





Article

Augmented CWT Features for Deep Learning-Based Indoor Localization Using WiFi RSSI Data

Paul Ssekidde ^{1,2,†}, Odongo Steven Eyobu ^{1,2,*,†}, Dong Seog Han ³ and Tonny J. Oyana ²

- ¹ Department of Networks, School of Computing Informatics Technology, Makerere University, Plot 56, Pool Road, P.O. Box 7062 Kampala, Uganda; pssekidde@cis.mak.ac.ug
- ² Geospatial Data and Computational Intelligence Lab, School of Computing Informatics Technology, Makerere University, Plot 56, Pool Road, P.O. Box 7062 Kampala, Uganda; toyana@cis.mak.ac.ug
- ³ School of Electronics Engineering, Kyungpook National University, 80 Daehak-ro, Buk-gu, Daegu 41566, Korea; dshan@knu.ac.kr
- * Correspondence: sodongo@cis.mak.ac.ug
- † These authors contributed equally to this work.

Abstract: Localization is one of the current challenges in indoor navigation research. The conventional global positioning system (GPS) is affected by weak signal strengths due to high levels of signal interference and fading in indoor environments. Therefore, new positioning solutions tailored for indoor environments need to be developed. In this paper, we propose a deep learning approach for indoor localization. However, the performance of a deep learning system depends on the quality of the feature representation. This paper introduces two novel feature set extractions based on the continuous wavelet transforms (CWT) of the received signal strength indicators' (RSSI) data. The two novel CWT feature sets were augmented with additive white Gaussian noise. The first feature set is CWT image-based, and the second is composed of the CWT PSD numerical data that were dimensionally equalized using principal component analysis (PCA). These proposed image and numerical data feature sets were both evaluated using CNN and ANN models with the goal of identifying the room that the human subject was in and estimating the precise location of the human subject in that particular room. Extensive experiments were conducted to generate the proposed augmented CWT feature set and numerical CWT PSD feature set using two analyzing functions, namely, Morlet and Morse. For validation purposes, the performance of the two proposed feature sets were compared with each other and other existing feature set formulations. The accuracy, precision and recall results show that the proposed feature sets performed better than the conventional feature sets used to validate the study. Similarly, the mean localization error generated by the proposed feature set predictions was less than those of the conventional feature sets used in indoor localization. More particularly, the proposed augmented CWT-image feature set outperformed the augmented CWT-PSD numerical feature set. The results also show that the Morse-based feature sets trained with CNN produced the best indoor positioning results compared to all Morlet and ANN-based feature set formulations.

Keywords: indoor localization; data augmentation; continuous wavelet transform; deep learning



Citation: Ssekidde, P.; Steven Eyobu, O.; Han, D.S.; Oyana, T.J. Augmented CWT Features for Deep Learning-Based Indoor Localization Using WiFi RSSI Data. *Appl. Sci.* **2021**, *11*, 1806. <https://doi.org/10.3390/app11041806>

Academic Editors: Subhas Mukhopadhyay and Wonjoon Kim

Received: 18 November 2020
Accepted: 10 February 2021
Published: 18 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Localization is known to be useful in various domains, such as medical care, police, firefighting, industries, museums, airports, and shopping malls. Specifically, the positions of defined objects in the mentioned domains are sought for using localization systems. A few examples of such objects include pedestrians in a building or airport, auto-motives on the road, foreign objects in the human body, mobile animals in a game park, and merchandise on a shelf. A key example is the precise positioning of mobile objects is one key challenge in navigation studies.

Positioning systems are classified as either indoor or outdoor depending on their development purposes and areas of application. Outdoor positioning systems have had a huge number of breakthroughs in terms of the accuracy and precision of the tools in the market today. Such outdoor positioning tools include the global positioning system (GPS), the BeiDou, and the Quasi-Zenith Satellite System (QZSS). The same outdoor tools are not accurate for indoor environments given the high levels of signal interference, absorption, multipath propagation, and fading effects in indoor scenarios. Therefore, indoor positioning solutions are currently a major work on progress in the navigation research community.

Various technologies, including infrared, Bluetooth, ultra-Wide band (UWB), wireless fidelity (Wi-Fi), RFID, cameras, sound, and magnetic systems, are basically used as workbenches for localization data generation. Different indoor positioning solutions based on these technologies [1–10] have been studied. Positioning approaches proposed in [1–4] are based on image data collected from indoor environments, while those in [5–8] were based on received signal strength indicator (RSSI) data.

Conventional image-based positioning approaches [1–4] use images of various kinds generated from the target indoor environment. Examples of such images include photographic images [11], radar signal graphical data [12], and RSSI CWT graphs [13]. Such images are used in the prediction or estimation of indoor positions and hence are used as fingerprints. This is what is referred to as combining image-based positioning with fingerprint-based positioning. So far, very limited studies [13] show that fingerprint data has been applied in deep learning systems for localization purposes. Notably, the most prominent fingerprint data transformed to feature sets are the RSSI.

In order to build deep learning-based localization systems, indoor feature representation is very critical. It is said features which compose the fingerprints used as indoor datasets. Popular feature representation approaches include the use of the time domain, frequency domain, and joint time-frequency domain. Examples of time domain features include mean, variance, standard deviation, root mean square, mean crossing rate, and many more. Examples of frequency domain features include spectral energy commonly referred to as PSD and the discrete fast Fourier transform coefficient. An example of a joint time-frequency feature is the wavelet coefficient. The CWT, therefore, falls under the time-frequency domain.

Studies such as Soro et al. [13] have used CWT-based images to build a feature set from RSSI data usable for indoor localization using CNN and K-nearest neighbors (k-NN). Notably, Soro et al. [13] was the first to use CWT-based images to solve indoor localization problems using deep learning. In said study, the Morlet mother wavelet was used to build CWT graphical images which were the basis of the indoor fingerprint regarded as the dataset.

This paper proposes a CWT-based approach for formulating the indoor fingerprint using RSSI data but introduces data augmentation in order to improve on the quality of the realized CWT-based dataset for purposes of improving the localization accuracy. Data augmentation is a technique which is used specifically to increase the size of a dataset due to data scarcity [14]. However, augmented data are expected to be representative of the real experimental data; hence, they require careful creation. A few of the various approaches used for creating augmented data include using Gaussian noise, local averaging, rotation, and time warping. Additionally, both the Morlet and the Morse mother wavelets were extensively experimented on to evaluate the performance of the augmented CWT features. The generalized Morse and Morlet wavelets were chosen so as to capture the general performance of limiting special forms of wavelets in both families. The key contributions of this paper are therefore summarized as follows:

- A dataset based on CWT scalogram image data. The CWT images were augmented with CWT scalograms derived from recorded RSSI signals complemented with additional white Gaussian noise (WGN).

- A numerical CWT PSD dataset derived from CWT scalogram images. The numerical CWT PSD dataset was augmented with numerical PSD's of CWT scalograms derived from recorded RSSI signals complemented with additional WGN.
- A performance analysis of the proposed datasets together with conventional feature set formulations. All datasets were trained using CNN and ANN to estimate the room a human subject is in and also estimate the exact position of the human subject in a room.

The rest of this article is organized as follows: Section 2 discusses the related works in indoor localization. Section 3 presents the preliminaries. Section 4 presents the proposed feature extraction algorithm and augmentation method. Section 5 presents the experimental setup. Section 6 presents the results and discussions. Finally, Section 7 concludes this article.

2. Related Works

This section mainly focuses on related works on deep learning approaches used for localization and feature extraction techniques particularly in indoor environments. The aim of this section is to point out the performances of different studies to aid analysis and validation of the proposed algorithm in this study.

In [15], a localization model was proposed by combining CNN with Gaussian process regression (GPR) using Wi-Fi RSSI data. RSSI fingerprints were captured from several access points in order to generate a CNN training dataset which was used for feature extraction after the reprocessing stage. In order to improve the performance of the model, the GPR algorithm was applied for coordinate adjustment of the target node. However, this approach exhibits a high computational complexity in order to achieve better prediction accuracy. Although in this study we also used CNNs in the experiments, our computational complexity in terms of training time was low, which is attributed to the proposed CWT feature sets. Our real-time positioning time complexity was low when using the trained deep learning model.

Shao et al. [10] developed a novel hybrid location image based on sensor data series from both WiFi and magnetic field data. CNN was applied to classify the location of images for a defined reference point. Two steps for learning were adopted in order to acquire patterns of different networks. Additionally, these steps are: first, an early-stop method for preventing overfitting caused by small datasets for WiFi fingerprint features; secondly, long-term training of stable magnetic field for learning the different image features. This proposed method achieved a localization error of 1 m considering different smartphone orientations, users, and even use patterns.

Dou et al. [16] proposed a top-down searching method using a deep reinforcement learning agent to handle dynamics in indoor environments using WiFi fingerprints. The model uses a hierarchical search strategy that begins from the whole area towards the target node. This method does not require prior knowledge of the floor plan. For a floor size of 2500 square meters, 75% accuracy was reported.

Njima et al. [17] developed a localization system that transfers the online prediction complexity to an offline preprocessing step using a convolutional neural network. The fingerprints were generated based on 3D radio images formulated using RSSI data, and CNNs were used for target training and prediction. Using finger print data, Kurtosis values were calculated and used to provide statistical parameters that distribute global information to the local filters. However, Njima et al. [17], considered localization accuracy and computational complexity as a trade-off, yet scalability and energy-efficiency are sometimes key metrics towards high accuracy of the target.

Ibrahim et al. [18] presented a CNN-based approach for indoor localization using RSSI time-series from WiFi access points. They demonstrated the approach on the UJIIndoorLoc dataset considering multi-building and multi-floor data and obtained 100% accuracy on both building and floor prediction using a mobile node. Four feature set formulations were implemented, namely, single RSS vector, averaged RSS time-series, concatenated

RSSI time-series, and RSSI time-series images. Feedforward DNNs and CNN were used in the prediction of the building floor and specific node location coordinates (longitude and latitude). The study reported that the RSSI time-series image dataset formulation yields a better result. However, it requires huge datasets, which is expensive in terms of manpower and time for data collection.

Hsieh et al. [19] generated a deep learning-based approach for indoor localization through combining received signal strength (RSS) and channel state information (CSI). However, CSI-based approaches need extra hardware modifications in the mobile nodes in order to generate the data, and this makes them expensive compared to RSSI-based methods that do not require additional hardware.

Abbas et al. [20] presented a WiDeep, a deep learning-based indoor localization system that achieves a fine-grained and robust accuracy in the presence of noise. It associates a stacked denoising auto encoder deep learning model with a probabilistic framework to reduce noise in WiFi fingerprints and capture the complex relationship between WiFi AP signals that run on the Android device and their locations. WiDeep employed regularization to avoid overfitting in deep learning models. WiDeep was evaluated with two separate test beds, a large university building and a small domestic apartment with different WiFi APs. The model achieved mean localization accuracy of 2.64 m and 1.21 m for the larger and the smaller test beds, respectively.

Chen et al. [21] created a sensor fusion framework which combines WiFi, Pedestrian Dead Reckoning (PDR), and landmarks. The data obtained by all these sensors were fused using the Kalman filter for position estimation. The performance of the system was evaluated by comparing the proposed fusion approach with individual approaches of WiFi weighted path loss (WPL) and PDR with landmarks. The proposed approach produced an average localization error of 1 m, which is an interesting performance.

Kumar et al. [22] proposed an indoor localization of vehicles that uses a CNN with deep learning to identify and localize vehicles in a car-park. Based on the study, a vehicle can localize itself with landmark-based positioning and internal car sensors, but is costly and requires precise knowledge of the enclosed area. The approach used off-the-shelf cameras mounted in the car-park and vehicle-to-infrastructure communication to enable all vehicles to obtain an indoor position. In this work, a pre-trained Google Net model was applied for object class classification of the images obtained from a public dataset known as ImageNet and other small images captured by cameras in their car garage. The evaluation of system was done on an eValet underground car garage equipped with network cameras monitoring several lanes. The approach generated an accuracy of 65.69% on the test data.

Wang et al. [23] developed DeepFi, a novel deep-learning-based indoor fingerprinting system using CSI. The experiment was conducted using one fixed AP and a mobile terminal. The mobile device receives transmitted wireless signals from the access point and stores the raw CSI values in its firmware. The CSI values from the firmware were read and assigned weights. The study proposed using the weights in a deep network to represent fingerprints and apply a greedy learning algorithm for weight training in order to reduce complexity. DeepFi achieved good accuracy; nevertheless, capturing fingerprints using CSI is not cost effective and requires a powerful machine to perform localization tasks.

Li et al. [24] proposed an indoor localization system using the affinity propagation clustering algorithm and the particle swarm optimization-based artificial neural network (PSO-ANN). Both algorithms were used on WiFi-based RSS data. The clustering technique was implemented with the purpose of reducing the location error. The PCA algorithm was used to reduce the dimensions of RSS fingerprint data. Additionally, the PCA was used with a purpose of reducing the computational complexity during the ANN training process. The affinity propagation clustering was also performed in an unsupervised approach in order to reduce the computational complexity. The proposed algorithm produced an error of 2.971 m.

Anindya et al. [5] proposed a sigma-point Kalman smoother (SPKS)-based location and tracking algorithm as a superior alternative for indoor positioning. In their experi-

mentation, different low cost sensors were used to capture the indoor environment data, including RSSI, infrared motion sensor data, and binary foot switches. Data from all these sensors were fused using a sigma-point Kalman smoother to determine the human user's indoor position. Their study presents superior results compared to using the extended Kalman filter and a commercially available positioning tool.

Wang et al. [25] proposed the use of a support vector machine (SVM) and an improved Kalman filter (KF) to estimate a target position based on RSSI. They applied the SVM in order to produce the estimated positions which are enhanced using the Kalman filter. However, Wang et al. report that the combination of SVM and the KF introduces a higher run-time. The proposed approach in our study only uses only a deep learning-trained model and does not exhibit real-time positioning delays except during the training of the deep learning model.

Pubudu et al. [26] proposed a robust extended Kalman filter (REKF) algorithm which estimates the positions of sensors which are attached to mobile robots in an indoor environment. This applications for which the approach was proposed are delay tolerant networks (DTNs). The approach of enabling sensor node mobility was done to eliminate the processing constraints of static sensor nodes and the need for static reference beacons. The positioning results show that the proposed REKF achieved localization accuracy within 1 m in the indoor environment.

Soro et al. [13] used a CWT to extract time-frequency features from the transformed image and used a CNN to train and learn the patterns. A K-nearest neighbors (KNN) model was then applied to define the target location. The method achieved good performance compared to some existing models. The work in [13] is the closest match to an existing feature extraction approach using the CWT for indoor localization compared to the idea in this paper. The difference is that this paper introduces the concept of data augmentation in order to further improve on the localization performance exhibited by the CWT feature set without data augmentation. Additionally, extensive experimentation was done to evaluate the performance of the augmented CWT features by applying Morlet and Morse.

Various data augmentation approaches have been developed to improve on the performance of deep learning models. Some of these include local averaging [14] and jittering [27] for time series data. In this paper, jitter-caused additive Gaussian noise is used as a tool for augmenting the RSSI signal.

Table 1 shows a summary of existing feature set formulations used in indoor localization studies. Each of the feature set formulations is characterized in terms of the technical dataset composition and whether its augmented or not.

Table 1. Some existing deep learning feature set formulations used in indoor localization studies.

Dataset	Feature Characteristics	Augmentation Characteristics
Soro et al. [13]	CWT images	none
Ibrahim et al. [18]	RSSI time images	none
Njima et al. [17]	RSSI images	none
Kumar et al. [22]	Camera images	none
Wang et al. [23]	Channel state information	none
Ibrahim et al. [18]	RSSI	none
Our Proposed	(1). CWT images (2). Numerical CWT PSD	AWGN AWGN

3. Preliminaries

In this section, we give a brief summary of the mother wavelet functions used in this study, PCA and CNN.

3.1. Morlet and Morse Wavelet Functions

The analyzing function is always referred to as the mother function. The Morlet and Morse are the two selected mother functions used to generate the CWT in this study. The Morlet function ψ is denoted by:

$$\psi(t) = e^{-\frac{t^2}{2}} \cos(5t) \quad (1)$$

where t is time in seconds and $\psi \in L^2(\mathbb{R})$ such that L is the norm. Figure 1a represents the standard Morlet wavelet with $0.5 \leq \gamma \leq 0.5$. ψ is a bandwidth and time-limited function with a zero-mean and unit energy [28]. The CWT requires shifted and stretched versions of the analyzing mother wavelet ψ , which is compared with the signal. The stretched or compressed versions of the Morlet mother wavelet as dilated wavelets can be denoted by $\psi_{\tau,\gamma}$ which is defined in Equation (2). Figure 1b represents the standard Morse mother wavelet.

$$\psi_{\tau,\gamma}(t) = \frac{1}{\sqrt{\gamma}} \psi\left(\frac{t-\tau}{\gamma}\right) \quad (2)$$

where γ is the scaling or dilation factor that controls the width of the wavelet and τ is the translation parameter controlling the location of the wavelet. Scaling a wavelet means stretching it (if $\gamma > 1$) or compressing it (if $\gamma < 1$), and translating it means shifting its position in time. The generalized Morse wavelet is a two-parameter family of wavelets defined in the frequency domain by:

$$\psi_{\beta,\gamma}(\omega) = K_{\beta,\gamma} H(\omega) \omega^\beta e^{-\omega^\gamma} \quad (3)$$

where $K_{\beta,\gamma}$ is a normalizing constant and $H(\omega)$ is the Heaviside unit step function. To be a valid wavelet, one must have $\beta > 0$ and $\gamma > 0$. By varying these parameters, the generalized Morse wavelets can be given a broad range of characteristics while remaining exactly analytic.

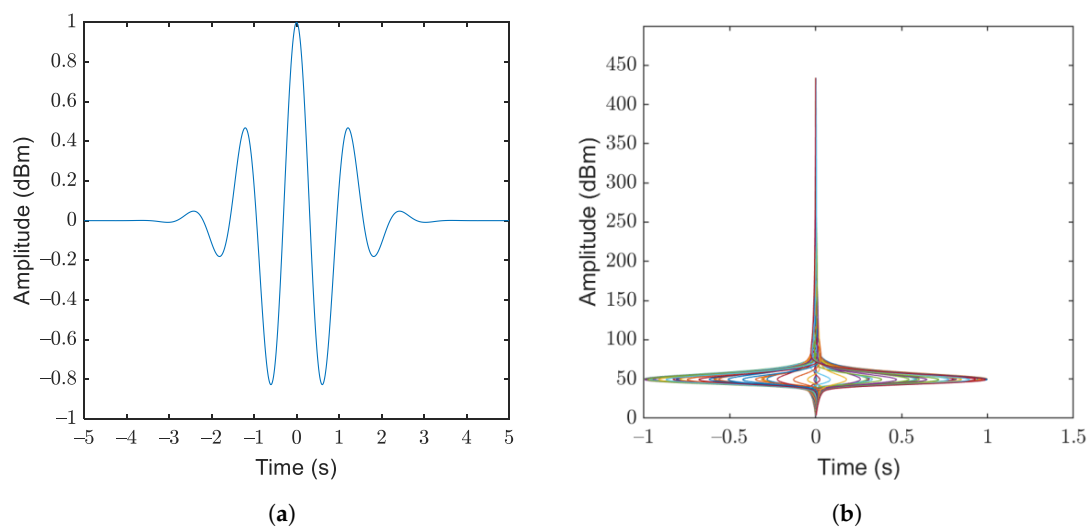


Figure 1. (a) A standard Morlet mother wavelet. (b) A standard Morse mother wavelet.

Fourier transformations are popular in creating feature representations for deep learning systems. However, under the Fourier transform, the time information of a time series is lost [29]. Due to this loss, it is hard to distinguish transient relations or to identify when the structural changes do occur. Therefore, wavelet analysis was introduced given that it performs the estimation of spectral properties of a time series as a function of time while revealing how different periodic components of the time series change over time. It is for this reason that the CWT was used in the feature set formulation in this study.

We have implemented the generalized Morse and Morlet wavelets so as to capture the general performances of limiting special forms of wavelets in both families. In summary, the Morlet wavelet represents a one-parameter family of functions, while the Morse wavelet represents a two-parameter family of functions. The limiting or special forms of the generalized Morse wavelets can be found if the frequency argument ω is rescaled by the peak frequency $\omega_{\beta, \gamma}$, defining a new frequency-domain wavelet [30], as seen in Equation (3). Some special forms of the Morse family include the Cauchy wavelet and the Bessel wavelets. Only the generalized Morse and Morlet wavelets were used in this study so as to capture the general performances of limiting special forms of wavelets in both families.

Using the Morlet or Morse mother wavelets, a CWT of a formulated RSSI vector was obtained. The CWT graphs were adopted as images used as fingerprints for training a CNN to estimate the human user position. Similarly, the CWT numerical data, which were essentially the PSD data—in this text referred to as the CWT-PSD data—were used as fingerprint information to train an ANN to estimate the human user position.

3.2. White Gaussian Noise

The Gaussian distribution was used to build synthetic RSSI data. Descriptive statistics were applied on the recorded RSSI data and the variance used in the Gaussian distribution function. The observed data were normally distributed. The synthetic RSSI data which contained additional WGN were mainly tuned by normalization to be between 0 and 1 before a CWT was performed to generate the frequency-time data. The data were then appended to the originally recorded RSSI data to make up the proposed features for deep learning purposes.

3.3. Principal Component Analysis

PCA is a non-parametric method used for extracting relevant information from confusing datasets [31]. PCA reduces the complexity of a dataset by dimension reduction, hence revealing the hidden underlying factors in data. PCA has for long been applied in image processing [32] for size reduction without compromising its quality. Based on PCA's capabilities as a tool, we have adopted it as appropriate tool to trim or reduce the size of data attained during the CWT data numerical preparation process.

3.4. Convolutional Neural Networks

In this subsection, brief descriptions of the CNN and its components, and how these components interact in order to learn the image patterns and finally predict the target node are presented. Although CNNs are technically not different from ANNs, the traditional form of an ANN is limited in processing image data, as it exhibits high computational complexity. The CNN architecture was built for image-focused tasks, and is hence a better alternative when dealing with image datasets [33]. Since this study used both image and numerical data, CNN and ANN were used for the experiments in this study. The CNN architecture consists of convolution layers, pooling layers, and fully connected layers, as illustrated by Figure 2. To improve the performance of the CNN model, dropout, batch normalization, and regularization were implemented as measures to control challenges such as overfitting.

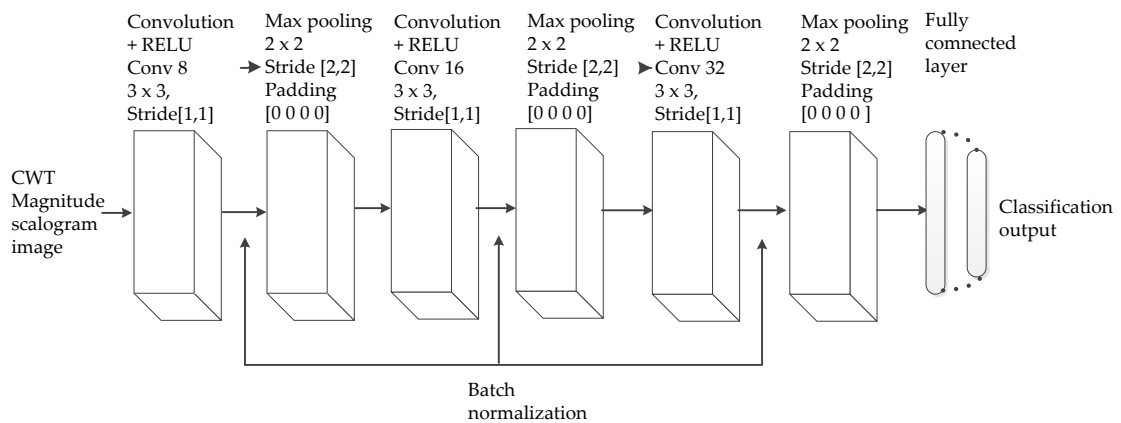


Figure 2. The general structure of the convolutional neural network considered.

4. Proposed Algorithm

The proposed algorithm for predicting the position of a human subject in an indoor environment is applied for two purposes. The first is to identify the room where the human subject holding the smartphone receiving the RSSI data is. The second is to estimate the precise position of a human subject holding a mobile phone in a room through the use of Wi-Fi RSSI fingerprints captured from Wi-Fi access points that are fixed at different locations. These two goals are achieved through the process flow illustrated in Figure 3. However, first, the description of the initial RSSI feature set formulation is presented next before the next subsections on augmentation.

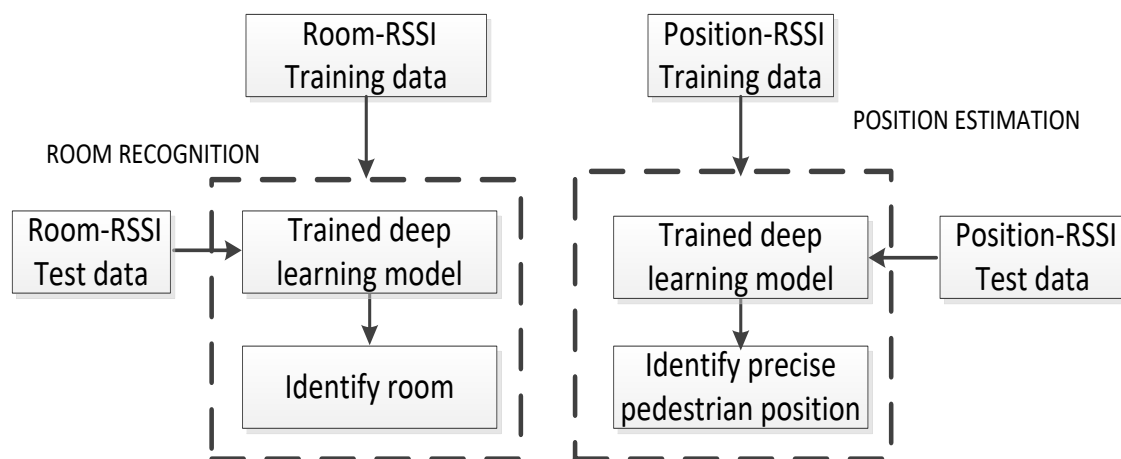


Figure 3. Model of the system.

Raw RSSI from different APs are captured at different reference points in each room. The window period for each RSSI capture is 1 s. The captured RSSI data are considered as a signal after normalization. Raw RSSI data from the APs are collected and normalized, especially for gradient-based optimization methods in order to accelerate the learning process and minimize the risk of algorithm divergence [7]. Normalization of the RSSI data is done to take care of APs whose signals have not been received at the reference point and those APs with a very low signal strength. We adopted the normalization approach used in [15,34] which is as follows: (1) All APs which were invisible are set to -120 dBm such that the normalization result is zero. (2) All RSSI below -110 dBm are set to zero, because at that level, the signal is susceptible to noise which may render it unreachable. Generally, all RSSI data after normalization are between 0 and 1. It is these normalized RSSI data for

each window of 1 s that are considered as the normalized RSSI signal for which the CWT is performed with the purpose of extracting the CWT graphical image and the CWT-PSD numerical data. At this stage, all CWT images are considered as the initial dataset except for the CWT-PSD, which must still under go dimension reduction using PCA.

4.1. Augmented CWT Image Feature Set Formulation

For each time window segment, after CWT generation, white Gaussian noise is added to each normalized RSS fingerprint signal, after which a CWT is again generated for the purpose of augmenting the initial CWT image feature set. Figure 4 includes the augmented CWT image formulation.

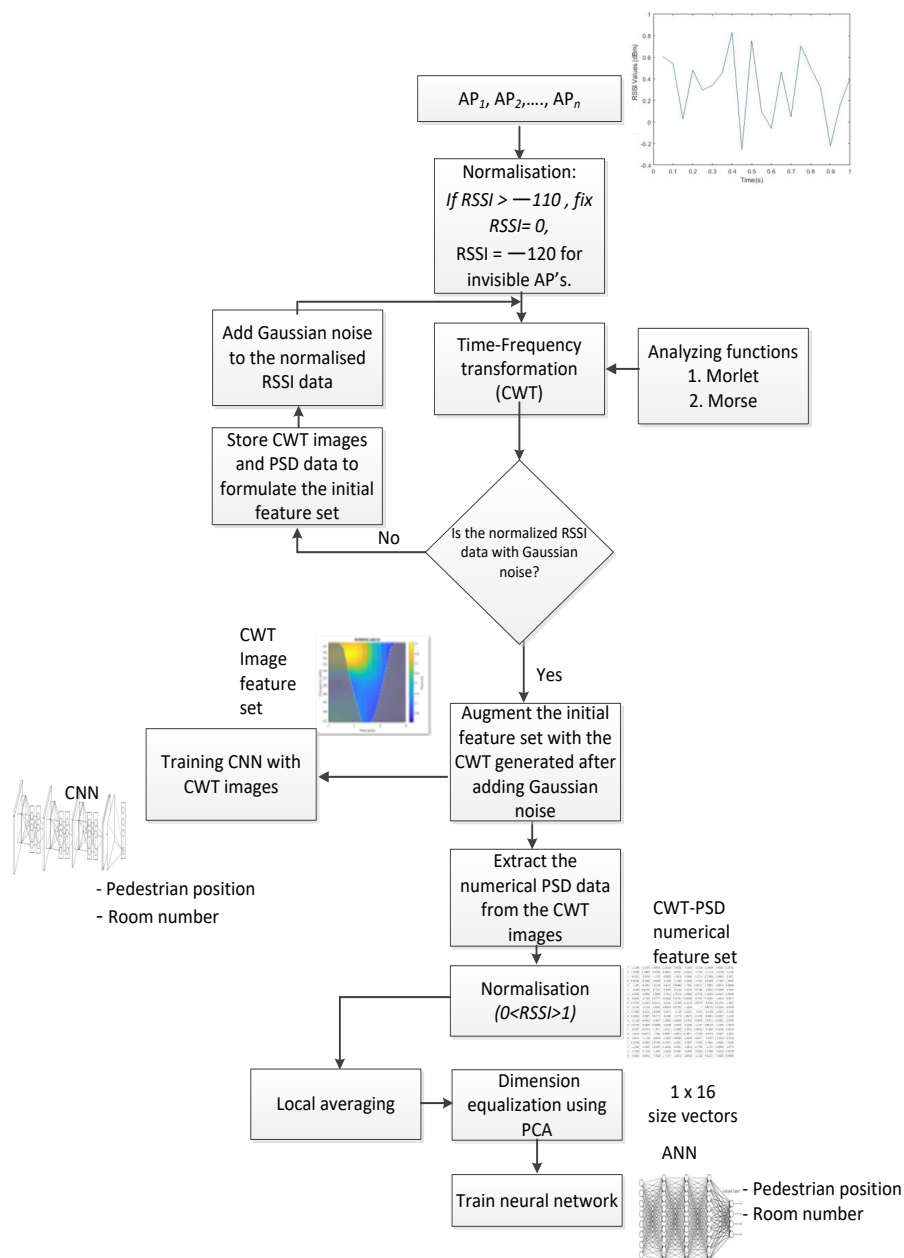


Figure 4. Generation of the augmented CWT features dataset.

4.2. Augmented CWT-PSD Numerical Feature Set Formulation

Given the variation in the dimensions of the CWT matrix numerically, a process of dimension equalization using the PCA is necessary. This process was done after locally averaging the numerical data. The said CWT matrix numerical data are the CWT PSD data.

The PSD data used from each CWT image are of vector size 1×16 after PCA equalization. These PSD data are referred to as the initial CWT-PSD numerical feature set. For each time window segment, after PSD extraction, white Gaussian noise is added onto each normalized RSS fingerprint signal, after which the PSD data are again extracted for purposes of augmenting the initial CWT-PSD numerical feature set. The process flow of the augmented CWT-PSD numerical feature set is included in Figure 4.

5. Experimental Setup

The experiment was conducted on the third floor of Block A, in the Network department at the college of computing and information science of Makerere University, as indicated in Figure 5. Five offices (rooms) were considered with over thirty one (31) access points labeled as AP1–AP31.

A customized Wi-Fi analyzer application installed in a smart mobile phone was used to capture the RSSI values from APs within defined reference points seen in Figure 5.

Fifteen (15) reference points for each room were considered by obtaining the RSSI vector at each point for room identification. The RSSI from different APs for each reference point were captured thrice at different times, each lasting for a period of 1 s. The RSSI data captured from the APs formulated a RSSI vector which was normalized and considered as a signal. This implies that each reference point had three RSSI data vectors which were captured in 3 s. In total, 45 RSSI vectors were captured for each room, since each room contains fifteen reference points. It is the 45 RSSI vectors which were processed for room identification. A CWT was performed on each RSSI vector signal and the CWT graphical image was extracted. The CWT-PSD numerical data were also extracted from each CWT graphical image of each RSSI vector signal. Since the window size was fixed to 1 s and the RSSI capture was done three times for each reference point, each reference point had three CWT graphical images. The CWT-PSD data generated a dataset size equivalent to the number of rows in the three different CWT-PSD 2-dimensional matrices attained in the 3 s. The separation distance between the reference points was one to two meters.

Three (3) reference points from each of the five rooms used in the experiment were selected to collect RSSI data for specific reference point estimation. Three reference points which were close to each other were sampled out to be representative of all the reference points in a room, given that the separation distance between all the points was uniform. Again, in order to take care of the RSS variations exhibited at the various positions in a room, 200 RSSI vectors were collected at each reference point, which implies that a total time of 200 s was used at each point. The assumption was that by increasing the time-space for RSSI data acquisition, we would be able to capture the RSS diversity at the different points in the room which were represented by the three sampled reference points. Collectively, 15 reference points in all the five rooms were considered, making the total number of RSSI data vectors 3000. It was the 3000 RSSI vectors which were used to estimate each of the selected specific positions in each room. For each RSSI vector, after augmentation using additive Gaussian noise, the feature set size was doubled. A summary of the dataset sizes is presented in Table 2.

For purposes of clarity, in the proceeding text in this paper, we have chosen to refer to the augmented CWT feature set and the augmented numerical CWT PSD feature set derived in Figure 4 as the proposed augmented CWT feature set and the CWT-PSD feature set respectively. We have chosen to follow this nomenclature in order to signify the CWT feature set as the key superior feature set formulation presented in this paper.

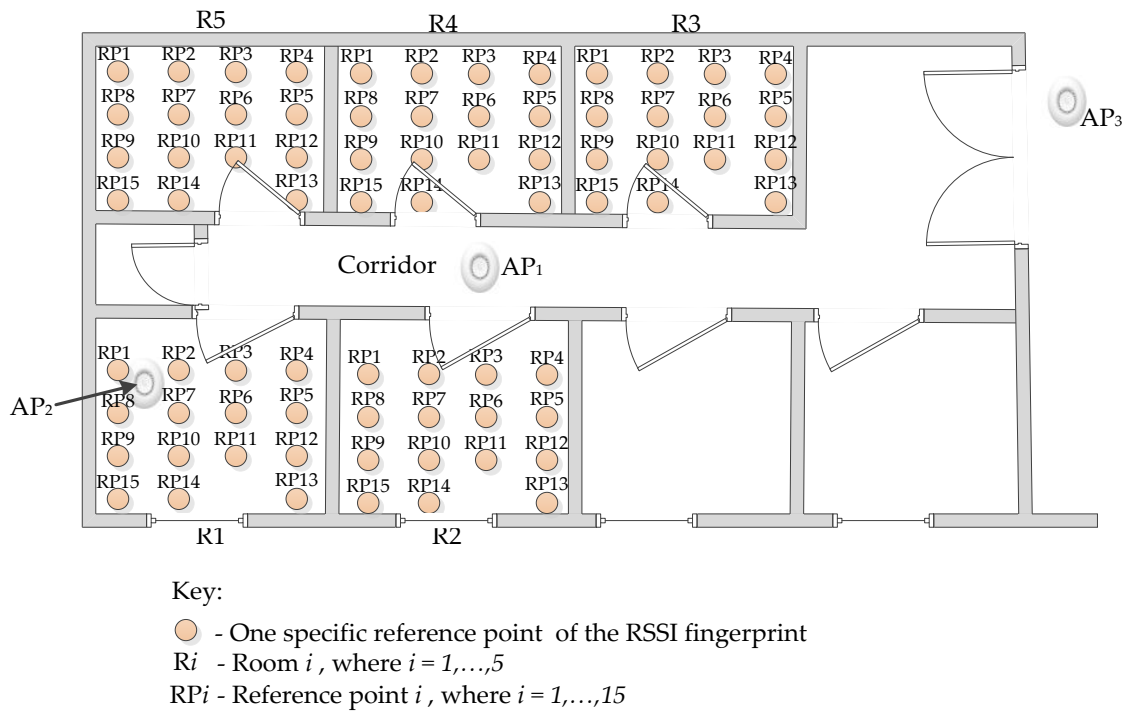


Figure 5. Fingerprint data collection scenario.

Table 2. Dataset sizes.

Feature Set	Room Identification		Specific Reference Point Estimation	
	Training Data Size	Test Data Size	Training Data Size	Test Data Size
RSSI vector [34]-CWT and FFT	170	55	2250	750
Unaugmented CWT	170	55	2250	750
Proposed augmented CWT	340	110	4500	1500
Unaugmented Numerical CWT-PSD	169	34	2250	450
Augmented numerical CWT-PSD	338	68	4500	900

For purposes of determining the optimal CNN and ANN configuration parameters, arbitrary values of the number of layers (10–30), learning rates (0.01–0.5), and maximum iterations (100–200) were tested. The best performing parameter values were adopted, as shown in Tables 3 and 4.

Table 3. CNN parameters and configurations.

Description	Value
Fully connected	15 fully connected layers
Softmax	softmax
Learning rate	0.01
Maximum iterations	150
Iterations reached on training	150

Table 4. ANN configuration.

Description	Value
Number of hidden nodes	2500
Neural Network	FeedForward
Train data, Validation data and Test data samples	75%, 10% and 15% respectively

Table 5 shows the performance metrics which were used in the performance evaluation of the proposed feature set.

Table 5. Performance metrics.

Measure	Description
Accuracy	The percentage of correctly predicted samples out of the total number of samples.
Precision	is the fraction of relevant instances among the retrieved instances.
Recall	is the fraction of the total amount of relevant instances that were actually retrieved
Prediction error	A percentage measure of the number of wrongly predicted positions given the total number of positions.
Localization error	The Euclidean distance between the actual and predicted position.

For this study, a real-time positioning system using the proposed algorithm was developed. The hardware composition was a smartphone dedicated to receive the signals from the APs and then route them to a server containing the RSSI data processing and localization system.

6. Results and Discussion

In this section, we report the performance of our proposed augmented CWT feature sets while comparing it with existing feature set formulations used for indoor localization applied in deep learning models. The two existing feature set formulations used for validating our proposed feature set were those proposed in [13,34]. In [13] joint a CWT-based feature set was proposed but without any augmentation. In [34], the feature set was composed of single RSSI images in which nine dummy values were added to each RSSI vector to generate a 2D image for CNN training and classification. We further show the effect of the Morlet and Morse wavelet transformed data on the room and specific position performance. Firstly, we present results for room identification in Section 6.1. Secondly, the performance results for specific reference point recognition are presented in Section 6.2.

6.1. Room Identification Results

Table 6 shows the confusion matrices for room identification based on the Morse wavelet. Table 6a shows the performance exhibited by the CWT features using CNN with an accuracy of 92.73%. Table 6b shows the proposed augmented approach using CNN with an accuracy of 97.3%. The proposed approach exhibits a superior performance because of the augmented features generated by addition of Gaussian noise to the original RSSI signal before CWT transformation. The results in Table 6a are attributed to a dataset with 255 images, and the augmented dataset whose results are in Table 6b was of size 450 images. Based on the accuracy results, the better performance of the augmented dataset implies that the Gaussian noise introduces more distinguishable features which are desirable for an RSSI classification.

Table 6c represents the proposed CWT-PSD without data augmentation using ANN with an accuracy of 88.2%. Table 6d shows the proposed augmented CWT-PSD using ANN with an accuracy of 70.6%. Contrary to the image-based CNN classification in Table 6b, the augmented PSD using ANN did not perform better than the CWT-PSD without augmentation represented by the confusion matrix in Table 6c. This phenomenon reveals that augmented images provide more distinguishable patterns compared to augmented numerical data. Studies such as [35] show that augmentation greatly increases the performance of CNN-based classification.

Table 6. Confusion matrices for room identification: (a–e) using the Morse mother wavelet, and (f) using FFT.

CWT without Augmentation Using CNN							Proposed Augmented CWT Using CNN					
True labels	R1	9	0	0	0	0	R1	21	0	0	0	0
	R2	0	11	0	0	0	R2	0	22	0	0	0
	R3	1	0	9	0	0	R3	0	0	21	1	0
	R4	1	0	0	11	0	R4	0	0	1	21	0
	R5	0	0	2	0	11	R5	1	0	0	0	22
		R1	R2	R3	R4	R5		R1	R2	R3	R4	R5
Predicted labels							Predicted labels					
(a)							(b)					
CWT-PSD without Augmentation Using ANN							Augmented CWT-PSD Using ANN					
True labels	R1	6	0	0	0	0	R1	4	0	0	1	2
	R2	0	5	1	0	0	R2	0	19	1	0	0
	R3	0	0	8	0	0	R3	0	0	6	0	0
	R4	1	0	0	6	0	R4	9	0	0	10	1
	R5	1	0	0	1	5	R5	1	0	0	5	9
		R1	R2	R3	R4	R5		R1	R2	R3	R4	R5
Predicted labels							Predicted labels					
(c)							(d)					
Single RSS Vector [34] CWT Using CNN							Single RSS Vector [34] FFT Using CNN					
True labels	R1	9	0	0	0	0	R1	9	0	2	0	0
	R2	0	11	0	0	0	R2	0	11	0	0	0
	R3	0	0	10	0	0	R3	0	0	8	0	0
	R4	0	0	1	11	0	R4	1	0	0	10	1
	R5	2	0	0	0	11	R5	1	0	1	0	10
		R1	R2	R3	R4	R5		R1	R2	R3	R4	R5
Predicted labels							Predicted labels					
(e)							(f)					

The dataset size of the proposed CWT-PSD without data augmentation using ANN was 255 vectors, and that of the augmented CWT-PSD was 450. Despite the fact that a huge dataset increases classification performance, a huge dataset can also degrade classification performance, especially if the augmented data cause a huge change in the distribution of the data, causing complexity in separating the data. This analysis implies that the Gaussian noise is not the best for augmenting numeric data in a classification problem.

Table 6e shows the performance of the single RSSI vector in [34] with CWT as the feature extraction approach with an accuracy of 95.7%. Table 6f shows the performance of the single rssi vector in [34] with FFT as the feature extraction approach with an accuracy of 87.3%. The accuracy results for the CWT features compared to the FFT features in the single RSS vector [34] approach are better. This points to the ability of CWT to better distinguish features in different classes. Generally, the proposed augmented CWT feature set performed better than all the other mentioned feature sets (Table 6). It is also clear that the CNN model outperformed the ANN in classification. The general results in Figure 6 also show that the Morse wavelet provides a better feature set compared to the Morlet wavelet.

For the rest of the figures in this article, we use the acronyms “A, B, C, D, E, F, G” to represent the CNN-based Morse wavelet-generated feature set, the CNN-based Morlet wavelet-generated feature set, the ANN-based Morse wavelet-generated feature set, the ANN-based Morlet wavelet-generated feature set, the CNN-RSSI [34] Morse-based feature set, the CNN-RSSI [34] Morlet-based feature set, and the CNN-RSSI [34] FFT-based feature set.

Figure 7a–d are presented to compare the performance of a deep learning approach used together with a wavelet. This should provide information on which deep learning approach should be used together with a specific wavelet transform to achieve good classification performance in terms of precision and recall. Figure 7a shows that for the augmented dataset, the Morse wavelet-generated features, when used in the CNN deep learning model, generated the best performance in terms of precision. The worst performance in Figure 7a is exhibited by the Morlet wavelet features used in ANN.

Figure 7b shows that for the unaugmented dataset, the Morlet wavelet-generated feature set used with the ANN model offered the best precision. Additionally, the worst was the Morse wavelet dataset applied to the ANN model. The ANN-based Morse precision performance for R4 was 0%, which is an indicator that the the time, R4 was not correctly identified. From the original RSSI data, R4’s RSSI signal has a huge similarity with all the other rooms’ RSSI signals. This gives us the insight that without augmentation, it is highly likely that the classification of R4 would generate a very minimal precision.

Figure 7c,d show the recall performance for the augmented and unaugmented datasets respectively. The results still show that the CNN-based Morse wavelet model performed better in the augmented feature set and the ANN-based Morlet wavelet model performed worst in the unagmented feature set. Generally, on average, the CNN-based Morse wavelet dataset performed best for accuracy, recall, and precision for the augmented dataset, and the ANN-based Morlet was preferable for the unaugmented dataset.

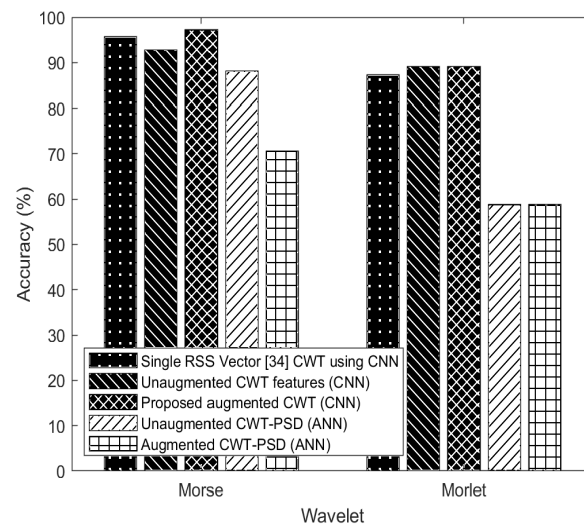


Figure 6. Accuracy results for room identification.

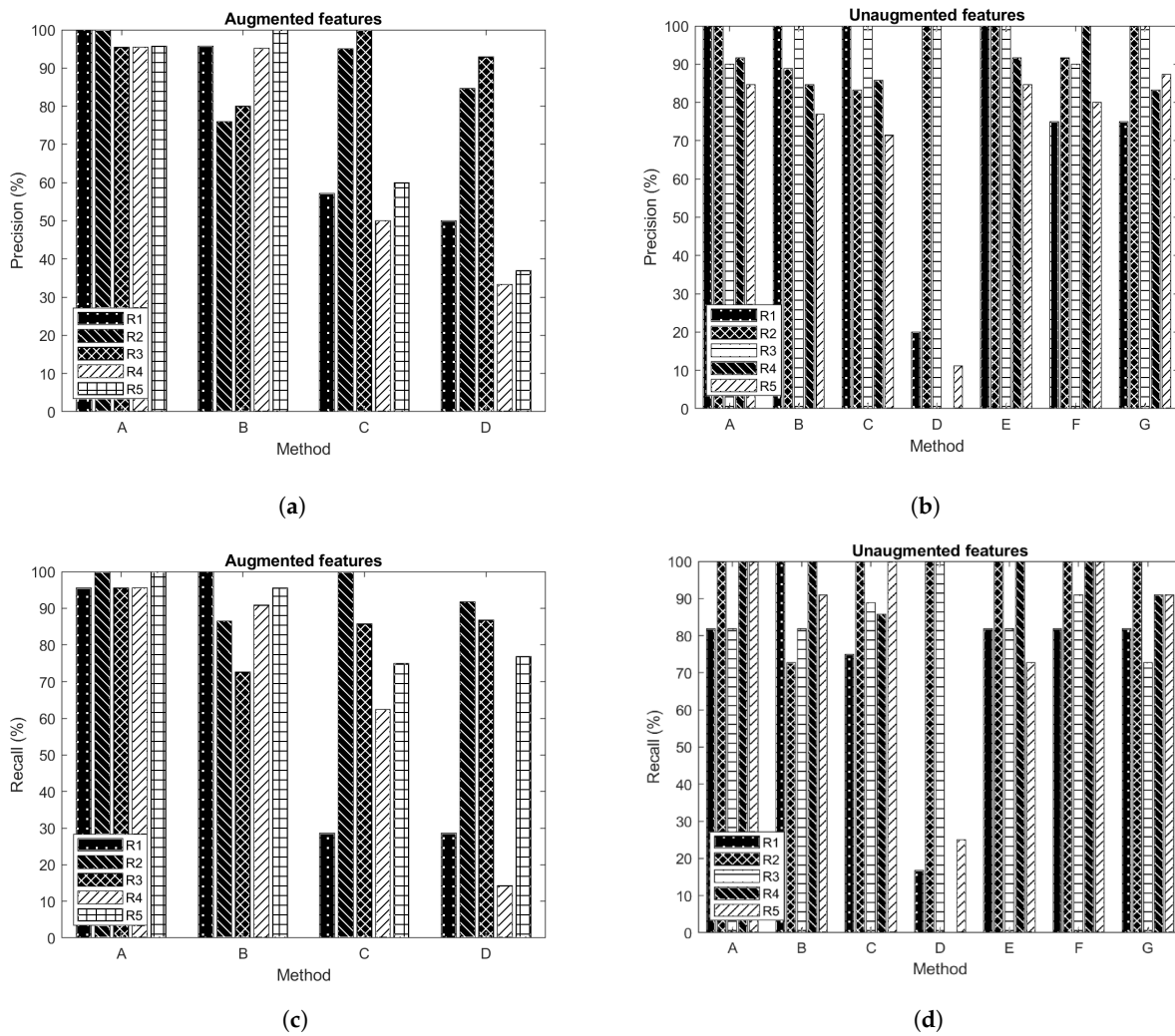


Figure 7. Room identification precision (a,b) and recall (c,d) results for augmented and unaugmented features respectively.

Table 7a–d present the confusion matrices for room identification based on the Morlet wavelet for the different feature set formulations mentioned in this study based on the Morlet wavelet. The best accuracy attained was 94.55% attained by the single RSSI Vector [34] using the CNN. Our proposed augmented CWT achieved 89.1%, which is the same as that attained by the unaugmented CWT. The RSSI feature set [34] is not an augmented feature set, yet it performed better than the augmented feature sets, implying that augmentation is not very effective when the Morlet wavelet is used for feature set generation. The augmented CWT-PSD using ANN also exhibited the same accuracy of 58.8% with the unaugmented PSD-CWT using ANN. This still shows the ineffectiveness of using augmentation when using the Morlet wavelet.

Table 7. Confusion matrices for room identification using the Morlet mother wavelet.

CWT without Augmentation Using CNN [13]							Proposed Augmented CWT Using CNN					
True labels	R1	11	0	0	0	0	R1	22	0	0	1	0
	R2	0	8	0	0	0	R2	0	19	5	0	1
	R3	0	0	9	0	0	R3	0	3	16	1	0
	R4	0	0	2	11	0	R4	0	0	1	20	0
	R5	0	3	0	0	10	R5	0	0	0	0	21
	R1	R2	R3	R4	R5		R1	R2	R3	R4	R5	
Predicted labels						Predicted labels						
(a)						(b)						
CWT-PSD without Augmentation Using ANN							Augmented CWT-PSD Using ANN					
True labels	R1	1	0	0	3	1	R1	4	0	0	1	3
	R2	0	8	0	0	0	R2	0	11	2	0	0
	R3	0	0	10	0	0	R3	0	1	13	0	0
	R4	0	0	0	0	2	R4	4	0	0	2	0
		R1	R2	R3	R4	R5		R1	R2	R3	R4	R5
Predicted labels						Predicted labels						
(c)						(d)						
Single RSS Vector CNN [34]												
True labels	R1	9	0	1	0	2						
	R2	1	11	0	0	0						
	R3	0	0	9	0	1						
	R4	0	0	0	11	0						
	R5	1	0	1	0	8						
	R1	R2	R3	R4	R5							
Predicted labels						Predicted labels						
(e)												

6.2. Estimation Results of Specific Reference Points

Figures 8 and 9a–d present the accuracy, precision, and recall results for the identification of specific reference points in a room. For the purpose of visualizing the results, only five points out of the fifteen reference points which formulated a finger print in each room were selected for analysis of precision and recall.

In Figure 8, the accuracy of the proposed augmented CWT is nearly the same as that of the unaugmented CWT for the Morse wavelet, at 94.93% and 94.80% respectively. However, both feature sets outperformed all the other feature sets, for both Morse and Morlet wavelets. Generally, the Morse-based feature sets performed better than the Morlet. This is significant for the unaugmented CWT-PSD using ANN.

The precision results arrived at using CNN for both the augmented and unaugmented datasets are high in Figure 9a–d. It is also clear from Figure 9a,c that the ANN-Morlet based precision and recall results are low for the augmented feature sets compared to the ANN-based, Morlet-based precision results for the unaugmented features in Figure 9b,d. We can deduce that ANN-based classification using Morlet-based feature sets does not exhibit improved performance when data augmentation is applied. For the case of this study, we can more specifically deduce that augmentation using Gaussian noise does not improve the performance of ANN-based learning models compared to CNN-based learning models.

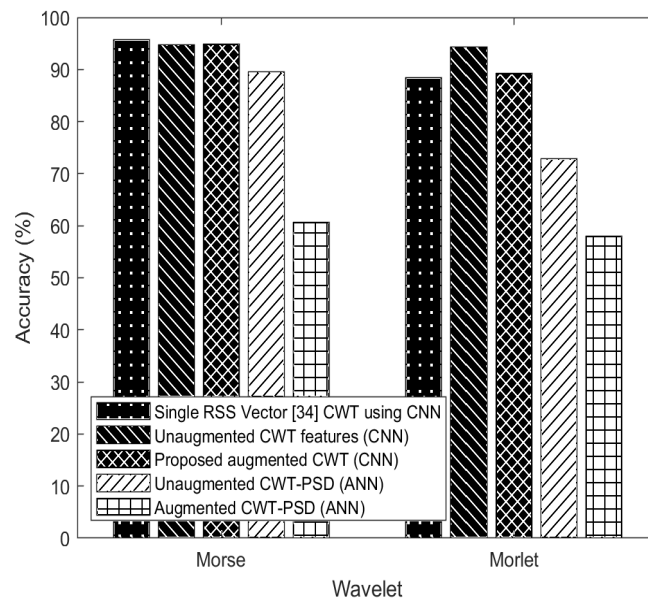


Figure 8. Accuracy results for specific reference point estimation.

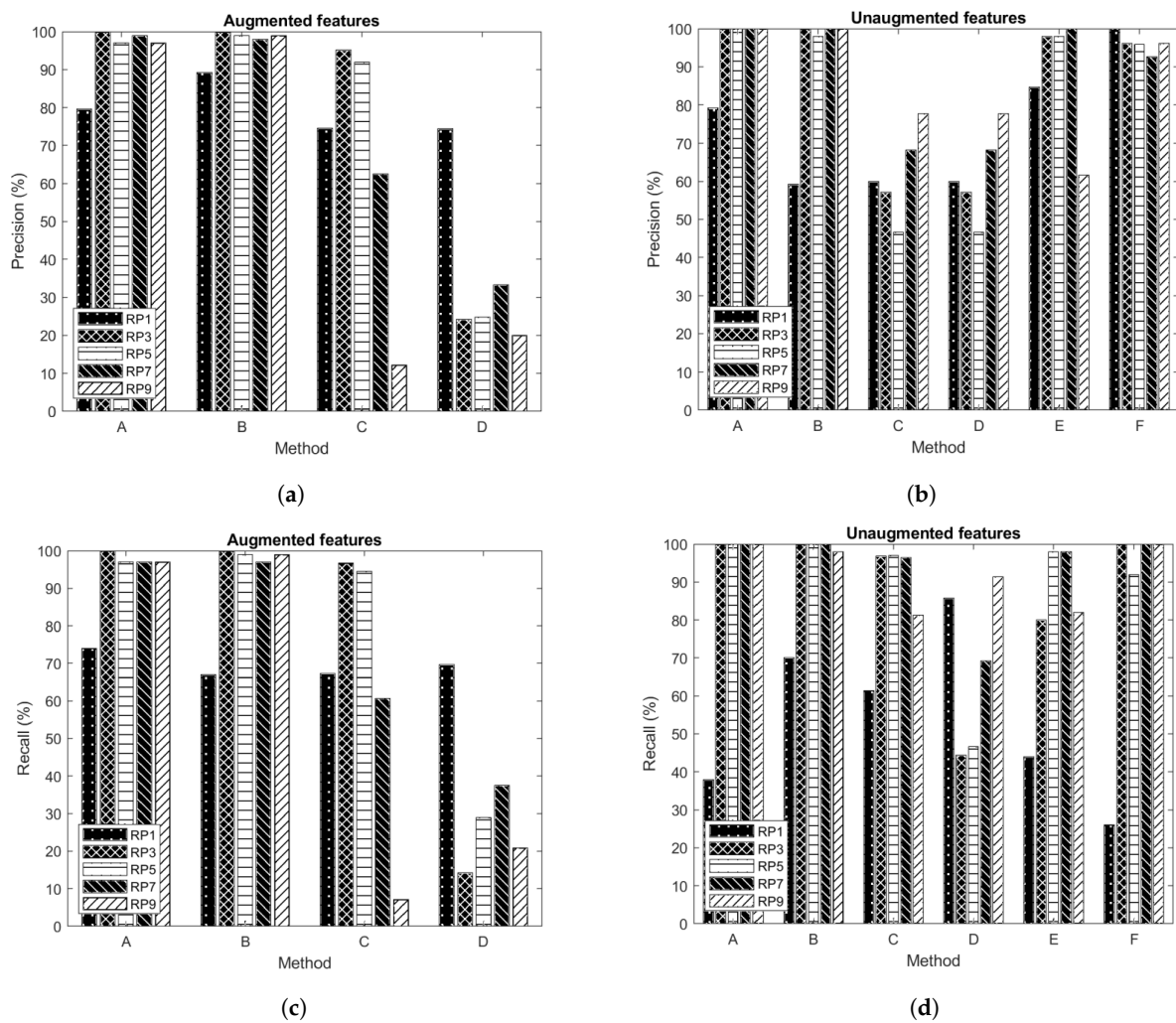


Figure 9. Specific reference point precision (a,b) and recall (c,d) results for augmented and unaugmented features respectively.

In summary, for both room prediction and specific reference point estimation, the pattern observed was that there existed better accuracy, precision, and recall performance for Morse over Morlet for ANN, and they were nearly the same for CNN. This result is backed up by the fact that technically, the Morse wavelet belongs to a two-parameter family of wavelets which generates distinct joint-time frequency data space compared to the Morlet, which is a single-parameter family of wavelets. The distinct joint-time frequency numerical data are what composes the CWT-PSD, which is used as input to the ANN model. The joint-time frequency graph which is referred to as the CWT image in this paper is used as input to the CNN model. The close performances of Morse and Morlet when using the CNN are attributed to the similarity of CWT graphical image extracts.

6.3. Prediction Error Analysis

The localization prediction error is categorized as: (1) poor recognition pointing to the same room and (2) poor recognition pointing to a different room. Poor recognition pointing to the same room means that the learning model predicted a different reference point, but in the same room wherein the actual true reference point exists, whereas poor recognition pointing to a different room means that the predicted reference point is in a different room to where the true reference point exists.

For the proposed augmented CWT Morse-based feature set, the poor recognition pointing to a different room accounts for 4.7% of the total number of dataset instances, while the poor recognition pointing to the same room instances account for 0.2% of the total number of dataset instances. The summarized analyses for the entire feature set formulations and deep learning models are presented in Figure 10a,b.

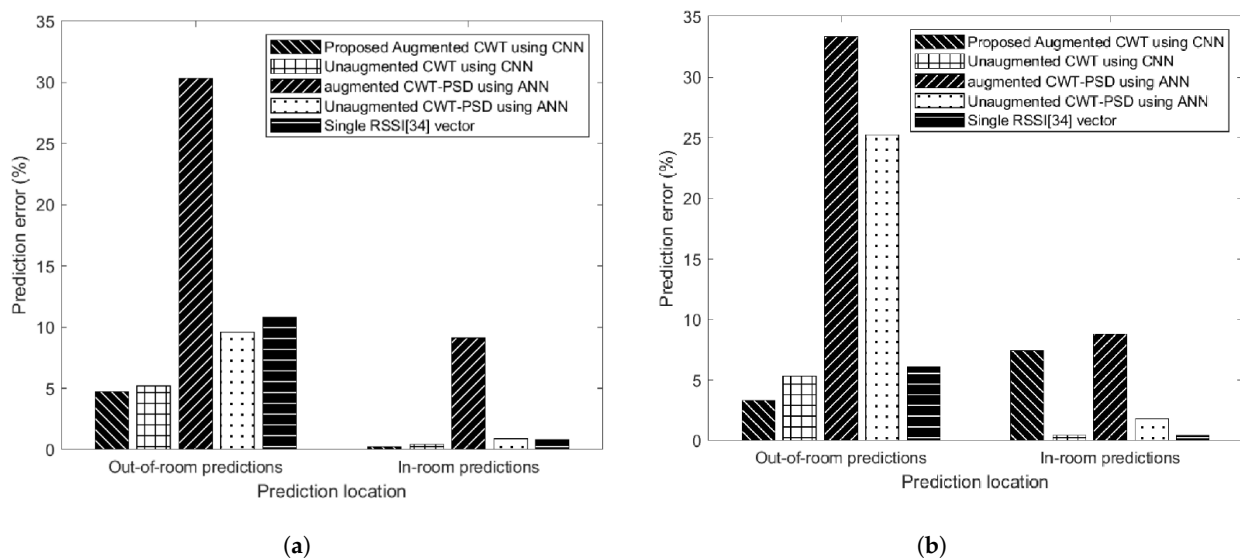


Figure 10. (a) Morse and (b) Morlet-based prediction error results.

In Figure 10a, the proposed CWT using CNN generates the least error for out-of-room predictions. This result is for the Morse-based feature set. This implies that the number of predicted reference positions which are different from the actual reference positions are the least compared the the other feature sets. The augmented CWT-PST using CNN offered the highest number of wrongly predicted positions. This is mainly because the larger the size of the dataset, the greater the number of errors, yet this factor may not affect the accuracy. However, in the in-room predictions, the unaugmented CWT-PSD using CNN offers the least prediction error compared to all the other feature sets. Its prediction error performance appears better than that of the proposed augmented CWT features. However, although the number of wrongly predicted positions for the proposed augmented CWT is

6.4. Localization Error Analysis

We present localization error results based on a real-time positioning system developed for the study and an analysis of the confusion matrices.

Our real-time system correctly recognized the room in which the human subject was holding the phone 29 out of 35 times. The specific reference points were correctly recognized 41 out of 50 times. Since it is not very representative to use location distance error for room identification, we have put effort into using location error for evaluating the specific reference point estimation.

In Hsieh et al., a localization error mathematical model was formulated based on prediction error. In this study, we formulated a Cartesian coordinate system shown in Figure 11 for the floor plan in Figure 5 in order to compute the localization error. Basically, we computed the Euclidean distance between two reference points. Once the predicted reference point is the same as the actual reference point, the localization error is zero; if not, then the localization error is the Euclidean distance between the actual and the predicted reference point. The cumulative distribution functions (CDF) of the localization error exhibited for both the Morse and Morlet-based feature sets' errors are presented in Figure 12a,b. The CDF results begin showing convergence at about 3 m, indicating a general localization accuracy realized by the CWT features. However, faster convergence was exhibited by the augmented feature sets, and this occurred close to 2 m. Conclusively, our proposed approach provides improved location accuracy compared to a sizable number of existing indoor positioning solutions, some of which were used for validating this study.

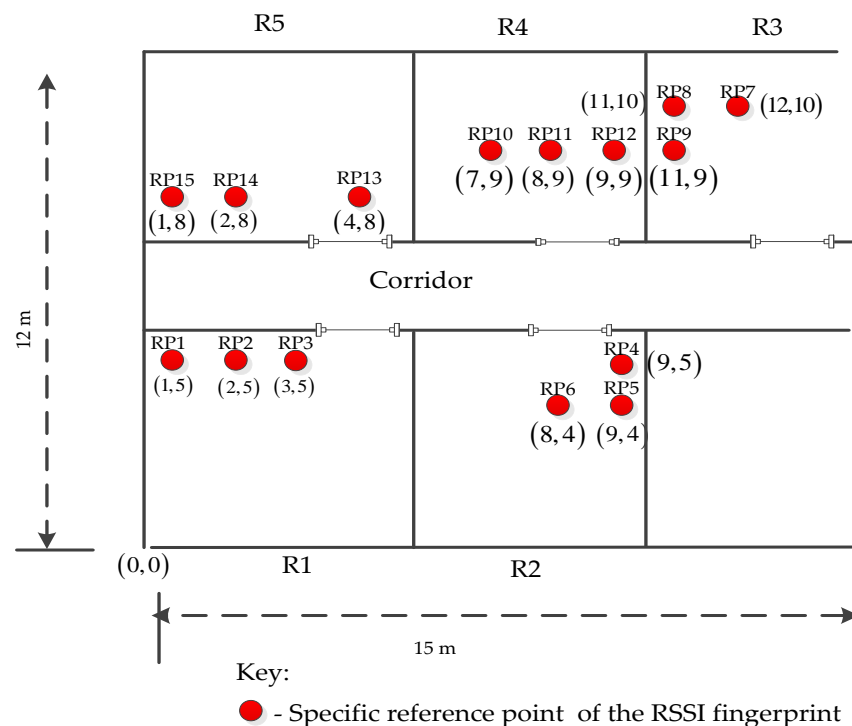


Figure 11. A location error determination system using a Cartesian coordinate graph.

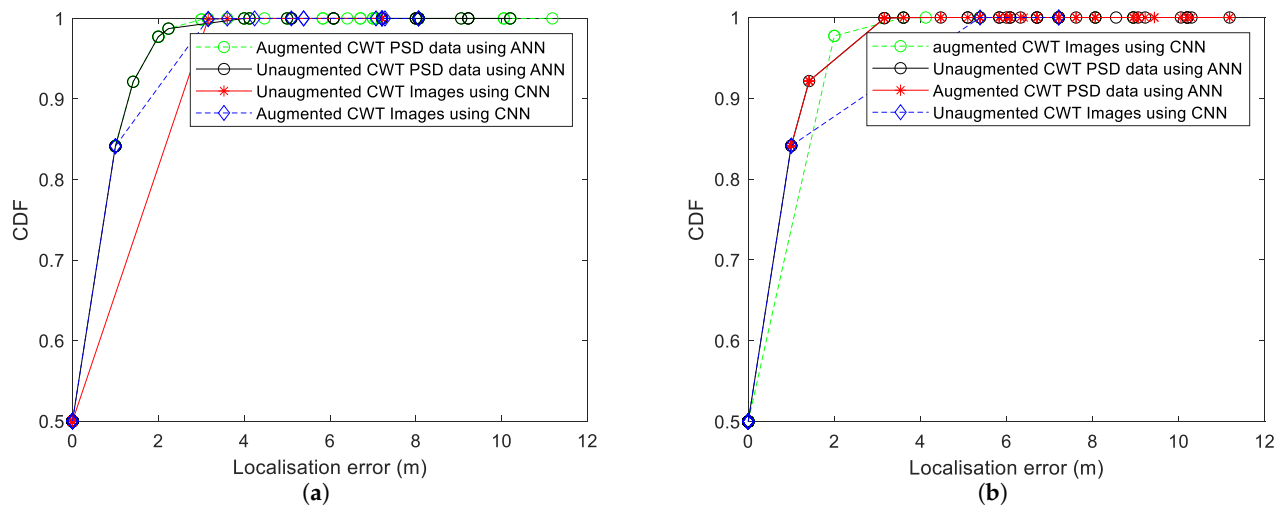


Figure 12. CDF of localization error based on (a) Morse-based feature sets and (b) Morlet-based feature sets.

7. Conclusions and Future Work

In this study, a new positioning solution for indoor environments was proposed; two deep learning-based algorithms were developed using WiFi RSSI data. The first algorithm uses images generated from received signal strength indicators' (RSSIs) continuous wavelet transforms with white Gaussian noise, and then trains a convolutional neural network. The second algorithm uses power spectral density data extracted from the CWT images, which is processed and then used to train an artificial neural network (ANN) to determine the position of an object. Position accuracies of 97.3% and 70.6% for CNN and ANN for room identification, and 94.93% and 60.6% for specific reference points were obtained. Due to the fact that CNN's performance surpasses that of ANN, it is considered as the best algorithm to use in indoor location with Wi-Fi fingerprints in this study, because it achieved better results than the other models did.

In future, we look forward to applying other data augmentation techniques to test whether they can significantly improve on the Morlet-based feature sets when ANN is applied. We also intend to further experiments the Morse wavelet abilities by using various dilating values to generate feature sets for indoor fingerprint data. Additionally, we intend to experiment with the proposed approach on inertial sensor motion (IMU) data for human subject tracking in indoor environments.

Author Contributions: Writing—original draft preparation, P.S. and O.S.E.; writing—review and editing, O.S.E., T.J.O., and D.S.H. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by Mak-Sida Project 381 at Makerere University, Uganda. This research received no other external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kawaji, H.; Hatada, K.; Yamasaki, T.; Aizawa, K. Image-based indoor positioning system: Fast image matching using omnidirectional panoramic images. In Proceedings of the 1st ACM International Workshop on Multimodal Pervasive Video Analysis, Firenze, Italy, 29 October 2010; pp. 1–4.

2. Liang, J.Z.; Corso, N.; Turner, E.; Zakhor, A. Image-based positioning of mobile devices in indoor environments. In *Multimodal Location Estimation of Videos and Images*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 85–99.
3. Lu, G.; Kambhampettu, C. Image-based indoor localization system based on 3d sfm model. In *Intelligent Robots and Computer Vision XXXI: Algorithms and Techniques*; International Society for Optics and Photonics: San Francisco, CA, USA, 2014; Volume 9025, p. 90250H.
4. Walch, F.; Hazirbas, C.; Leal-Taixe, L.; Sattler, T.; Hilsenbeck, S.; Cremers, D. Image-based localization using lstms for structured feature correlation. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 627–637.
5. Paul, A.S.; Wan, E.A. RSSI-based indoor localization and tracking using sigma-point Kalman smoothers. *IEEE J. Sel. Top. Signal Process.* **2009**, *3*, 860–873. [[CrossRef](#)]
6. Ahn, H.S.; Yu, W. Environmental-adaptive RSSI-based indoor localization. *IEEE Trans. Autom. Sci. Eng.* **2009**, *6*, 626–633.
7. Wu, K.; Xiao, J.; Yi, Y.; Chen, D.; Luo, X.; Ni, L.M. CSI-based indoor localization. *IEEE Trans. Parallel Distrib. Syst.* **2012**, *24*, 1300–1309. [[CrossRef](#)]
8. Bacciu, D.; Gallicchio, C.; Micheli, A.; Chessa, S.; Barsocchi, P. Predicting user movements in heterogeneous indoor environments by reservoir computing. In Proceedings of the IJCAI Workshop on Space, Time and Ambient Intelligence (STAMI), Barcellona, Spain, 16 July 2011; pp. 1–6.
9. Wang, Y.; Ye, Q.; Cheng, J.; Wang, L. RSSI-based bluetooth indoor localization. In Proceedings of the 2015 11th International Conference on Mobile Ad-Hoc and Sensor Networks (MSN), Shenzhen, China, 16–18 December 2015; pp. 165–171.
10. Shao, W.; Luo, H.; Zhao, F.; Ma, Y.; Zhao, Z.; Crivello, A. Indoor positioning based on fingerprint-image and deep learning. *IEEE Access* **2018**, *6*, 74699–74712. [[CrossRef](#)]
11. Xu, H.; Yang, Z.; Zhou, Z.; Shanguan, L.; Yi, K.; Liu, Y. Indoor localization via multi-modal sensing on smartphones. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Heidelberg, Germany, 12–16 September 2016; pp. 208–219.
12. Marck, J.W.; Mohamoud, A.; vd Houwen, E.; van Heijster, R. Indoor radar SLAM A radar application for vision and GPS denied environments. In Proceedings of the 2013 European Radar Conference, Nuremberg, Germany, 9–11 October 2013; pp. 471–474.
13. Soro, B.; Lee, C. Joint Time-Frequency RSSI Features for Convolutional Neural Network-Based Indoor Fingerprinting Localization. *IEEE Access* **2019**, *7*, 104892–104899. [[CrossRef](#)]
14. Steven Eyobu, O.; Han, D.S. Feature representation and data augmentation for human activity classification based on wearable IMU sensor data using a deep LSTM neural network. *Sensors* **2018**, *18*, 2892. [[CrossRef](#)] [[PubMed](#)]
15. Zhang, G.; Wang, P.; Chen, H.; Zhang, L. Wireless Indoor Localization Using Convolutional Neural Network and Gaussian Process Regression. *Sensors* **2019**, *19*, 2508. [[CrossRef](#)] [[PubMed](#)]
16. Dou, F.; Lu, J.; Wang, Z.; Xiao, X.; Bi, J.; Huang, C.H. Top-Down Indoor Localization with Wi-Fi Fingerprints Using Deep Q-Network. In Proceedings of the 2018 IEEE 15th International Conference on Mobile Ad Hoc and Sensor Systems (MASS), Chengdu, China, 9–12 October 2018; pp. 166–174.
17. Njima, W.; Ahriz, I.; Zayani, R.; Terre, M.; Bouallegue, R. Deep CNN for Indoor Localization in IoT-Sensor Systems. *Sensors* **2019**, *19*, 3127. [[CrossRef](#)] [[PubMed](#)]
18. Ibrahim, M.; Torki, M.; ElNainay, M. CNN based indoor localization using RSS time-series. In Proceedings of the 2018 IEEE Symposium on Computers and Communications (ISCC), Natal, Brazil, 25–28 June 2018; pp. 01044–01049.
19. Hsieh, C.H.; Chen, J.Y.; Nien, B.H. Deep learning-based indoor localization using received signal strength and channel state information. *IEEE Access* **2019**, *7*, 33256–33267. [[CrossRef](#)]
20. Abbas, M.; Elhamshary, M.; Rizk, H.; Torki, M.; Youssef, M. WiDeep: WiFi-based accurate and robust indoor localization system using deep learning. In Proceedings of the 2019 IEEE International Conference on Pervasive Computing and Communications, Kyoto, Japan, 11–15 March 2019; pp. 1–10.
21. Chen, Z.; Zou, H.; Jiang, H.; Zhu, Q.; Soh, Y.C.; Xie, L. Fusion of WiFi, smartphone sensors and landmarks using the Kalman filter for indoor localization. *Sensors* **2015**, *15*, 715–732. [[CrossRef](#)] [[PubMed](#)]
22. Kumar, A.K.T.R.; Schäufele, B.; Becker, D.; Sawade, O.; Radusch, I. Indoor localization of vehicles using deep learning. In Proceedings of the 2016 IEEE 17th International Symposium on a World of Wireless, Mobile And Multimedia Networks (WoWMoM), Coimbra, Portugal, 21–24 June 2016; pp. 1–6.
23. Wang, X.; Gao, L.; Mao, S.; Pandey, S. CSI-based fingerprinting for indoor localization: A deep learning approach. *IEEE Trans. Veh. Technol.* **2016**, *66*, 763–776. [[CrossRef](#)]
24. Li, N.; Chen, J.; Yuan, Y.; Tian, X.; Han, Y.; Xia, M. A Wi-Fi indoor localization strategy using particle swarm optimization based artificial neural networks. *Int. J. Distrib. Sens. Netw.* **2016**, *12*, 4583147. [[CrossRef](#)]
25. Wang, X.; Liu, X.; Wang, Z.; Li, R.; Wu, Y. SVM+ KF Target Tracking Strategy Using the Signal Strength in Wireless Sensor Networks. *Sensors* **2020**, *20*, 3832. [[CrossRef](#)] [[PubMed](#)]
26. Pathirana, P.N.; Bulusu, N.; Savkin, A.V.; Jha, S. Node localization using mobile robots in delay-tolerant sensor networks. *IEEE Trans. Mob. Comput.* **2005**, *4*, 285–296. [[CrossRef](#)]
27. Cui, Z.; Chen, W.; Chen, Y. Multi-scale convolutional neural networks for time series classification. *arXiv* **2016**, arXiv:1603.06995.
28. Nowicki, M.; Wietrzykowski, J. Low-effort place recognition with WiFi fingerprints using deep learning. In *International Conference Automation*; Springer: Warsaw, Poland, 2017; pp. 575–584.

29. Conraria, L.A.; Soares, M.J. The Continuous Wavelet Transform: A Primer. No. 23/2010 NIPE-Universidade do Minho. 2010. Available online: <https://ideas.repec.org/p/nip/nipewp/16-2011.html> (accessed on 17 November 2020).
30. Lilly, J.M.; Olhede, S.C. Generalized Morse wavelets as a superfamily of analytic wavelets. *IEEE Trans. Signal Process.* **2012**, *60*, 6036–6041. [[CrossRef](#)]
31. Shlens, J. A tutorial on principal component analysis. *arXiv* **2014**, arXiv:1404.1100.
32. Mudrova, M.; Procházka, A. Principal component analysis in image processing. In Proceedings of the 213th Annual MATLAB Technical Computing Conference, Prague, Czech Republic, 2005.
33. O’Shea, K.; Nash, R. An introduction to convolutional neural networks. *arXiv* **2015**, arXiv:1511.08458.
34. Jang, J.W.; Hong, S.N. Indoor localization with wifi fingerprinting using convolutional neural network. In Proceedings of the 2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN), Prague, Czech Republic, 3–6 July 2018; pp. 753–758.
35. Frid-Adar, M.; Diamant, I.; Klang, E.; Amitai, M.; Goldberger, J.; Greenspan, H. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing* **2018**, *321*, 321–331. [[CrossRef](#)]