Journal of Theoretical and Applied Information Technology
30th September 2017. Vol.95. No.18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

MAJORITY VOTING OF ENSEMBLE CLASSIFIERS TO IMPROVE SHORELINE EXTRACTION OF MEDIUM RESOLUTION SATELLITE IMAGES

[1]SYAIFULNIZAM ABD MANAF, [1]NORWATI MUSTAPHA, [1]MD NASIR SULAIMAN, [1]NOR AZURA HUSIN, [1]M. N. SHAH ZAINUDDIN, [2]HELMI ZULHAIDI MOHD SHAFRI

[1]Intelligent Computing Research Group, Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Malaysia

[2]Geospatial Information Science Research Centre, Faculty of Engineering, Universiti Putra Malaysia, Malaysia

E-mail: [1]nizamkpt2020@gmail.com

ABSTRACT

Coastal zones are constantly exposed to changes caused by natural processes, anthropogenic activities or both, which can precariously alter the coastal landscapes of many countries. Thus, monitoring of coastal zones is needed to provide important information about current conditions of a country's coastal areas by examining changes that are taking place. In this respect, such monitoring can be carried out by traditional ground survey, airborne aerial photo, or remote sensing. However, the former is more effective and efficient as it can extract vital boundary information from satellite images using appropriate image analysis. Nonetheless, shoreline extraction has a number of challenges, and many methods have been proposed to improve such extraction, such as the use of machine learning methods. Thus, this study was carried out to determine the most effective ensemble voting classifier based on two different types of classifiers, comprising 11 single classifiers and 4 ensemble classifiers. Performance criteria of the classifiers were based on the overall accuracy, training time, and testing time. The analysis of the experimental data revealed several interesting results. First, for the combination of single and ensemble classifiers, ensemble classifiers with majority voting of Random Forest and Support Vector Machine RBF kernel were the most effective classifiers, attaining high overall accuracy. Second, for the combination of two single classifiers, Multilayer Perceptron and k-Nearest Neighbor attained high overall accuracy, rendering them as the most effective classifiers in this category of classifiers. Third, there were trade-offs between performance measures, as increased overall accuracy was accompanied by longer training and testing time.

in the performance of such classifiers as both of voting-based ensemble classifiers increased significantly.

Keywords: *Shoreline Extraction, Image Classification, Satellite Images, Majority Voting, Ensemble Classifier*

1. INTRODUCTION

Coastal zones are constantly exposed to changes caused by natural processes and anthropogenic activities or both. Such shoreline changes are largely due to highly dynamic natural processes caused by a host of factors, such as tides, winds, waves, water currents, sediments, and oceanic temperatures, among others. In addition, anthropogenic or humans' activities arising from high concentrations of populations along coastal areas have significantly contributed to the degradation of shorelines, despite the many benefits accrued from socio-economic activities of such areas[1]. Inevitably, these changes are continually reshaping coastal areas of many countries on a massive and unpredictable scales [2]. Thus, coastal zones monitoring is needed to provide important information about the current condition of a country's coastal areas by examining changes that are taking place at its border, which is referred to as a coastline or a shoreline. According to [3], the terms 'shoreline' and 'coastline' are commonly used by the coastal community and remote sensing fraternity, respectively. Essentially, a shoreline is defined as a physical line that acts as an interface that physically separates land and water, thus creating a boundary between them [4][5]. Hence,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

shoreline extraction provides the boundary information of land and water, which helps monitor erosions or accretions of coastal zones [6]. In this regard, such monitoring can be performed faster and more accurately by using remote sensing rather than by using traditional field survey. Clearly, with remote sensing, satellite images of large, inaccessible geographical areas can be captured and analyzed more effectively and efficiently, as opposed to using traditional field survey or airborne aerial photography, which are relatively laborious and imprecise. Moreover, the former method can be used to extract vital boundary information from satellite images using appropriate image analysis.

Additionally, the extraction of selected features of shoreline can be readily performed on available data to reveal important spatial and temporal characteristics. Effectively, this extraction helps control and manage historical records of geographical changes that have been taking place, thus making the prediction of future changes of coastlines more precisely and reliable. Irrespective of the nature of coastal areas, acquiring information of a shoreline entails a shoreline indicator, which can accurately represent the true position of a particular shoreline of interest [3]. In general, the shoreline indicators can be divided into three main categories: (i) an indicator that is based on visually detectable features, (ii) an indicator that is based on a specific tidal datum, and (iii) an indicator that is based on features that may be imperceptible to the human's eyes[4]. In this study, the researchers used the third category of indicators as the shoreline indicator.

The main aim of this study is to develop and evaluate the most effective and efficient ensemble machine learning technique for the extraction of shoreline of the Langkawi Island (which is located at the North West coast of Peninsular Malaysia) based on Landsat TM satellite images. For this study, the researchers used pixel-based approaches to classify land-water classes using majority voting of ensemble classifiers based on two different combinations of 15 heterogeneous machine learning classifiers. In particular, the majority voting ensemble classifiers were employed to perform the classification process. In fact, the techniques used in this study were the improved versions of techniques used in a study of [7], who compared and evaluated the performances of single classifiers on a same dataset.

To aid discussion, this paper is structured as follows: Section II discusses the related works, Section III details the materials and methods used, Section IV reports the experimental results of the proposed ensemble machine learning techniques, and Section V highlights the discussion and conclusion of the paper.

2. RELATED WORKS

To date, many techniques to extract shorelines from optical multispectral satellite images have been proposed and evaluated. In general, such techniques to extract shorelines from satellite images can be divided into three main techniques, namely the image processing techniques, image classification techniques, and satellite-derived indices. The image processing techniques are based on segmentation[8], edge detection[9], wavelet[10], and density slicing[11], histogram threshold[12]. The satellite image classification techniques are basically machine learning (ML) techniques, which can be further divided into two types of classifications: (i) supervised classification, such as Maximum Likelihood, Mahalanobis Distance, Minimum Distance, Neural Network and Support Vector Machines[13][14][15][16] and (ii) unsupervised classification, such as ISODATA[14] and fuzzy c-means[17]. The satellite-derived indices techniques include band rationing[18][19][12] and normalized difference water index (NDWI)[16] [20][21].

For satellite image classification techniques based on supervised classification, Lipakis et al. [22] employed several machine-learning algorithms to extract both spectral and spatial information of shorelines. However, their study did not specify the actual type of machine-learning algorithms used for such extraction of shoreline information. In contrast, Masria et al. [12] used SVM to detect temporal changes of coastal areas around Rosetta Promontory, Egypt, which recorded a range of overall accuracies, ranging 97% to 100%.

Interestingly, Sekovski et al [14] used both unsupervised and supervised classification methods in their study to compare the impacts of such different methods on the detection of bodies of water. For the former, they employed ISODATA; for the latter, they used Maximum Likelihood, Mahalanobis Distance, Parallelepiped and Minimum Distance. On the other hand, Rokni et al. [13] proposed an approach based on the integration of image fusion and image classification techniques for surface water change detection. Their study

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

revealed several interesting results, in particular Gram-Schmidt-ANN and Gram Schmidt-SVM approaches were observed to achieve relatively higher accuracies that those of other three image fusion techniques that were based on the combinations of Maximum Likelihood, ANN, and SVM.

Recently, Choung and Jo [16] proposed a machine learning-based method to extract shorelines using SVM and to compare its accuracy with that of a water index-based method. Remarkably, the former method was observed to be more accurate than that of the latter method in extracting shorelines of coastal zones consisting of numerous irregular shapes and color shades, among others. Even though there were some studies of land cover mapping [23][24], which have been reported in the literature, research on shoreline extraction using ensemble methods based on image classification techniques is seriously lacking. Hence, this study was carried out with the main aim to investigate the suitability of several majority voting ensemble classifiers in extracting shorelines of a particular study area, namely the Langkawi Island.

3. MATERIALS AND METHODS

The methods used to extract shorelines from satellite images consisted of four phases. The four phases were pre-processing, satellite image classification, accuracy assessment, and post-processing, as depicted in Figure 1. Prior to these extraction phases, a study area was chosen and data acquisition was carried out before performing data pre-processing.

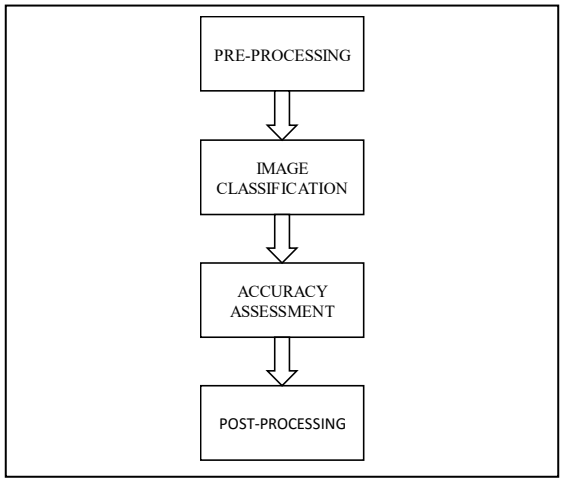

Figure 1: The Four Phases of Methodology of This Study

3.1 Study Area

Langkawi Island is located at the North West coast of Peninsular Malaysia, as shown in Figure 2. Geographically, this island is located at 6º 15'N and 6º 29'N latitude and 99º 37'E and 99º 57'E longitude, covering a total area of about 478.48 km^2. This exotic island consists of many small islands dotting around a main island, of which the latter was chosen as the study area.

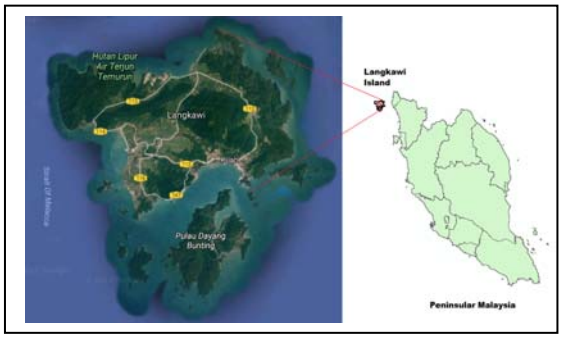

Figure 2: The Study Area of the Research

In fact, Langkawi Island is one of the districts of the state of Kedah. The main landmass of this island consists of six sub-districts, such as Ayer Hangat, Bohor, Kedawang, Kuah, Padang Matsirat, and Ulu Melaka. As a pristine island, Langkawi Island is famous for its picturesque mountainous landscapes, glittering white-sandy beaches, and long stretch of hills and paddy fields [25].

After gaining a duty-free status in 1987, Langkawi has become one of the popular tourist destinations, enabling it to gain a profitable stream of revenues from the tourism industry [26]. The establishment of Langkawi Development Board (LADA) in 1990 further boosted the economic developments of Langkawi. LADA plays an important role in planning and implementing development projects in this island. Currently, LADA and other local authorities and relevant government agencies, such as the Langkawi Municipal Council, the Langkawi Land and Mineral Office, and the Tourism Malaysia, are working in concert to ensure the development projects will proceed smoothly without adversely affecting ecological and geological aspects of the island [25].

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.2 Data Acquisition

The data used in this research consisted of two scenes of multispectral Landsat-5 Thematic Mapper (TM) data. Two scenes, instead of one, were used to ensure they would sufficiently cover the whole study area. These two scenes were acquired on two different dates, as shown in Table 1. As shown, the first scene was acquired on 24 December 2010 at 11.29 am (local time), while the second scene was acquired on 5 August 2011 at 11.28 am (local time).

Table 1: Description of Satellite Image Data

Sensor	Acquired Date	Acquired Time (GMT)	Path/Row
Landsat TM	24/12/2010	3:29:49	129/56
Landsat TM	05/08/2011	3:28:41	128/56

Over the years, Landsat TM satellite orbited the Earth more than 150,000 times, transmitting over 2.5 million images of land surface conditions around the world. From 1984 to 2013, this satellite provided multispectral images of the Earth's surface at an altitude of 705 km with 8-bit radiometric resolution, thus making it the longest operating satellite sensor[12]. The repeat cycle was within 16 days with 185 km swath width. Landsat TM satellite had 7 spectral bands, including a thermal band, such as blue: 0.45–0.52 µm; green: 0.52–0.60 µm; red: 0.63–0.69 µm; NIR (near infrared): 0.76–0.90; SWIR 1: 1.55-1.75 µm; Thermal: 10.40-12.50; and SWIR 2: 2.05-2.35 µm. All these bands had 30m spatial resolution, except the Thermal band that had 120m spatial resolution.

3.3 Pre-processing

Pre-processing of satellite images is an important step for attaining high classification accuracy and easing computational complexity. In this pre-processing phase, 5 processes were performed, namely radiometric correction, atmospheric correction, mosaicking, geometric correction, and subset of the study area, as shown in Figure 3. The main aim of the pre-processing phase was to clean satellite images from errors caused by satellite sensors, such as atmospheric, radiometric, and geometric errors.

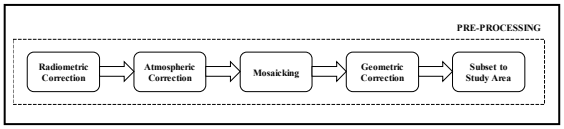

Figure 3: Pre-processing Phase

For this study, such errors were introduced when the Landsat TM sensors captured reflected solar energy and converted data to radiance values before rescaling them to digital numbers (DNs), ranging between 0 and 255. Hence, radiometric correction was performed using bias and gain values to convert DNs to radiance (L_λ) values based on the scene being captured [20], as expressed by Equation 1.

$$L_\lambda = \text{Gain} * \text{Pixel value} + \text{Offset} \quad (1)$$

Then, atmospheric correction was performed by converting the radiance data to Top of Atmosphere (ToA) reflectance[27] that helped remove the atmospheric scattering effects of the image data. This process used the minimum value of a band, which represented the background signature of the band.

At the end of the pre-processing stage, geometric corrections using ground control points (GCP) were used to transform the data from global coordinate system WGS84 to Rectified Skew Orthomorphic (RSO) local projection such as to precisely align all data on the same map[28]. In this stage, image registration using image-to-image geometric correction was performed. The 30 GCPs were located between the geo-referenced image coordinates and the new image coordinates to cover the whole study area. The RMS value for all GCPs was 0.457, which was far below 1.00. With such small RMS value, registration errors or distortions of the image that could increase the overall error count was effectively avoided.

In addition, mosaicking and subset processes were two processes carried out to combine two or more images and to trim the combined images such that they would fit neatly with the study area. Specifically, ENVI 5.3 was used for pre-processing tasks at this stage. Once thoroughly cleaned and perfectly fitted, such images in the reflectance format would be used in the classification phase.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.4 Satellite Image Classification

The main aim of the satellite image classification phase was to identify land and water classes that helped determine the boundaries that separated these two regions. In this phase, supervised image classification methods were used to classify land and water classes of satellite images as shown in Figure 4.

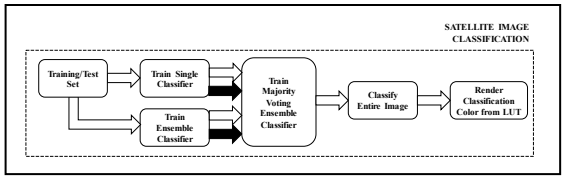

Figure 4: Satellite Image Classification Phase

For this study, only one training set was created in the form of polygons, consisting of 260 and 65 generated polygons for land and water classes respectively, as depicted in Figure 5. The recommended training sample size for each class should be higher than 10 - 30 times the number of bands[29]. They were selected by image interpretation of image regions captured by the satellite, which were equivalent region of the fields. The same training set was also used as a testing set in the cross validation scheme to determine the classification accuracy. Actually, the training set was created to build the model, while the testing set was produced to measure its performance. Jeffries-Matusita distance and Transformed Divergence were used[30] to ensure separability of the training and testing set, yielding a separability index of 1.97, which was close to 2.0 that represents perfect separability.

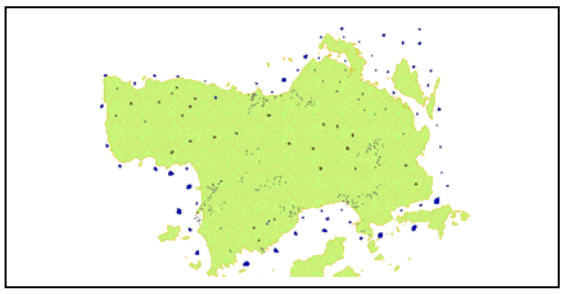

Figure 5: Training and Testing Set

Clearly, given the high number of techniques available, choosing the most effective ML classifier for this domain problem was a challenging task. As such, for this study, the researchers used 11 single ML classifiers and 4 ensemble ML classifiers. The single ML classifiers used in this study were Decision Tree (DT), Naïve Bayes (NB), k-Neareast Neighbour (kNN), Linear Discriminat Analysis (LDA), Quadratic Discriminant Analysis (QDA), Logistic Regression (LR), SGD Classifier, SVM-linear (SVM-L), SVM-RBF (SVM-R), SVM-polynomial (SVM-P), and Multi-Layer Perceptron Artificial Neural Network (MLP). The ensemble ML classifiers selected for this study were AdaBoost (ADB), Gradient Boosting (GDB), Random Forest (RF), and Extra Trees (ET). After the model had learned part of image regions from the training samples, the whole image regions were then classified. In addition, a lookup table (LUT) consisting of all color classes was used to render the classified image with appropriate colors. In this study, brown and blue colors were used to represent land and water, respectively.

In theory, ensemble methods are deemed highly effective, because aggregated decision of multiple classifiers is far more superior to a single decision of a single classifier. Fundamentally, ensemble methods are models composed of multiple weaker models that are independently trained, where each model's prediction is combined with other models' predictions to provide an overall prediction[31]. In this respect, such combination entailed the researcher to carefully select the appropriate types of classifiers and the efficient ways to combine them. With this approach, a reliable model was created that helped improve the prediction and classification accuracy. In fact, different models of ensembles could be generated from the same basic algorithm based on different subsets of the data or from different algorithms based on the same dataset. Interestingly, on certain occasions, ensembles can perform better than a single model due to the diversity of base models in solving certain problems. Furthermore, according to [32][33], the combination of various classifiers can result in better accuracy compared to those of standalone classifiers. Hence, the ensemble approach based on majority voting was proposed for this research.

The majority voting is one of the simplest and most intuitive ensemble combination techniques, as shown in Figure 6. In general, the majority vote classifier is defined as a series of votes, which is represented as follows:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$C(X) = \arg\max \sum_{j=1}^{p} w_j \, I(h_j(X) = i) \quad (2)$$

For this combination scheme, a classification of an unlabeled instance is performed on a class that obtains the highest number of votes (the most frequent vote), and such a method is also known as the Plurality Vote (PV).

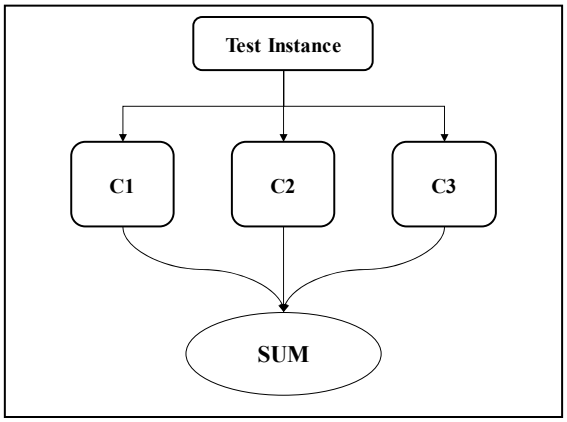

Figure 6: Majority Voting Classifier

3.5 Accuracy Assessment

In the accuracy assessment phase, the overall accuracy, training time, and testing time were used as performance indicators, as shown in Figure 7. In particular, the overall accuracy was used as the primary performance indicator, because it is widely used in the evaluation of satellite image classification methods. Whereas, the other two factors were treated as additional performance indicators in this study.

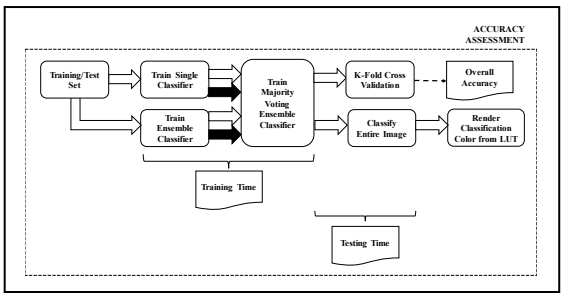

Figure 7: Accuracy Assessment Phase

For cross validation, the 10-fold cross-validation method was used to deal with overfitting and class imbalance problems. Essentially, the underlying mechanism of k-fold cross-validation helps partition an original sample randomly into k subsamples of equal size. A single subsample was used to test the model as validation data, while the remaining (k-1) subsamples were used as training data [34].

3.6 Post-processing

Finally, in the post-processing phase as shown in Figure 8, the resultant classified image was converted into GIS vector format using ENVI 5.3, which could be further processed using ArcGIS 10.3. Moreover, smoothing processes were performed on the classified images to generalize the image data consisting of smooth polygons. Then, the resultant classified image underwent a raster-to-vector conversion process to convert it into GIS vector format. Subsequently, polygon-to-line conversion process was performed on this converted image to change it into vector lines format. Finally, line smoothing was applied to the final shoreline to smooth the straight edges and angular corners of features such that curves would be seamlessly connected at their vertices.

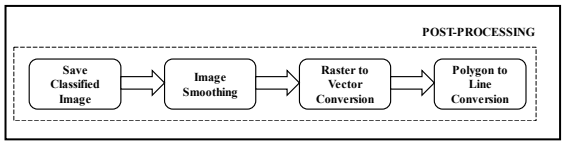

Figure 8: Post-processing Phase

4. RESULTS

Initially, the analysis of the experimental data of single classifiers was carried out on a high-performance workstation, namely Dell Precision 3620 machine, which was equipped with 3.4GHz Intel i7-6700 Quad Core Processor and 16 GB RAM, running on the Microsoft Windows 7, which was a 64-bit operating system.

In order to perform majority voting ensemble on the same dataset, the workstation's RAM was upgraded to 28GB to avoid any problems related to insufficient memory. Such upgrading was needed as 16GB memory could only support single ML classifiers but not all ensemble ML classifiers.

4.1 Experimental Setup

The study area and specific type of satellite images for this study had to be identified before performing other processes. As such, we chose Landsat TM satellite images of the whole

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Langkawi Island, which is located at the north-west coast of Peninsular Malaysia, as the study area. In the pre-processing phase, each image was cleaned by a series of processes such as atmospheric and geometric corrections to rid it of errors generated by satellite sensors. Then two additional processes such as mosaicking and image subset were performed to ensure the cleaned images were configured to the same coordinate system that helped overlay it on the same map. In the classification phase, 260 and 65 polygons were generated for the training and testing samples of the image for land and water classes, respectively. Image classification was performed on the 30m image, resulting in a classified image of land and water classes.

In the first run, we classified the geo-references and fitted satellite images with the use of 15 ML classifiers (11 single and 4 ensemble classifiers). The single ML classifiers used in this study were DT, NB, kNN, LDA, QDA, LR, SGD, SVM-L, SVM-R, SVM-P, and MLP, while the ensemble ML classifiers selected were ADB, GDB, RF, and ET. In the experiments, the performances of these 15 classifiers were assessed in terms of overall accuracy, training time, and testing time. Then, in the final run, we classified the same satellite image using all possible combinations of any two classifiers (as mentioned above). The performances of ensemble voting classifiers were also assessed in terms of overall accuracy, training time, and testing time. Later, we carefully examined the accuracy performance results to select relevant classifiers for further post-processing. In the post-processing phase, the resultant classified image was generalized using sieve and clump processes before converting it into GIS vector format using raster-to-vector conversion process. Subsequently, polygon-to-line conversion process was performed on this converted image to change it into vector lines format. Finally, line smoothing was applied to the shoreline to obtain a final extracted shoreline.

4.2 Experimental Results

Table 2 summarizes the results of satellite image classifications based on the 15 classifiers, of which 11 were single classifiers and the remaining 4 were ensemble classifiers. Clearly, the results showed MLP was the most effective method, attaining the highest overall accuracy at 99.55%. As anticipated, the other four ensemble classifiers also achieved impressive overall accuracy, placing them among the top five classifiers in terms of accuracy performance. In contrast, at 98.58% of overall accuracy, QDA was the least effective method, albeit having the fastest training and testing time. Evidently, these results reinforce the suitability of ensemble methods for classifying land and water classes. From these promising results, the researchers undertook further investigation to determine whether ensemble voting methods would be suitable for such study in this domain.

Table 2: Performance Results of Satellite Image Classification using 15 Classifiers (11 Single Classifiers and 4 Ensemble Classifiers).

ML Classifier	OA (%)	Training Time (s)	Testing Time (s)
DT	99.28	0.098	0.576
NB	98.75	0.051	0.713
kNN	99.21	0.378	8.21
LDA	99.30	0.06	0.534
QDA	98.58	0.042	0.65
LR	99.36	0.256	0.75
SGD	99.01	0.125	0.598
SVM-L	99.31	1.032	3.311
SVM-R	99.08	8.758	19.867
SVM-P	98.91	26.89	28.211
MLP	99.55	3.512	4.067
ADB	99.49	1.295	2.213
GDB	99.50	2.101	9.908
RF	99.41	0.237	1.256
ET	99.37	0.201	1.353

Table 3 summarizes the performance results of satellite image classifications of selected ensemble classifiers, which are shown in Table 4, Table 5, and Table 6. Clearly, RF+SVM-R was the most effective ensemble classifier, as evidenced by its high overall accuracy at 99.59%. Comfortably at the second place was kNN+MLP, which attained 99.57% of overall accuracy. Equally impressive were the remaining classifiers such as MLP+SVM-R, MLP+SGD, and RF+NB, which equally achieved 99.56% of overall accuracy.

Table 5 shows the results of overall accuracy of ensemble voting classifiers based on two single classifiers. The accuracy results of single ensemble voting classifiers, with each based on the combination of two different classifiers, are highlighted in this table. However, the results of the same ensemble voting classifiers were not displayed, because the voting had not resulted in any improvements in the overall accuracy. Overall, all the ensemble voting classifiers achieved high overall accuracy, with each classifier's accuracy

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

surpassing more than 90.00%, except for QDA+NB.

Table 3: Performance Results of Satellite Image Classification of Selected Ensemble Voting Classifiers

Voting ML Classifier	OA (%)	Train Time (s)	Test Time (s)
MLP + SVM-R	99.56	13.334	28.451
SGD + MLP	99.56	3.863	7.738
kNN + MLP	99.57	4.334	15.866
RF + NB	99.56	0.594	5.114
RF + SVM-R	99.59	9.638	22.731
RF + ADB	99.55	4.709	16.149

Given that MLP, as a single classifier, had achieved the highest overall accuracy, it was used as a benchmark to which other voting results were compared. Such comparison revealed that only two voting classifiers achieved higher performances beyond this benchmark, namely MLP+KNN and MLP+SVM-R, with the former and the latter attaining the highest and the second highest overall accuracy at 99.57% and 99.56%, respectively. In contrast, with 98.74% of overall accuracy, QDA+NB was the least effective ensemble voting classifier.

In addition, the remaining ensemble classifiers were evaluated to determine whether their performances would improve after voting. The evaluation showed that a few of these classifiers did improve their accuracy performances. Surprisingly, some of the classifiers did not make any improvements, such as kNN+SGD, MLP+DT, SVM-R+DT, SVM-P+DT, QDA+NB, LDA+ SGD, LDA+SVM-P, LR+SGD, MLP+SGD, and SVM-P +MLP. These results were significantly different from those of DT and LR, which only achieved 99.28% and 99.36% of the overall accuracy, respectively.

As demonstrated, some classifiers were not able to improve their accuracy after performing ensemble voting. Closer examination revealed that such classifiers were those that had been combined with LDA, namely LDA+SGD and LDA+SVM-P. Similarly, no improvements in accuracy were observed for other MLP classifiers, except MLP+kNN and MLP+SVM-R that achieved high overall accuracy.

Table 6 shows the overall accuracy results of voting of an ensemble classifier and other single classifiers. The results showed that RF+SVM-R was the most effective ML, attaining 99.59% of overall accuracy. A close second was RF+NB, which achieved 99.57% of the overall accuracy. The same results indicated that RF+LDA was the least effective ML, which managed to attain 99.16% of the overall accuracy.

Table 4 summarizes the results of voting of two ensemble MLs. Evidently, all combinations of ensemble classifiers were observed to be quite ineffective, except for the combination of ET and RF. Even for the latter combination, the improvement was not significantly high. Arguably, the lack of improvements was due to overfitting of classifiers that led to poor accuracy.

Table 4: Performance Results of Satellite Image Classification Using Ensemble Voting Classifiers of Ensemble Classifiers.

Voting ML Classifier	GDB	ADB	RF	ET
GDB	-	99.50	99.34	99.49
ADB	99.50	-	99.55	99.49
RF	99.34	99.55	-	99.42
ET	99.49	99.49	99.42	-

The best result from Table 3 was chosen to do further post-processing tasks. The classification process produced the image classification map as shown in Figure 9, which clearly highlights the differentiation between the land class (in brown) and the water class (in blue). Later, post-classification processes were performed on the image classification map to extract the shoreline from the satellite image. Figure 10 shows the extracted shoreline output after the completion of the extraction processes.

Figure 9: Image Classification Map Result

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Figure 10: Extracted Shoreline Output

5. DISCUSSIONS AND CONCLUSION

In this study, a series of experiments involving 11 single machine-learning classifiers and 4 ensemble classifiers were carried out, revealing several interesting and important results. For single classifiers, MLP was the most effective image classification method, having achieved the highest overall accuracy [4]. For ensemble classifiers using majority voting algorithm, RF+SVM-R recorded the highest overall accuracy. In addition, RF+NB, MLP+kNN, and MLP+SVM-R were among the most effective classifiers. In contrast, LDA was the least effective classifier based on the ensemble voting approach. The results also showed that the training and testing time of voting ensemble classifiers increased substantially, despite their high overall accuracy. Clearly, the overall accuracy of such classifiers decreased quite significantly when more than two classifiers were combined.

There were, however, some limitations in this study, which may preclude the generalization of the research findings. Firstly, this research only used Landsat TM data as the study area. Ideally, Landsat sensors, such as Landsat ETM+ or Landsat OLI can be utilized to provide a more robust and comprehensive data to improve the reliability of such research findings. Furthermore, such sensors can also be applied to examine other important study areas. Secondly, the resultant extraction spatial data were based on 30m satellite resolution, below which detecting any changes might be difficult. Thirdly, the majority voting of ensemble classifiers was only based on the combination of two different classifiers. Admittedly, the researchers did try to use several combinations based on three or more classifiers, but the results were disappointingly poor, showing a substantial decrease in the overall accuracy. Presumably, the maximum number of classifiers to combine that can help achieve optimum overall accuracy is limited to two only.

For future research, object-based image analysis (OBIA) can be applied to each classifier that has been implemented using the pixel-based approach. With OBIA, the combination of segmentation algorithms and ML classifiers can further improve the overall accuracy of extracted shorelines. In addition, deep learning can be performed with the use of a more powerful workstation, which should ideally have bigger RAM and high-performance GPU card. For validation assessment, ML classifier results can be further analyzed using ground truth data, such as reference shoreline provided by the local authority. As there is limitation with 30m dataset, future research can use satellite images of higher resolution to improve the accuracy of shoreline extraction.

ACKNOWLEDGEMENT:

The authors would like to thank the Ministry of Science, Technology and Innovation of Malaysia for the ScienceFund Research Grant (No: 01-01-04-SF2291); the Malaysian Remote Sensing Agency, the Department of Survey and Mapping of Malaysia, the Federal Department of Town and Country Planning Peninsular Malaysia, the Department of Agriculture Malaysia, and the United States's Geological Survey for the research data; and Universiti Putra Malaysia for the logistical assistance rendered to the researchers.

REFRENCES:

[1] N. H. M. Salleh, R. Othman, S. H. M. Idris, S. A. Halim, M. S. Shukor, I. Yussof, M. Samsudin, and M. Omar, "Development of tourism industry and its impact on Langkawi Island community," *J. Ekon. Malaysia*, vol. 48, no. 2, pp. 71–82, 2014.

[2] M. F. Mohamad, L. H. Lee, and M. K. H. Samion, "Coastal Vulnerability Assessment towards Sustainable Management of Peninsular Malaysia Coastline," *Int. J. Environ. Sci. Dev.*, vol. 5, no. 6, pp. 533–538, 2014.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[3] R. Gens, "Remote sensing of coastlines: detection, extraction and monitoring," *Int. J. Remote Sens.*, vol. 31, no. 7, pp. 1819–1836, 2010.

[4] E. H. Boak and I. L. Turner, "Shoreline Definition and Detection: A Review," *Journal of Coastal Research*, vol. 214. pp. 688–703, 2005.

[5] S. Dolan, Robert, S. Fenster, Michael, and J. Holme, "Temporal Analysis of shoreline Recession and Accretion," *J. Coast. Res.*, vol. 7, no. 3, pp. 723–744, 1991.

[6] A. Faiboona, S. Pradita, and S. Ritphringb, "Monitoring Shoreline Change Using Geo-Informatics : A Case Study at Chalatat Beach in Songkhla Province , Thailand," vol. 66, no. 0, 2011.

[7] Syaifulnizam Abd Manaf, Norwati Mustapha, Md. Nasir Sulaiman, Nor Azura Husin, and Mohd Radzi Abdul Hamid, "Artificial Neural Networks for Satellite Image Classification of Shoreline Extraction for Land and Water Classes of the North West Coast of Peninsular Malaysia," *Adv. Sci. Lett.*, vol. 4, no. 2, pp. 400–407, 2016.

[8] S. M. Semenov, N. A. Abushenko, and A. S. Chichigin, "Discrimination of Shorelines on Satellite Images from Boundary-Point and Halftone Information," *Mapp. Sci. Remote Sens.*, vol. 36, no. 4, pp. 245–255, 2016.

[9] Y. Zhang, X. Li, J. Zhang, and D. Song, "A Study on Coastline Extraction and Its Trend Based on Remote Sensing Image Data Mining," *Abstr. Appl. Anal.*, vol. 2013, no. Article ID 693194, p. 6, 2013.

[10] S. Yu, Y. Mou, D. Xu, X. You, L. Zhou, and W. Zeng, "A New Algorithm for Shoreline Extraction from Satellite Imagery with Non-separable Wavelet and Level Set Method," *Int. J. Mach. Learn. Comput.*, vol. 3, no. 1, pp. 158–163, 2013.

[11] P. S. Frazier and K. J. Page, "Water Body Detection and Delineation with Landsat TM Data," *Photogramm. Eng. Remote Sens.*, vol. 66, no. 12, pp. 1461–1467, 2000.

[12] A. Masria, K. Nadaoka, A. Negm, and M. Iskander, "Detection of Shoreline and Land Cover Changes around Rosetta Promontory, Egypt, Based on Remote Sensing Analysis," *Land*, vol. 4, no. 1, pp. 216–230, 2015.

[13] K. Rokni, A. Ahmad, K. Solaimani, and S. Hazini, "A new approach for surface water change detection: Integration of pixel level image fusion and image classification techniques," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 34, no. September, pp. 226–234, 2015.

[14] I. Sekovski, F. Stecchi, F. Mancini, and L. Del Rio, "Image classification methods applied to shoreline extraction on very high-resolution multispectral imagery," *Int. J. Remote Sens.*, vol. 35, no. 10, pp. 3556–3578, 2014.

[15] A. Yousef and K. Iftekharuddin, "Shoreline extraction from the fusion of LiDAR DEM data and aerial images using mutual information and genetic algrithms," *Proc. Int. Jt. Conf. Neural Networks*, pp. 1007–1014, 2014.

[16] Y. Choung and M. Jo, "Comparison between a Machine-learning-based Method and a Water-index-based Method for Shoreline Mapping Using a High-Resolution Satellite Image Acquired in Hwado Island , South Korea," *J. Sensors*, 2017.

[17] R. Dewi, W. Bijker, A. Stein, and M. Marfai, "Fuzzy Classification for Shoreline Change Monitoring in a Part of the Northern Coastal Area of Java, Indonesia," *Remote Sens.*, vol. 8, no. 3, p. 190, 2016.

[18] C. Lira and R. Taborda, *Advances in Applied Remote Sensing to Coastal Environments Using Free Satellite Imagery*. 2014.

[19] M. G. M. Sarwar and C. D. Woodroffe, "Rates of shoreline change along the coast of Bangladesh," *J. Coast. Conserv.*, vol. 17, no. 3, pp. 515–526, 2013.

[20] Y.-J. Choung and M.-H. Jo, "Shoreline change assessment for various types of coasts using multi-temporal Landsat imagery of the east coast of South Korea," *Remote Sens. Lett.*, vol. 7, no. 1, pp. 91–100, 2016.

[21] Z. Du, W. Li, D. Zhou, L. Tian, F. Ling, H. Wang, Y. Gui, and B. Sun, "Analysis of Landsat-8 OLI imagery for land surface water mapping," *Remote Sens. Lett.*, vol. 5, no. 7, pp. 672–681, 2014.

[22] M. Lipakis and N. Chrysoulakis, "Shoreline extraction using satellite imagery," in *In: Pranzini, E. and Wetzel, E. (eds): Beach Erosion Monitoring. Results from BEACHMED/e-OpTIMAL Project*

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

(Optimization des Techniques Integrées de Monitorage Appliquées aux Lottoraux) INTERREG IIIC South. Nuova Grafica Fiorentina, Florence, Italy, 2005, pp. 81–96.

[23] M. W. A. Halmy and P. E. Gessler, "The application of ensemble techniques for land-cover classification in arid lands," *Int. J. Remote Sens.*, vol. 36, no. 22, pp. 5613–5636, 2015.

[24] A. Gidudu, B. Abe, and T. Marwala, "Land Cover Mapping Using Ensemble Feature Selection Methods," 2008.

[25] G. R. Hashim, "Langkawi ' s Sustainable Regeneration Strategy and Natural Heritage Preservation E nvironment A sia," *EnvironmentAsia*, vol. 8, no. 2, pp. 1–8, 2015.

[26] N. Samat, "Assessing Land Use Land Cover Changes in Langkawi Island : Towards Sustainable Urban Living," *Malaysian J. Environ. Manag.*, vol. 11, no. 1, pp. 48–57, 2010.

[27] S. Gilmore, A. Saleem, and A. Dewan, "Effectiveness of DOS (Dark-Object Subtraction) method and water index techniques to map wetlands in a rapidly urbanising megacity with Landsat 8 data," *CEUR Workshop Proc.*, vol. 1323, no. March, pp. 100–108, 2015.

[28] T. Nguyen, "Optimal ground control points for geometric correction using genetic algorithm with global accuracy," *Eur. J. Remote Sens.*, vol. 48, pp. 101–120, 2015.

[29] C. Li, J. Wang, L. Wang, L. Hu, and P. Gong, "Comparison of classification algorithms and training sample sizes in urban land classification with landsat thematic mapper imagery," *Remote Sens.*, vol. 6, no. 2, pp. 964–983, 2014.

[30] M. Bochow, "Improving Class Separability - A Comparative Study of Transformation Methods for the Hyperspectral Feature Space," in *Proceedings of 4th EARSeL Workshop on Imaging Spectroscopy. New quality in environmental studies.*, 2005, pp. 439 – 447.

[31] S. Mao, L. Jiao, L. Xiong, S. Gou, B. Chen, and S. K. Yeung, "Weighted classifier ensemble based on quadratic form," *Pattern Recognit.*, vol. 48, no. 5, pp. 1688–1706, 2015.

[32] M. W. a. Halmy and P. E. Gessler, "The application of ensemble techniques for land-cover classification in arid lands," *Int. J. Remote Sens.*, vol. 36, no. 22, pp. 5613–5636, 2015.

[33] N. C. Oza and K. Tumer, "Classifier ensembles: Select real-world applications," *Inf. Fusion*, vol. 9, no. 1, pp. 4–20, 2008.

[34] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," *Int. Jt. Conf. Artif. Intell.*, vol. 14, no. 12, pp. 1137–1143, 1995.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 5: Performance Results of Satellite Image Classification Using Ensemble Voting Classifiers of Single Classifiers.

Single ML Classifier	DT	NB	kNN	LDA	QDA	LR	SGD	SVM-L	SVM-R	SVM-P	MLP
DT	-	-	-	-	-	-	-	-	-	-	-
NB	99.50	-	-	-	-	-	-	-	-	-	-
kNN	99.32	99.39	-	-	-	-	-	-	-	-	-
LDA	99.30	99.30	99.30	-	-	-	-	-	-	-	-
QDA	99.36	98.74	99.33	99.30	-	-	-	-	-	-	-
LR	99.54	99.36	99.44	99.30	99.37	-	-	-	-	-	-
SGD	99.40	99.30	99.19	99.16	99.33	99.34	-	-	-	-	-
SVM-L	99.48	99.34	99.38	99.30	99.37	99.38	99.31	-	-	-	-
SVM-R	99.23	99.37	99.23	99.30	99.26	99.42	99.32	99.35	-	-	-
SVM-P	99.27	99.43	99.36	99.27	99.45	99.46	99.48	99.40	99.29	-	-
MLP	99.54	99.55	99.57	99.30	99.55	99.55	99.56	99.55	99.56	99.52	-

Table 6: Performance Results of Satellite Image Classification Using Ensemble Voting Classifiers of Single and Ensemble Classifiers.

ML Classifier	DT	NB	kNN	LDA	QDA	LR	SGD	SVM-L	SVM-R	SVM-P	MLP
GDB	99.48	99.53	99.50	99.29	99.55	99.53	99.48	99.49	99.52	99.45	99.55
ADB	99.49	99.53	99.51	99.29	99.55	99.54	99.52	99.50	99.51	99.46	99.55
RF	99.31	99.57	99.43	99.16	99.40	99.39	99.38	99.55	99.59	99.27	99.42
ET	99.41	99.43	99.41	99.29	99.43	99.48	99.34	99.42	99.38	99.41	99.54