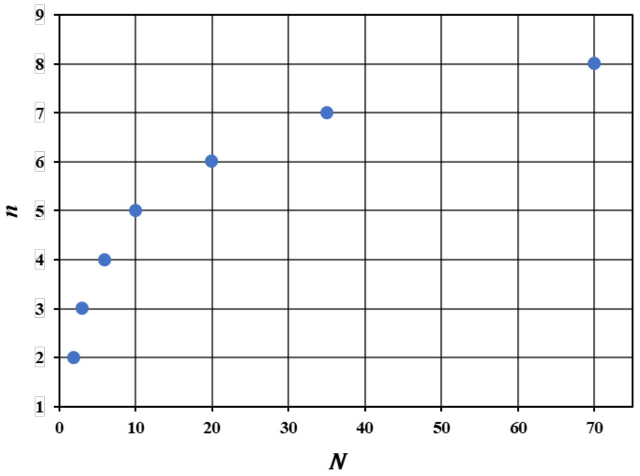


Figure 5: Required number of stages n for given output port counts N with low crosstalk

switch. Although the nested MZI switch has somewhat smaller ratio of the output ports with low crosstalk, it has advantages that its phase shifters are placed in parallel and gives a shorter circuit length than the cascaded MZI optical switch.

Recently, a new concept was demonstrated that effectively reduces the wavelength dependence of an MZI switch [19]. This technique employs no PGCs but arm waveguides with different waveguide widths in each cascaded MZI. It is a future work whether this technique is possible to apply to the nested MZI switch with arrayed arm waveguides.

5. Conclusion

We examined the crosstalk characteristics of a 4-stage nested MZI optical switch with PGCs. It was confirmed that crosstalk is lower than −40 dB over a wavelength band of as wide as 200 nm ranging from 1450 to 1650 nm in six output ports of the 4-stage switch. We showed that crosstalks to these six ports are blocked in two MZI stages where one is in cross state and the other is in through state, respectively. Based on the results, we deduced the required number of stages for given output port counts with low crosstalk in multi-stage nested MZI switch. It is possible to construct 1 × 10, 1 × 20, 1 × 35 optical switches with 5, 6, 7-stages, respectively, and output port count reaches 70 in 8-stage nested MZI switch.

Conflict of Interest

We declare no conflict of interest.

Acknowledgment

This work was partly supported by JSPS KAKENHI Grant Number JP18K04286.

References

[1] N. Takato, K. Jinguji, M. Yasu, H. Toba, and M. Kawachi, “Silica-based single-mode waveguides on silicon and their application to guided-wave optical interferometers,” Journal of Lightwave Technology, **6**(6), 1003-1010, 1988, doi: 10.1109/50.4091.

[2] S. Sohma, S. Mino, T. Watanabe, M. Ishii, T. Shibata, and H. Takahashi, “Solid-state optical switches using planar lightwave circuit and IC-on-PLC technologies,” Proceedings of SPIE, **5625**, 767-775, 2004, doi: 10.1117/12.579684.

[3] H. Takahashi, T. Watanabe, T. Goh, S. Sohma, and T. Takahashi, “PLC optical switch that enhances the optical communication network,” NTT Technical Review, **3**(7), 17-21, 2005.

[4] C. R. Doerr and K. Okamoto, “Advances in silica planar lightwave circuits,” Journal of Lightwave Technology, **24**(12), 4763-4789, 2006, doi: 10.1109/JLT.2006.885255.

[5] T. Watanabe, K. Suzuki, T. Goh, K. Hattori, A. Mori, T. Takahashi, T. Sakamoto, K. Morita, S. Sohma, and S. Kamei, “Compact PLC-based transponder aggregator for colorless and directionless ROADM,” Technical Digest of 2011 Optical Fiber Communication Conference/National Fiber Optic Engineers Conference (OFC/NFOEC 2011), paper OTuD3, 2011, doi: 10.1364/OFC.2011.OTuD3.

[6] K. Sato, H. Hasegawa, T. Niwa, and T. Watanabe, “A large-scale wavelength routing optical switch for data center networks,” IEEE Communications Magazine, **51**(9), 46-52, 2013, doi: 10.1109/MCOM.2013.6588649.

[7] N. Farrington, A. Forencich, G. Porter, P.-C. Sun, J. E. Ford, Y. Fainman, G. C. Papen, and A. Vahdat, “A multiport microsecond optical circuit switch for data center networking,” IEEE Photonics Technology Letters, **25**(16), 1589-1592, 2013, doi: 10.1109/LPT.2013.2270462.

[8] Q. Cheng, M. Bahadori, M. Glick, S. Rumley, and K. Bergman, “Recent advances in optical technologies for data centers: a review,” Optica, **5**(11), 1354-1370, 2018, doi: 10.1364/OPTICA.5.001354.

[9] S. Katayose, Y. Hashizume, and M. Itoh, “Fabrication and demonstration of 1 × 8 silicon-silica multi-chip switch based on optical phased array,” Japanese Journal of Applied Physics, **55**, 08RB01, 2016, doi: 10.7567/JJAP.55.08RB01.

[10] O. Moriwaki and K. Suzuki, “Fast switching of 84 μs for silica-based PLC switch,” Technical Digest of 2020 Optical Fiber Communication Conference (OFC 2020), paper Th3B.5, 2020, doi: 10.1364/OFC.2020.Th3B.5.

[11] K. Hattori, M. Fukui, M. Jinno, M. Oguma, and K. Oguchi, “PLC-based optical add/drop switch with automatic level control,” Journal of Lightwave Technology, **17**(12), 2562-2571, 1999, doi: 10.1109/50.809678.

[12] H. Takahashi, T. Goh, T. Shibata, M. Okuno, Y. Hibino, and T. Watanabe, “High performance 8-arrayed 1 × 8 optical switch based on planar lightwave circuit for photonic networks,” Proceedings of 28th European Conference on Optical Communication (ECOC 2002), paper 4.2.6, 2002.

[13] T. Watanabe, Y. Hashizume, and H. Takahashi, “Double-branched 1×29 silica-based PLC switch with low loss and low power consumption,” Technical Digest of 17th Microoptics Conference (MOC 2011), paper J-2, 2011.

[14] K. Suzuki, T. Mizuno, M. Oguma, T. Shibata, H. Takahashi, Y. Hibino, and A. Himeno, “Low loss fully reconfigurable wavelength-selective optical 1 × N switch based on transversal filter configuration using silica-based planar lightwave circuit,” IEEE Photonics Technology Letters, **16**(6), 1480-1482, 2004, doi: 10.1109/LPT.2004.827419.

[15] T. Watanabe, K. Tasaki, T. Nagayama, and S. Fukushima, “Nested Mach-Zehnder interferometer optical switch with low crosstalk,” Technical Digest of 24th Microoptics Conference (MOC 2019), paper P-69, 2019, doi: 10.23919/MOC46630.2019.

[16] K. Tasaki, M. Tokumaru, T. Watanabe, T. Nagayama, and S. Fukushima, “Nested Mach-Zehnder interferometer optical switch with phase generating couplers,” Japanese Journal of Applied Physics, **59**, SOOB04, 2020, doi: 10.35848/1347-4065/ab8f09.

[17] T. Mizuno, H. Takahashi, T. Kitoh, M. Oguma, T. Kominato, and T. Shibata, “Mach–Zehnder interferometer switch with a high extinction ratio over a wide wavelength range,” Optics Letters, 3**0**(3), 251-253, 2005, doi: 10.1364/OL.30.000251.

[18] M. Kawasako, T. Watanabe, T. Nagayama, and S. Fukushima, “4-stage Mach-Zehnder interferometer optical switch with phase generating couplers,” Technical Digest of 26th Microoptics Conference (MOC 2021), paper P-40, 2021, doi: 10.23919/MOC52031.2021.

[19] T. Goh, K. Yamaguchi, and A. Yanagihara, “Multiband optical switch technology,” Technical Digest of 2022 Optical Fiber Communication Conference (OFC 2022), paper W4B.1, 2022, doi: 10.1364/OFC.2022.W4B.1.

Advances in Science, Technology and Engineering Systems Journal Vol. 7, No. 4, 147-154 (2022)

www.astesj.com

ASTESJ

ISSN: 2415-6698

Analysis of Different Supervised Machine Learning Methods for Accelerometer-Based Alcohol Consumption Detection from Physical Activity

Deeptaanshu Kumar*[,1], Ajmal Thanikkal[1], Prithvi Krishnamurthy[1], Xinlei Chen[1], Pei Zhang[2]

[1]*Electrical & Computer Engineering, Carnegie Mellon University, Pittsburgh, 15213, USA*

[2]*Electrical & Computer Engineering, University of Michigan, Ann Arbor, 48109, USA*

ARTICLE INFO

Article history:
Received: 31 May, 2022
Accepted: 16 August, 2022
Online: 29 August, 2022

Keywords:
Artificial Intelligence
Biomedical Engineering
Data Mining
Information Technology
Machine Learning
Systems Engineering
Mobile Computing
Embedded Systems

ABSTRACT

This paper builds on the realization that since mobile devices have become a common tool for researchers to collect, process, and analyze large quantities of data, we are now entering a generation where the creation of solutions to difficult real-world problems will mostly come in the form of mobile device apps. One such relevant real-life problem is to accurately and cheaply detect the over-consumption of alcohol, since it can lead to many problems including fatalities. Today, there are several expensive and/or tedious alternative procedures in the market that are used to test subjects' Blood Alcohol Content (BAC). This paper explores a cheaper and more effective alternative to address this problem by classifying if subjects have consumed too much alcohol by using accelerometer data from the subjects' mobile devices while they perform physical activity. In order to create the most accurate classification system, we conduct experiments with five different supervised machine learning methods and use them on two features derived from accelerometer data of two different male subjects. We then share our experiment results that support why "Decision Tree Learning" is the supervised machine learning method that is best suited for our mobile device sobriety classification system.

1. Introduction

This paper is an extension of work originally presented in the 2021 17th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob) [1].

1.1. Problem and Motivation

Alcohol misuse and abuse is responsible for great personal and economic harm in the United States (US) and around the world, where more than 88,000 people die from alcohol-related issues each year, which makes it the third leading preventable cause of death in the US [2]. Excessive drinking has been proven to damage the heart, liver, pancreas, and immune system [3]. In addition to the detrimental health effects, alcohol misuse has cost the US $223.5B in economic loss in 2006 alone [2].

The consumption of alcohol negatively affects individuals' brains and their central nervous systems. These effects only become worse with larger alcohol concentrations in the individuals' blood. More specifically, judgment, reaction time, balance, and psycho-motor performance start becoming compromised above a BAC of 0.02 – 0.05 [2]. These abilities are necessary to operate vehicles, so it is not surprising that almost 50% of traffic fatalities involve the (mis)use of alcohol.

Currently, there are a couple of different methods to test intoxication levels. These tests can be administered by drawing blood, monitoring breath (breathalyzer), and collecting urine/saliva/short strands of hair. All of these methods try to directly measure alcohol's presence in the individuals' bodies. On the other hand, field sobriety tests that are performed by law enforcement officials typically include physical tasks to gauge individuals' levels of impairment. Most of these physical movements involve performing tasks such as individuals walking backwards in a straight line or maintaining a steady posture while touching their noses with their arms stretched out. The advantages of such physical tests are that they are convenient and cheap. On the other hand, devices like the breathalyzer are fairly expensive, where costs typically range from $3,000-$5,000 per unit, and require frequent device calibration, comprehensive maintenance, and expensive repairs. However, physical tests are considered to

*Corresponding Author: Deeptaanshu Kumar, deeptaan@alumni.cmu.edu

https://dx.doi.org/10.25046/aj070419

be subjective since they are based on the officers' observation of the individuals, but are still cheaper than chemical tests.

As a result, we believe that mobile device-based systems can give individuals the best of both worlds by providing low-cost portable devices, which also measure individuals' sobriety levels objectively. The advantage of mobile computing is due to the fact fact that mobile devices can sense real-world data and respond to trends based on the data and/or their surrounding environments. If mobile devices can signal individuals that they are intoxicated by their physical responses, then fatalities can be mitigated across the world.

1.2. Proposed Solution

As seen in Figure 1, we propose a multi-stage detection system that makes use of mobile devices' accelerometers to capture real-time data from individuals. This data is then processed and fed into a data classifier, which can internally determine whether the individuals are intoxicated by comparing their data to an existing classification model that is based on historical data.

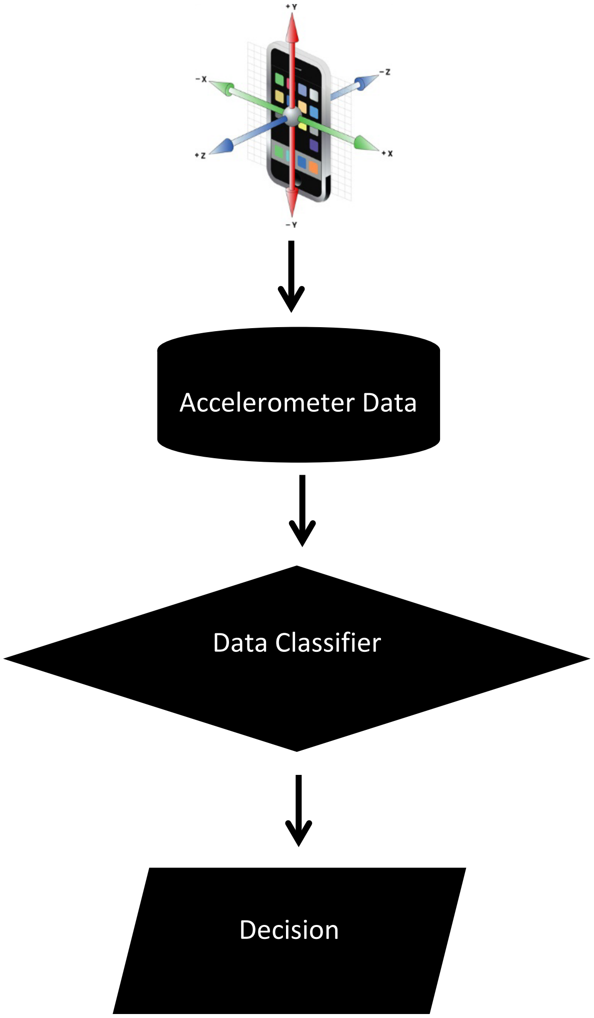


Figure 1: Block diagram of the system used for the experiments

In this paper, we take the first step towards implementing such a system by proposing robust accelerometer data features that can distinguish between sober and intoxicated individuals. We then test these features using five Supervised Machine Learning models to see their accuracy in predicting individuals' sobriety levels: Support Vector Machines (SVM), Decision Tree Learning, Boosting, K-Nearest Neighbors (KNN), and Neural Networks (NN) [4].

2. Related Works

Several researchers have focused their prior research on movement-pattern recognition by using accelerometers, especially since these days all mobile devices are embedded with highly accurate and precise accelerometers [5]. There have been many interesting mobile applications developed, which can detect the individuals' activities, such as daily exercises or crossing the street, by solely analyzing the accelerometer data.

There have also been several papers that have proposed different variations of mobile systems to detect individuals' intoxication levels:

1. Detecting abnormalities in individuals' gaits while they walk intoxicated [6–9].
2. Evaluating eye (iris) movements of individuals who are intoxicated [10,11].
3. Monitoring the steadiness of postures of intoxicated individuals [12].

These mobile systems can sense individuals' intoxication levels and log the location/time of the incidents. Although these papers show the individual differences in step variance time among individuals, the differences are relative to everyone's unique baseline, so the differences cannot be used to cleanly separate any intoxicated and sober individuals in the general population.

3. Methodology

3.1. Physiological Basis

One of the first symptoms that individuals exhibit as they become intoxicated is that they have decreased balance and motor coordination. This is because of the effect of alcohol on the brain's chemistry that it achieves by changing the neurotransmitters' levels. These neurotransmitters are entities that act as chemical messengers, which send critical human signals throughout the body. These signals include those that control thought processes, behaviors, and emotions. Many believe that out of all the neurotransmittors, alcohol specifically targets the GABA neurotransmitter [13].

3.2. Physical Activity Data Collection

With these factors in mind, we planned our experiments such that we could easily distinguish between intoxicated and sober subjects based on the subjects' abilities to balance their bodies and maintain steady postures. Our proposed mobile-based system logs subjects' accelerations along the x-, y-, and z-axes.

In order to collect the subjects' data, we created an Android app that runs on a Motorola Moto G mobile phone. This device contains a built-in API, which outputs the device's linear acceleration after negating the effects of gravity. Our Android app has a built-in button, which is used to start and stop data collection periods during our experiments.

For our experiments, we recruited two subjects, Subject 1 and Subject 2, to grip the mobile devices in their right hands, keep their right arms outstretched, and maintain those steady postures for 10 seconds. The subjects' right arms form 90-degree angles with their bodies while their left arms are kept by their sides. Additionally, their right feet are in front of the left feet in straight lines, so that their left toes are touching their right heels. During the experiments, the subjects keep their eyes to better test their balance and motor skills.

We first tested both subjects in sober states, before they consumed any alcohol. After that, both subjects consumed 3, 6, and 9 drinks of alcohol over a period of 120 minutes, while we recorded their data. To keep results consistent, we defined one drink in this paper to be 1.25 oz of 80 proof liquor i.e., vodka.

4. Results

After we finished our experiments with both subjects, we had four data points per subject, which came out to eight total data points for our initial analysis. For each of the data points, we plotted the subjects' accelerations across the x-, y- and z-axes against their times, which can be seen above in Figures 2 – 9.

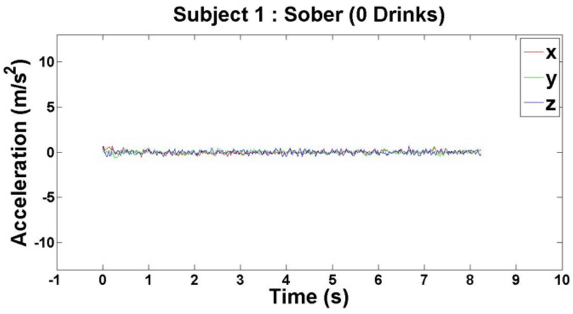
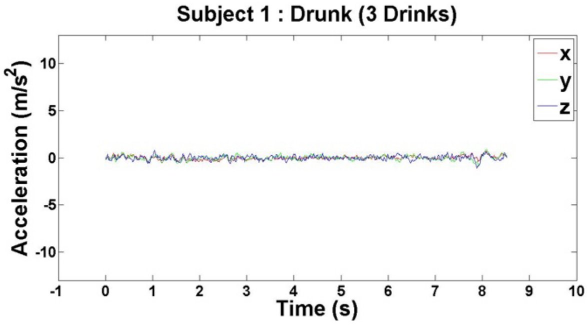


Figure 2: Accelerometer reading when Subject 1 has had 0 drinks



.

Figure 3: Accelerometer reading when Subject 1 has had 3 drinks

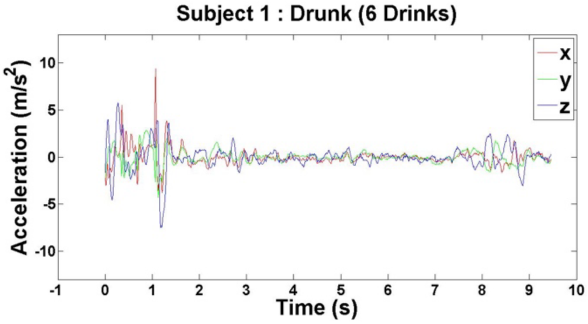


Figure 4: Accelerometer reading when Subject 1 has had 6 drinks

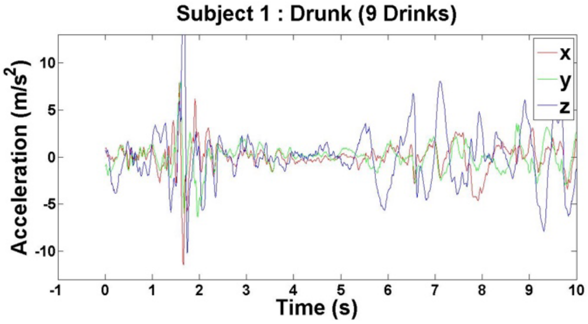


Figure 5: Accelerometer reading when Subject 1 has had 9 drinks

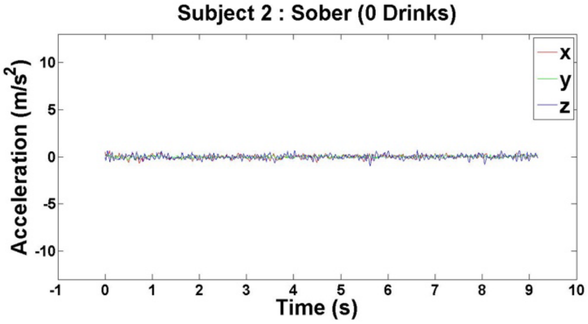


Figure 6: Accelerometer reading when Subject 2 has had 0 drinks

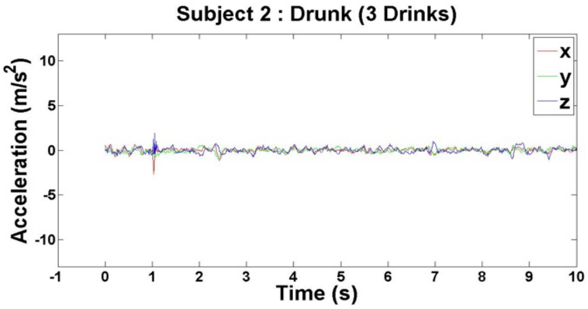


Figure 7: Accelerometer reading when Subject 2 has had 3 drinks

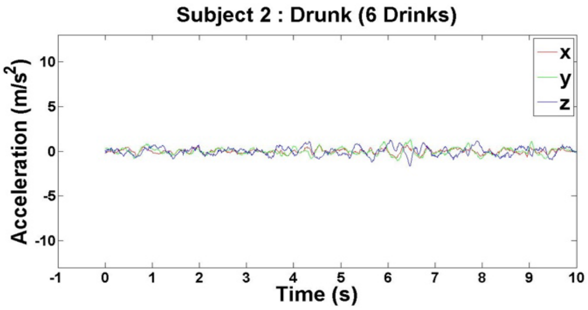


Figure 8: Accelerometer reading when Subject 2 has had 6 drinks

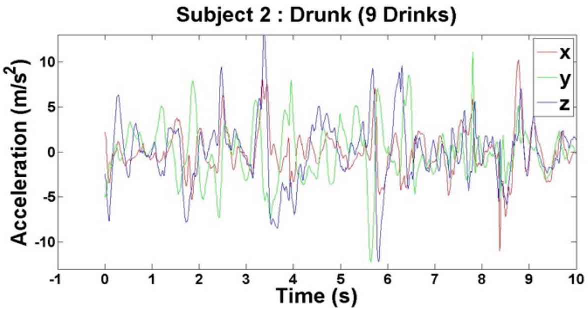


Figure 9: Accelerometer reading when Subject 2 has had 9 drinks

5. Supervised Learning Model Results

Based on the results of our experiments, we clearly see a distinction between the sober data (0 - 3 drinks) and intoxicated data (6 - 9 drinks). Additionally, there is a relationship between the number of drinks the subjects consumed and the "unsteadiness" of their corresponding data. In order to capture this change, we examined two features: variance and highest frequency.

Table 1: Tabulation of variance, highest frequency, and BAC for Subject 1

Number of Drinks	**Estimated BAC**	**Variance**	**Highest Frequency**
0	0	0.0382	0.1227
3	0.06	0.06059	0.11793
6	0.19	1.749063	2.860169
9	0.28	9.80903	1.5984

Table 2: Tabulation of variance, highest frequency, and BAC for Subject 2

Number of Drinks	**Estimated BAC**	**Variance**	**Highest Frequency**
0	0	0.0556	0.1092
3	0.06	0.06059	0.11793
6	0.19	0.214869	2.216312
9	0.28	12.0609	0.92764

The first distinguishing feature we found was the variance in the amplitude (as can be seen in Table 1). We calculated it for each axis, and then we selected the maximum variance from all the three axes. This helped make the feature more stable as the subjects held the phones in different positions, since we noticed that the variances transferred amongst the axes.

The second distinguishing feature we found was the highest frequency component of the time-series data (as can be seen in Table 2). However, the relationship was not easy to detect as in the case of the previous correlation between BAC and variance, despite the fact that the highest frequency for intoxicated data was still higher than the highest frequency for sober data.

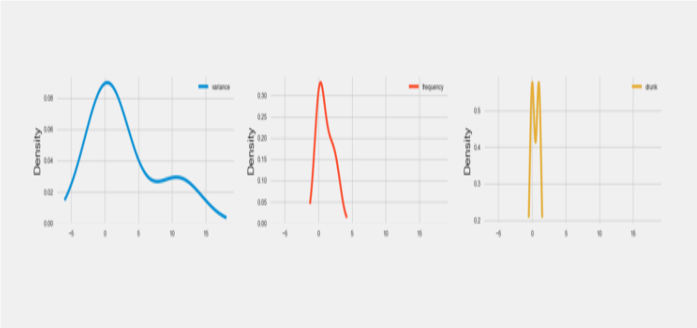


Figure 10: Breakdown of data in terms of datapoints across both subjects

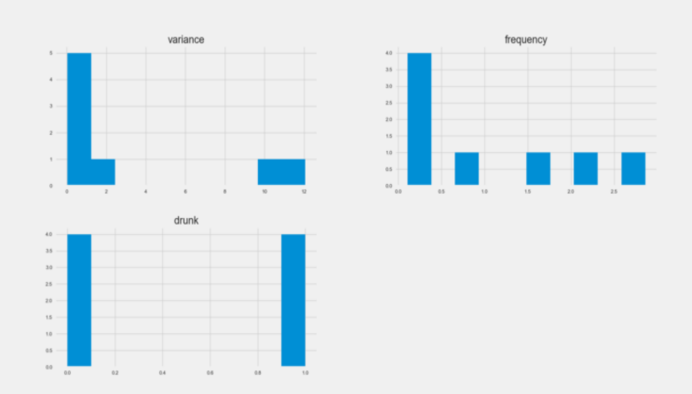


Figure 11: Density plots of the raw data for both subjects

Both the variance analysis and the highest frequency analysis were performed on Subject 1's data and Subject 2's data. Figures 10 – 11 given above indicate different levels of increasing variance and highest frequency, which we use to estimate the BAC of Subjects 1 and 2. The BAC is calculated using the number of drinks consumed and the body mass of each subject.

Based on these results, we tried the following Supervised Machine Learning methods to determine which methods were most effective at distinguishing between sober and intoxicated subjects, where we used Subject 1's data as the training dataset and Subject 2's data as the test dataset.

5.1. SVM

For the SVM implementation, we used the Python class "sklearn.svm.SVC" [14]. For our implementation, we used two different SVM models to see the different effects on the datasets (as shown in the Table 3). First, we used the original SVM model with the values of "random_state=None, kernel=poly". However, to see the effect of hyperparameter tuning the original SVM model, we updated several different parameters "random_state=0, kernel=rbf" [15,16].

Table 3: SVM model statistics

SVM Type	**Accuracy (%)**	**Execution Time (s)**
Default SVM	100	0.18
Adjusted SVM	67	0.19

5.2. Decision Tree Learning

For the decision tree implementation, we used the Python class "sklearn.tree.DecisionTreeClassifier" [17]. For our implementation, we used two different decision trees to see the different effects on the datasets (as shown in the Table 4). First, we used the regular decision tree with the default values of "gini" for the Gini impurity and "None" for the "max-depth". However, to see the effect of pruning the trees, we tested two different parameters: setting "entropy" for the information gain and setting the "max_depth" to "3" [18].

Table 4: Decision Tree Learning model statistics

Decision Tree Type	Accuracy (%)	Execution Time (s)
Default Decision Tree	67	0.71
Pruned Decision Tree	100	0.81

5.3. Boosting

For the decision tree implementation, we used the Python class “sklearn.tree. GradientBoost-ingClassifier” [19]. For our implementation, we used two different boosting classifiers to see the different effects on the datasets (as shown in the Table 5). First, we used regular gradient boosting with the default values of “n_estimators=100, learning_rate=0.1, max_depth=3” [20]. However, to see the effect of hyperparameter tuning the boosted model, we updated several different parameters “n_estimators=1000, learning_rate=1.0, max_depth=1” [16,18].

Table. 5. Gradient Boosting model statistics

Gradient Boosting Type	Accuracy (%)	Execution Time (s)
Default Gradient Boosting	67	1.13
Adjusted Gradient Boosting	67	1.12

5.4. KNN

For the KNN implementation, we used the Python class “sklearn.neighbors.KNeighborsClassifier” [21]. For our implementation, we used two different KNN models to see the different effects on the datasets (as shown in the Table 6). We created a loop to test the models on our dataset by performing hyperparameter tuning and updating the “n_neighbors” value to values 2 through 8 [16,22].

Table 6: KNN model statistics

KNN Type	Accuracy (%)	Execution Time (s)
Default KNN	75	0.54
Adjusted KNN	80	0.53

5.5. Neural Networks

For the neural network implementation, we used the Python class “keras.Sequential” [22]. For our implementation, we used two different models to see the different effects on the datasets (as shown in the Table 7). First, we used the regular keras model with two layers and 1000 epochs for the number of iterations across the data. However, to see the effect of an additional layer in my model, we tested the same keras model with a third layer and “epochs” value of “1000” for the number of cycles across the data [13,23].

Table. 7. NN model statistics

NN Type	Accuracy (%)	Epoch
Default NN	100	1000
Adjusted NN	67	1000

6. Analysis

6.1. SVM

This algorithm gave us some of the better results in both subjects’ datasets, as can be seen in Figure 12 below. This was potentially due to the fact that this algorithm is a good general-purpose classification algorithm, especially since we used the “rbf” kernel, which is known for its general and wide-spread use across many types of datasets [24].

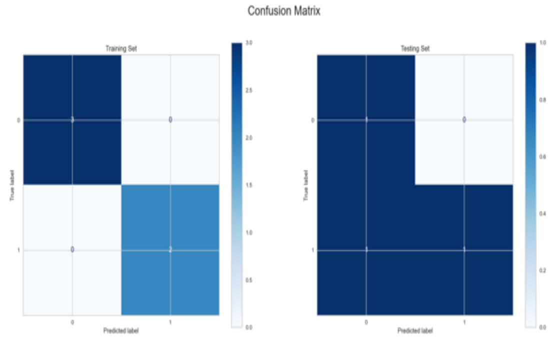


Figure 12: Confusion Matrix for SVM model

The main goal of this algorithm is to divide datasets into several classes in order to find a maximum marginal hyperplane (MMH). This can be done in the following two steps where first Support Vector Machines will generate hyperplanes iteratively that separate the classes in the best way and thereafter, they will choose the hyperplane that segregate the classes correctly [25].

We decided to perform hyperparameter tuning on our SVM and test between the rbf and polynomial kernels. We tested these because we know that the polynomial kernel is typical considered more generalized and therefore less efficient and accurate, but the rbf kernel is considered to be one of the most preferred kernel functions in SVM [24].

Based on the average runtime for our runs, we got 0.18 seconds per run as per wall clock time.

Unlike other algorithms, we didn’t do too many modifications. We only tested with different kernels, since we wanted to test a kernel that was good with generalized results to avoid overfitting/underfitting the long-tailed tailed variables. We didn’t show the results for the “poly” kernel because the “rbf” kernel performed better on the non-uniform datasets.

6.2. Decision Tree Learning

We generally got better training results with decision trees, as can be seen in Figures 12 – 13 above. This might be due to our pruning methods, or it could be because of the bi-directional tails in some of the attributes in the dataset that rendered the algorithm less effective.

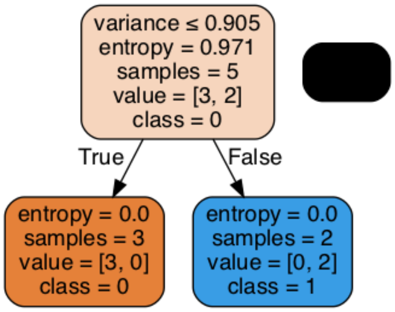


Figure 13. Node breakdown of Decision Tree model

This represents a flowchart-like tree structure, in that each internal node signifies a test on an attribute, each branch represents an outcome of the test, and each leaf node (terminal node) holds a class label [26]. This algorithm performs bests on non-linear datasets, which makes it a good choice (in theory) for our non-linear datasets

We decided to prune our trees using "criterion=entropy, max_depth=3". We chose entropy because there are features that have "uncertainty" as they the values are clustered very close to each other [23]. Additionally, we did not want to overfit the model based on the training data, so we limited the max_depth to 3 [23].

Based on the average runtime for our runs, we got 1.12 seconds per run as per wall clock time.

As we mentioned above, we decided to prune our trees using "criterion=entropy, max_depth=3". The results for the classifier for the default values were due to the fact that the model overfit on the training data and didn't give as good performance on the testing data.

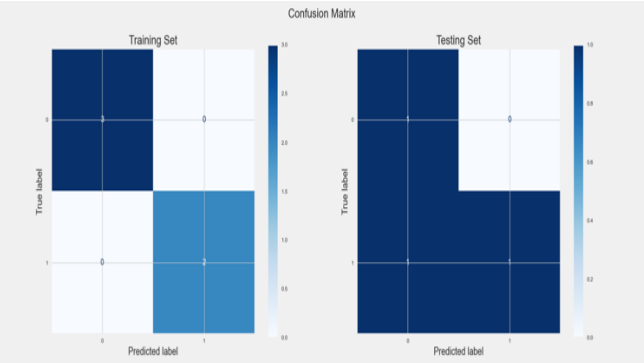


Figure 14: Confusion Matrix for Decision Tree model

6.3. Boosting

This algorithm provided interesting results because for both datasets because the larger the training data size, the better the test predictions, as can be seen in Figures 14 – 15 below. This could be due to the fact that, similar to our neural network algorithm implementation, we used a large number of iterations.

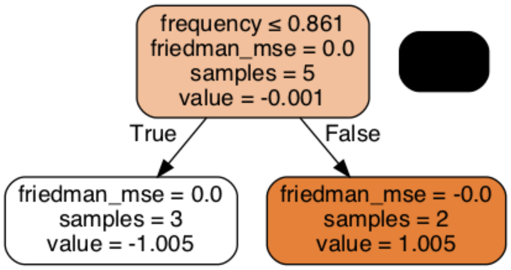


Figure 15: Breakdown for Boosting model

Since we were doing a Binary Classification problem, we used Gradient Boosting, which builds an additive model in a forward stage-wise fashion. This allows it to optimize arbitrary differentiable loss functions, where in each stage n_classes, regression trees are fit on the negative gradient of the binomial or multinomial deviance loss function [27].

We decided to perform hyperparameter tuning on our GradientBoostingClassifier and set "n_estimators=1000, learning_rate=1.0, max_depth=1". We chose these values because we know that Boosting algorithms are prone to overfitting, so by choosing a high number of boosting stages to perform while shrinking the contribution of each tree by learning_rate, we thought we could counteract that tendency [19].

We set "n_estimators=1000", so there were 1000 iterations per run. We decided to test with different values for "learning_rate" and "n_estimators" and only showed the results for "n_estimators=1000" and "learning_rate=1.0", because this combination showed the best performance. This is because we can have a large "large learning rate" with iterations or have a "slow learning rate" with more iterations. This combination was a good middle choice to get the best of both parameters [24].

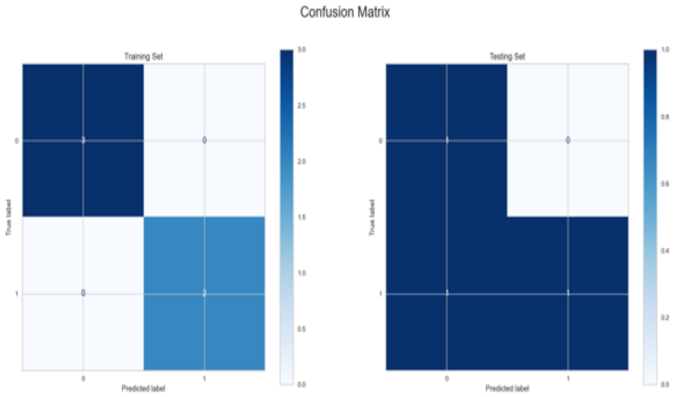


Figure 16: Confusion Matrix for Boosting model

6.4. KNN

This algorithm gave us expected results in that as our training dataset size increased, our test performance decreased significantly, as can be seen in Figures 16-17 below. This is a clear case of the algorithm overfitting on the training dataset and suffering as a result on the testing dataset.

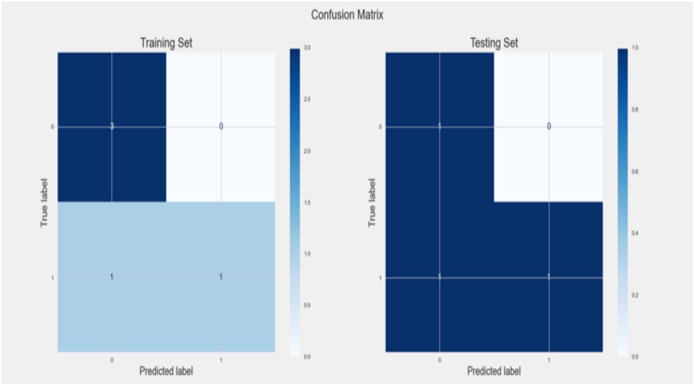


Figure 17: Confusion Matrix for KNN model

This algorithm uses the nearest neighbors in a dataset, where those data points that have minimum distance in feature space from our new data point. In this algorithm, "K" is the number of such data points we consider as part of our implementation of the algorithm. As a result, distance metric and K value are two important considerations while using the KNN algorithm [14].

We decided to perform hyperparameter tuning on our KNN model and test the k values from two through five. We tested these because we wanted to test the underfitting vs overfitting nature of our model and realized that each of the two datasets had different k values that performed the best.

Based on the average runtime for my runs, we got 0.54 seconds per run as per wall clock time.

The only tuning we did on our KNN classifier was testing the k values from two through eight. We didn't show the results for the other values because the performance was degrading after certain "k" values for both datasets.

6.5. Neural Networks

We noticed that this algorithm gave us the most consistent results across both datasets, as can be seen in Figure 18 below. This might be due to the large number of cycles we ran the algorithm on both datasets.

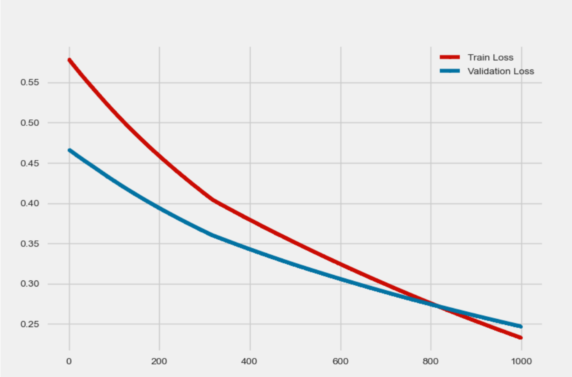


Figure 18: Loss function for NN model

This represents a mind-like algorithm which has different layers of nodes, or neurons, that get activated based on different parameters. In the case of classification datasets, like both of ours, the output layer classifies each example, applying the most likely label. Each node on the output layer represents one label. In turn, that node turns on or off according to the strength of the signal it receives from the previous layer's input and parameters [28].

We decided to perform hyperparameter tuning on our neural networks using two activation functions: "relu" and "sigmoid". We chose sigmoid because "the use of a single Sigmoid/Logistic neuron in the output layer is the mainstay of a binary classification neural network" [27].

We set epoch to 1000 for our tests, so there were 1000 cycles per run and we decided to test with two different stopping criteria: epoch 200 and epoch 1000. We didn't choose to show the results for the epoch 200 because both the accuracy was higher for the epoch 1000 neural networks due to the higher number of cycles that all of the data was being processed [29]. Additionally, we tested with two and three activations layers, but only showed results for two activation layers because the adding the third activation layer hurt accuracy on the test dataset due to the fact that we used "softsign" as our third layer, which skewed results negatively [30].

7. Conclusion

The data we collected from our two subjects confirms that the effect of alcohol consumption by individuals can be strong enough to alter accelerometer readings. These readings can then be used to classify individuals' sobriety levels. One of the limitations of our approach is that clear distinctions are mostly visible only for subjects who are well beyond the regular drinking amount, therefore, for intoxicated individuals who are not well beyond that drinking amount, the classifier model that we built is not as successful in establishing a clear difference between sober and intoxicated states. Additionally, in order to be able to classify "less intoxicated" vs "more intoxicated", we will require larger and more diverse datasets.

7.1. Effect of Cross-Validation

There is tremendous benefit to use a K-folds cross-validation over the standard random data splits because when we build K different models, we are able to make predictions on all of our data. This is especially helpful in smaller data sets so that the algorithm can recognize better patterns [24].

7.2. Definition of Best

For our analysis, we defined "best" as the algorithm that gave us the best balance of the highest test accuracy (not necessarily highest training accuracy) based on a training set size and the fastest execution time.

7.3. Best Classifier

Each algorithm has its own strengths and weaknesses and is therefore good/bad on different types of datasets. That being said, we saw that for datasets that contain both uniform and tailed distributions, such as our subjects' accelerometer data, the Decision Tree Learning was the best Supervised Machine Learning method and should be used as part of our mobile device sobriety classification system.

8. Future Work

In order to build on our research, we want to address the issue of limited test subjects and expand the experiments to include a larger sample size that is comprised of individuals with varied genders, body compositions, backgrounds, etc. [31].

Additionally, we want to explore real world-use cases that can be addressed by our research results. For example, our system can be used with data from any accelerometer, such as those found in car steering wheels. This means that if our system is embedded into cars, then cars' internal systems will potentially be able to warn drivers if their BAC levels are above the legal driving limit, which will directly reduce traffic fatalities caused by alcohol consumption.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

We owe our gratitude to our mentors for guiding us throughout our experiments: Professor Pei Zhang and Dr. Xinlei Chen. We also thank both of our anonymous subjects for participating in our experiments, so that we could monitor their data and use it as part of our analysis.

References

[1] D. Kumar, A. Thanikkal, P. Krishnamurthy, X. Chen, P. Zhang, "Accelerometer-Based Alcohol Consumption Detection from Physical Activity," in 2021 17th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), 415–418, 2021, doi:10.1109/WiMob52687.2021.9606257.

[2] R.W. Olsen, H.J. Hanchar, P. Meera, M. Wallner, "GABAA receptor subtypes: the 'one glass of wine' receptors," Alcohol, **41**(3), 201–209, 2007, doi:https://doi.org/10.1016/j.alcohol.2007.04.006.

[3] Alcohol Facts and Statistics. National Institute on Alcohol Abuse and Alcoholism, 2021, Online: http://www.niaaa.nih.gov/alcohol-health/overview-alcohol-consumption/alcohol-facts-and-statistics..

[4] D. Kumar, Supervised Learning, Georgia Institute of Technology, 2022.

[5] R. Li, G.P. Balakrishnan, J. Nie, Y. Li, E. Agu, K. Grimone, D. Herman, A.M. Abrantes, M.D. Stein, "Estimation of Blood Alcohol Concentration From Smartphone Gait Data Using Neural Networks," IEEE Access, **9**, 61237–61255, 2021, doi:10.1109/ACCESS.2021.3054515.

[6] C. Nickel, C. Busch, "Classifying accelerometer data via hidden Markov models to authenticate people by the way they walk," IEEE Aerospace and Electronic Systems Magazine, **28**(10), 29–35, 2013, doi:10.1109/MAES.2013.6642829.

[7] B. Suffoletto, P. Dasgupta, R. Uymatiao, J. Huber, K. Flickinger, E. Sejdic, "A Preliminary Study Using Smartphone Accelerometers to Sense Gait Impairments Due to Alcohol Intoxication," Journal of Studies on Alcohol and Drugs, **81**(4), 505–510, 2020, doi:10.15288/jsad.2020.81.505.

[8] S. Bae, D. Ferreira, B. Suffoletto, J.C. Puyana, R. Kurtz, T. Chung, A.K. Dey, "Detecting Drinking Episodes in Young Adults Using Smartphone-Based Sensors," in Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., Association for Computing Machinery, New York, NY, USA, 2017, doi:10.1145/3090051.

[9] J.A. Killian, K.M. Passino, A. Nandi, D.R. Madden, J. Clapp, "Learning to detect heavy drinking episodes using smartphone accelerometer data," CEUR Workshop Proceedings, **2429**, 35–42, 2019.

[10] T.H.M. Zaki, M. Sahrim, J. Jamaludin, S.R. Balakrishnan, L.H. Asbulah, F.S. Hussin, "The Study of Drunken Abnormal Human Gait Recognition using Accelerometer and Gyroscope Sensors in Mobile Application," in 2020 16th IEEE International Colloquium on Signal Processing & Its Applications (CSPA), 151–156, 2020, doi:10.1109/CSPA48992.2020.9068676.

[11] Z. Arnold, D. Larose, E. Agu, "Smartphone Inference of Alcohol Consumption Levels from Gait," in 2015 International Conference on Healthcare Informatics, 417–426, 2015, doi:10.1109/ICHI.2015.59.

[12] Aiello, Agu, "Investigating postural sway features, normalization and personalization in detecting blood alcohol levels of smartphone users," in 2016 IEEE Wireless Health (WH), 1–8, 2016, doi:10.1109/WH.2016.7764559.

[13] W. S, How Alcoholism Works, HowStuffWorks Science, 2021, Online: http://science.howstuffworks.com/life/inside-the-mind/ human brain/alcoholism4.htm

[14] Sklearn.svm.SVC, Scikit, 2021, Online: https://scikit-learn.org/ stable/modules/generated/sklearn.svm.SVC.html

[15] A Beginner's Guide to Neural Networks and Deep Learning, Pathmind, 2021, Online: https://wiki.pathmind.com/neural-network#logistic

[16] R. Pramoditha, Plotting the Learning Curve with a Single Line of Code." Medium, Towards Data Science, Medium, Towards Data Science, 2021, Online: https://towardsdatascience.com/plotting-the-learning-curve-with-a-single-line-of-code-90a5bbb0f48a

[17] Sklearn.tree.decisiontreeclassifier, Scikit, 2021, Online: https://scikit-learn.org/stable/modules/generated/sklearn.tree. Decision TreeClassifier.htmll

[18] Romuald_84, Boosting: Why Is the Learning Rate Called a Regularization Parameter?" Cross Validated, Cross Validated, 1963, Online: https://stats.stackexchange.com/questions/168666/boosting -why-is-the-learning-rate-called-a-regularization-parameter..

[19] Sklearn.neighbors.kneighborsclassifier, Scikit, 2021, Online: https://scikit-learn.org/stable/modules/generated/sklearn.neighbors. KNeighborsClassifier.html.

[20] Baeldung, Epoch in Neural Networks, Baeldung on Computer Science, 2021, Online: https://www.baeldung.com/cs/epoch-neural-networks.

[21] C. Chaine, Using Reinforcement Learning for Classfication Problems, Stack Overflow, 1965, Online: https://stackoverflow.com/questions/44594007/using-reinforcement-learning-for-classfication-problems.

[22] Htoukour, Neural Networks to Predict Diabetes, Kaggle, 2018, Online: https://www.kaggle.com/htoukour/neural-networks-to-predict-diabetes

[23] Alcohol's Effects on the Body. National Institute on Alcohol Abuse and Alcoholism, NIH, 2021, Online: http://www.niaaa.nih.gov/alcohol-health/alcohols-effects-body

[24] Sklearn.ensemble.gradientboostingclassifier, Scikit, 2021, Online: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble. GradientBoostingClassifier.html.

[25] Shulga, Dima, 5 Reasons Why You Should Use Cross-Validation in Your Data Science Projects, Medium, Towards Data Science, 2018, Online: https://towardsdatascience.com/5-reasons-why-you-should -use-cross-validation-in-your-data-science-project-8163311a1e79.

[26] Decision Tree, GeeksforGeeks, 2021, Online: https://www.geeksforgeeks.org/decision-tree/

[27] Decision Tree Algorithm - A Complete Guide, Analytics Vidhya, 2021, Online: https://www.analyticsvidhya.com/blog/2021/08/ decision-tree-algorithm/.

[28] Complete-Life-Cycle-of-a-Data-Science-Project, Complete Life Cycle Of A Data Science Project, 2021, Online: https://awesomeopensource.com/project/achuthasubhash/Complete-Life-Cycle-of-a-Data-Science-Project

[29] Machine Learning with Python - Algorithms, 2022, Online: Online: https://awesomeopensource.com/project/achuthasubhash/Complete-Life-Cycle-of-a-Data-Science-Project

[30] K. Team, Keras Documentation: The Sequential Class, Keras, 2021, Online: https://keras.io/api/models/sequential/

[31] D. Deponti, D. Maggiorini, C.E. Palazzi, "DroidGlove: An android-based application for wrist rehabilitation," in 2009 International Conference on Ultra Modern Telecommunications & Workshops, 1–7, 2009, doi:10.1109/ICUMT.2009.5345442.