Hybridization of SLIC and Extra Tree for Object Based Image Analysis in Extracting Shoreline from Medium Resolution Satellite Images

Syaifulnizam Abd Manaf[1*]   Norwati Mustapha[1]   Md Nasir Sulaiman[1]
Nor Azura Husin[1]   Helmi Zulhaidi Mohd Shafri[2]   Mohd Norhisham Razali[1]

[1] *Intelligent Computing Research Group, Faculty of Computer Science and Information Technology,
Universiti Putra Malaysia, Serdang, Selangor, Malaysia*
[2] *Geospatial Information Science Research Centre, Faculty of Engineering, Universiti Putra Malaysia,
Serdang, Selangor, Malaysia*
* Corresponding author's Email: nizamkpt2020@gmail.com

Abstract: Observation satellites orbiting the Earth provide important surveillance information that helps in identifying various types of land cover. As such, the resolution of satellite images is critical to ensure high accuracy in classifying land cover types. Clearly, high-resolution images are desirable, but such images are prohibitively expensive. Hence, the use of medium-resolution satellite images seems more economical and practical. Several techniques have been developed to monitor the conditions of land covers across the world, such as aerial photography, ground survey, and remote sensing. Among the three techniques, remote sensing is the best, given its ability to monitor vast geographical areas more accurately and much faster compared to those of the other two techniques. In recent decades, many countries have been relying on remote sensing to monitor the conditions of coastal areas by extracting shorelines from satellite images. To date, several pixel-based methods have been proposed for the extraction of shorelines, but most of these methods are fraught with problems. Thus, the object-based approach is proposed using a combination of segmentation algorithms, namely Felzenswalb, Quickshift, and SLIC, together with 15 machine learning classifiers, to classify segmented images of Langkawi Island. The performance of the segmentation algorithms and machine learning classifiers were assessed in terms of segmentation time and overall accuracy in four experimental settings comprising of three different parameters. The research findings showed that the proposed hybridization of SLIC segmentation and Extra Tree classifier was the most efficient and accurate technique compared to other combinations of techniques in extracting the shoreline of the study area. Specifically, SLIC was faster than Felzenswalb and Quickshift by as much as 37 times and 500 times, respectively. Together with Extra Tree classifier, SLIC managed to achieve 100% overall accuracy in the object-based classification in three out of four settings compared to the rest of the techniques tested in the study.

Keywords: Object based image analysis, SLIC segmentation, Extra tree classifier, Machine learning, Shoreline extraction.

1. Introduction

Over the years, natural processes and anthropogenic activities are continually reshaping the coastal landscapes of countries. Naturally, changes in shorelines are largely due to highly dynamic processes caused by a host of factors, such as tides, winds, waves, water currents, sediments, and oceanic temperatures, among others. In recent years, many people have settled in newly developed towns and cities, some of which are located along the coasts, bringing in some socio-economic benefits to societies. Inevitably, anthropogenic activities resulting from the heavy concentration of people has adversely affected coastal areas of many countries [1]. Without redress, such activities can significantly change the coastlines of countries, which can occur unpredictably [2].

Therefore, monitoring of coastlines is needed to provide important information about prevailing conditions of coastal areas of a country by examining changes that are taking place along its borders. According to [3], such a border is commonly referred to as a 'coastline" or a 'shoreline'.

A shoreline is defined as a physical line that acts as an interface that separates land and water, thus creating a boundary between the two [4]. As such, shoreline extraction provides important information of the boundary between land and water, which can help detect and monitor any signs of coastal erosions or accretions. Currently, coastline monitoring can be performed using several techniques, such as remote sensing, aerial photography, and field survey. Given that it uses satellite images that can cover vast land areas, remote sensing is deemed the most accurate and fastest technique among the three. Moreover, this technique can extract important boundary information from satellite images using appropriate image analyses. In contrast, the techniques based on traditional field survey or airborne aerial photography are relatively time-consuming, laborious, and imprecise.

More importantly, the extraction of selected features of shoreline from satellite data can reveal important spatial and temporal characteristics. With information of such characteristics, researchers can determine and predict future coastal changes more accurately. In addition, acquiring information of a shoreline entails a shoreline indicator to accurately represent the true position of a particular shoreline of interest [3]. Essentially, the shoreline indicator 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 are imperceptible to the human's eyes [4]. For this study, the researchers used the third category of indicators as the shoreline indicator.

Admittedly, accurate delineation of shoreline boundaries is difficult to achieve as coastlines will undergo a gradual, continual change due to soil erosions and accretions caused by fluctuating tides. Cleary, such a change can seriously impede the extraction process of the pixel-based approach, as evidenced by the presence of salt-and-pepper problem in vector GIS generated from the conversion of classified images into raster format. To overcome this problem, object-based image analysis (OBIA) can be used to improve the extraction process.

To date, several approaches based on OBIA have been proposed to extract shorelines from optical multispectral satellite images. For example, [2][3] implemented an OBIA using the rule-based approach and vegetation index, namely Normalized Difference Vegetation Index (NDVI), to extract coastlines from Quickbird multispectral images. Furthermore, [3] used Digital Elevation Model (DEM) as additional data to classify Quickbird and Landsat images. Likewise, [1] proposed the use of OBIA with the rule-based approach and water index, namely Normalized Difference Water Index (NDWI), to improve the classification accuracy. However, some of the rules can sometimes overlap with or oppose one another which effectively reducing the classification accuracy [5]. Interestingly, [6] proposed a set of rules for OBIA using the combination of NDVI and NDWI to distinguish types of coast. Even though these spectral indices can reliably discriminate land cover classes, they only cover small ranges and have the potential to overlap with other classes [7]. To overcome the drawbacks of NDWI, machine learning (ML) can be used in the pixel-based approach to improve the classification accuracy [8].

In view of the important points highlighted in the discussion, this study was carried out with the main aim to propose and determine the most effective and efficient hybridization of segmentation algorithm and ML technique for OBIA in the extraction of shorelines. Specifically, the extraction process was performed on the Landsat TM satellite images to determine the shorelines of the North West coast of Peninsular Malaysia. For the extraction process, the researchers performed OBIA using 15 ML classifiers (11 single classifiers and 4 ensemble classifiers) to classify land-water classes. The proposed OBIA techniques are in fact the alternative approaches to the pixel-based classification methods, which have been studied by [9].

To facilitate discussion, this paper is organized based on the following sections. Section 2 discusses OBIA, and Section 3 details the materials and methods used in the study. Section 4 elaborates the experimental results of the proposed OBIA technique. Finally, Section 5 discusses the main research finding and its implication on the current practice and concludes the discussion of the paper.

2. Object-based image analysis

OBIA is a sound methodology that is closely related to human perception. Fundamentally, OBIA treats segments or objects, not individual pixels, as the basic processing units. It starts with segmenting an image into smaller homogeneous regions called objects. Essentially, the segmentation combines relevant pixels into objects based on spectral, shape, and neighbourhood similarity. The spectral variable is represented by the mean values and standard

deviations of a specific spectral band. The shape variable includes the size, perimeter, and compactness of an image. The neighbourhood variable indicates the mean difference between an object and other darker entities. The relations among such variables help improve the value of the final classification, which is lacking in the pixel-based approaches. For OBIA method, each object is treated as part of the 'super-object', where the latter is obtained by combining several neighbouring objects. In addition, each object can be subdivided into smaller objects, called 'sub-objects'. Superpixels algorithm is the process that groups neighbouring pixels of an image that have similar features into perceptually meaningful homogeneous regions.

Segmentation of the OBIA method is the process of partitioning a projected image into multiple segments, known as superpixels. Each segment corresponds to a group of pixels, with one segment assigned to each pixel. The aim of segmentation is to represent the image as an object that is more meaningful and easy to analyse. As such, similar pixels should be grouped together. The segmentation assigns every pixel with a certain level such that pixels with the same level will share common visual characteristics. Furthermore, the aggregation of pixels as an object is based on certain common features, such as colour, texture, edge, spatial location, gradient, and others. Algorithms for generating superpixels can be broadly divided into two categories, namely the graph-based method and the gradient-ascent-based method. For this study, only three segmentation algorithms were considered, namely Felzenszwalb (Felz), Quickshift (QS), and Simple Linear Iterative Clustering (SLIC).

2.1.1. Felz segmentation

Felz segmentation algorithm is a graph-based method, which was proposed by [10], that uses an efficient graph-based representation of local neighborhoods. The weight of an edge $w(v_i, v_j)$ between two vertices is used as the level of dissimilarity of two points (differences in intensity, range, color, location, etc.). It splits nodes such that edges inside a component will have relatively lighter weights and edges between components will have relatively heavier weights. The algorithm has a single scale parameter that influences the actual size and the number of segments, which can vary depending on the local contrast.

2.1.2. Quickshift segmentation

QS segmentation is a gradient-ascent-based method that uses the mode-seeking segmentation algorithm proposed by [11], which is similar to that of the mean-shift method. The algorithm segments a satellite image by identifying clusters of pixels based on the joint spatial and color dimensions. The superpixels produced by QS, which are not fixed, are influenced by three main parameters, namely the *ratio(λ)*, *kernel size(σ)*, and *maximum distance(τ)*, where λ is a trade-off between the distance in the color-space (R,G,B) and the distance in the spatial-space (R,G,B,X,Y). Lower ratios indicate high spatial importance; in contrast, higher ratios place greater emphasis on colors. The kernel size (σ) controls the scale of the local density approximation. For the hierarchical segmentation, the parameter τ assigns a particular level to segments of the subtrees.

2.1.3. SLIC segmentation

SLIC a gradient-ascent-based method that utilizes K-means for superpixel segmentation. In this method, the pixels of an image are initially clustered and then iteratively refined until certain criteria are met to form the superpixels [12]. Essentially, SLIC uses the colours of pixels to cluster them based on their similarity and proximity to one another on the image plane. SLIC uses a 5-dimensional vector to generate superpixels by assigning certain weights to parameters *L, a,* and *b* of the CIELAB color space and by measuring the similarity of spatial coordinates based on Euclidean distances.

Interestingly, this technique can generate compact superpixels at lower computational cost. In fact, the "compactness" parameter can effectively generate a good trade-off between the similarity and proximity of pixels. In addition, *n-segments* can select any number of centres for k-means. Overall, SLIC is an efficient method for clustering pixels as it only compares superpixels' centres, as opposed to k-means method that compares all clusters' centres.

In addition to the segmentation algorithms, ML classifiers are an integral part of the building block in the OBIA. In this study, 15 machine learning classifiers were used, including the extra-tree classifier.

2.2 Extra-Tree classifier

An extremely randomized tree classifier, known as Extra-Tree (ET) classifier, was proposed by [13]. Essentially, ET is an ensemble classification method of decision tree classifier, which differs from classic decision trees in the way they are built. This means that in order to find the best split to separate the samples of a node into two groups, random splits are drawn from each of the randomly selected features,
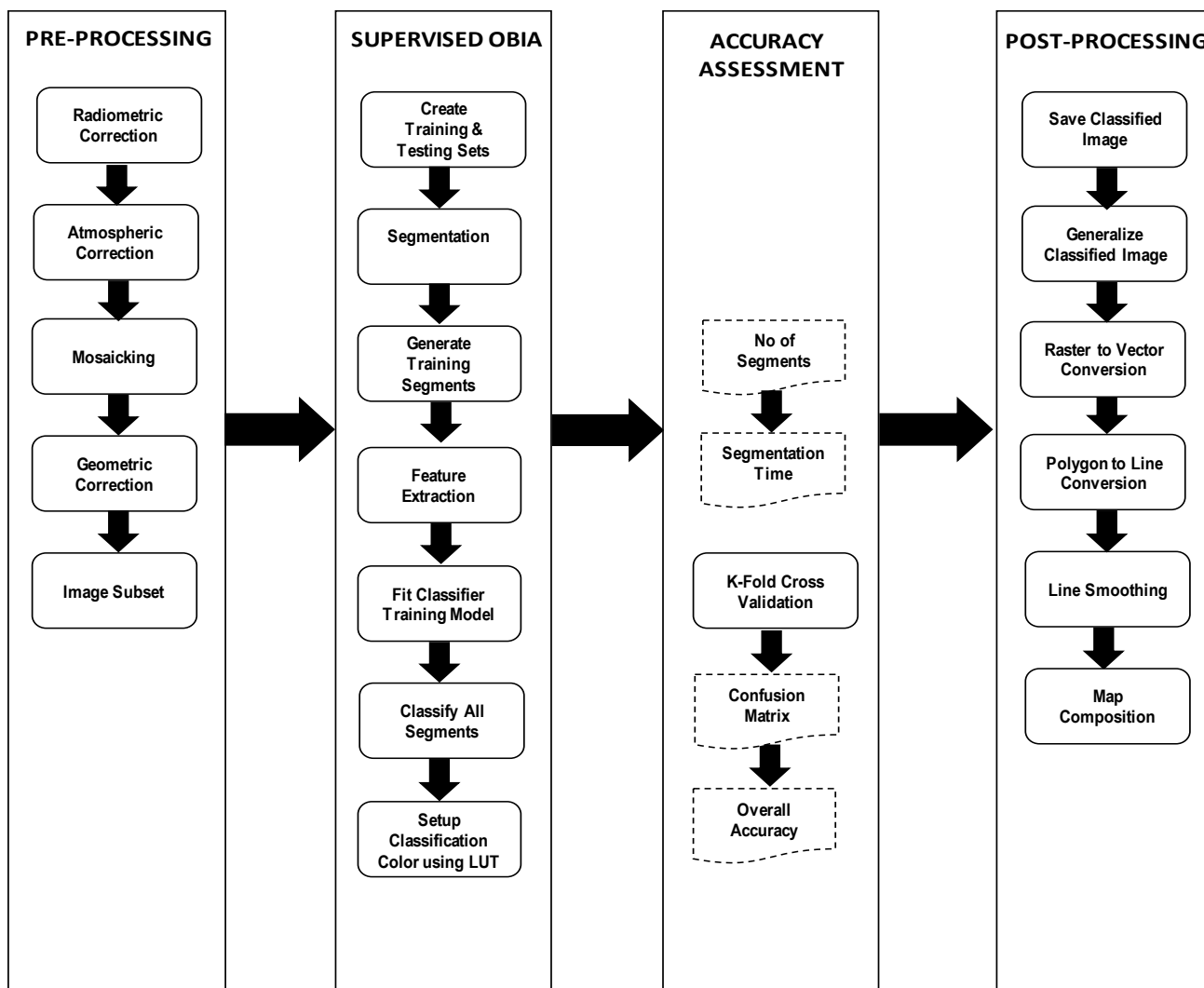
Figure.1 Methodology of this research

Figure.2 Study area of the research

namely *max_features,* from which the best split will be chosen. As such, ET classifier fits with various sub-samples of a dataset by averaging out the variance problems of single decision tree method, which helps improve the predictive accuracy and the control of over-fitting.

3.   Materials and methods

In this study, the methods used to extract shorelines from satellite images consisted of four phases, namely pre-processing, supervised OBIA, accuracy assessment, and post-processing, as depicted in Fig. 1. Prior to the pre-processing phase, a study area was chosen from where relevant data were collected.

3.1 Study area

The chosen area of this study was Langkawi Island, which is located at the North West coast of Peninsular Malaysia, as shown in Fig. 2. This island is located at 6° 15'N and 6° 29'N latitude and 99° 37'E and 99° 57'E longitude, covering a total area of about 478.48 km^2. This exotic island consists of a major landmass, with many smaller islands dotting its coastlines.

Table 1. The dates of data acquisition of two scenes of multispectral Landsat-5 TM

Data	Path/Row	Acq. date	Acq. time (GMT)
Landsat TM	129/56	24/12/2010	03:29:49
Landsat TM	128/56	05/08/2011	03:28:41

3.2 Data acquisition

In this research, two scenes of multispectral Landsat-5 Thematic Mapper (TM) were used as the research data to ensure the study area would be sufficiently covered. The researchers collected the data of the two scenes on two different dates. The first scene was acquired on 24 December 2010 at 11.29 am local time (GMT+8), while the second scene was acquired on 5 August 2011 at 11.28 am. Table 1 summarizes the dates of data acquisition of the two scenes.

For the last few decades, the Landsat TM satellite had orbited the Earth more than 150,000 times, transmitting over 2.5 million images of land surfaces of 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 at that time.

The repeat cycle was within 16 days with 185 km swath width. Landsat TM satellite has an 8-bit radiometric resolution with 256 grey value levels that specify the image brightness. It also has 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) µm: 0.76–0.90 µm; Short-wave Infrared (SWIR) 1: 1.55-1.75 µm; SWIR 2: 2.05-2.35 µm; and thermal: 10.40-12.50 µm. All these bands have spatial resolutions of 30 m, except the thermal band that has a spatial resolution of 120 m.

3.3 Pre-processing

Pre-processing of satellite images is an important step for attaining high classification accuracy and easing computational complexity. In the pre-processing phase, five processes were performed, namely radiometric correction, atmospheric correction, mosaicking, geometric correction, and subset of the study area. 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.

Radiometric calibration was immediately performed on the image once the study area and image data were established. Such calibration converted the satellite image to radiance (L_λ) image, as expressed by Eq. (1).

$$L_\lambda = Gain \times Pixel\ value + Bias \qquad (1)$$

The pixel value is a digital number (DN) that ranges from "0" to "255", and the radiance of each image band depends on the gain and bias values. After calibrating the image data, dark object subtraction (DOS) method was applied to the radiance image to remove haze components caused by the additive atmospheric scattering effects of the image data [14]. This method assumes there is a high probability that at least a few pixels of an image will be black with near-zero percent reflectance (e.g., water, cloud, and shadow), which is due to atmospheric scatterings called path radiance that need to be removed [15].

Essentially, the pixel value represents the value of an image's element that must be subtracted from a particular spectral band to remove the first-order scattering component. The removal process uses the minimum value of a band, which represents its background signature. Given that TM images were the data of this study, the use of the DOS method was deemed appropriate, as it had been successfully used in earlier generation of Landsat sensors. The DOS method includes a simple multiplicative correction to deal with the effects of atmospheric transmittance. The land surface reflectance can be estimated by using Eq. (2).

$$p = \frac{\pi(L_{sat} - L_p)d^2}{E_0\ cos_2(\theta_2)} \qquad (2)$$

In this formula, L_{sat}, L_p, and d are the satellite radiance, path radiance, and Earth-sun distance (in astronomical unit), respectively. Mosaicking process was then applied to combine the two scenes of the image into a single large image based on the same coordinate system. Then, image registration process was performed using image-to-image geometric correction process with root-mean-square (RMS) value of 0.457 for 30 ground control points (GCPs). Later, Rectified Skew Orthomorphic (RSO) Kertau was utilized to project the image data (of the West Coast of Peninsular Malaysia) onto a local projection system. Finally, image sub-setting was performed to fit the image to the study area, yielding an image of 2,513 x 1,830 pixels. Once cleaned, the image was used in the classification phase.

3.4 Supervised OBIA

In this phase, the OBIA used image segmentation as the preliminary step to enhance the classification process after the required training and testing samples had been established. Arguably, the proper use of the supervised OBIA relies heavily on the user's knowledge of such analysis. With such knowledge, proper training samples can be developed to build the appropriate model of ML algorithms to classify different types of land covers. The training samples were represented by a number of groups of homogeneous pixels, each of which defined a particular type of land cover based on the closest spectral features.

On the other hand, the testing samples are usually used to measure the performances of such models, which in this case involved the land and water classes. Using the land ground truth data as a reference, the training and testing sets were created from the random sampling of several randomly placed observations. For this study, only one set of training and testing samples was created in the form of polygons, comprising 260 and 65 polygons for land and water classes, respectively. To ensure separability of the training and testing set, Jeffries-Matusita distance and Transformed Divergence were used [9], yielding a separability index of 1.97, which was close to 2.0 that indicated perfect separability.

The hybridization involved two processes, namely segmentation, and classification. For the segmentation process, the quality of the segmentation helped determine the classification output. Depending on the type of segmentation algorithms and parameters used, this process may produce a number of dissimilar segments. In this study, the segmentation algorithms used were Felz, QS, and SLIC, and the chosen parameters for such algorithms were determined by trial and error. Then, the training and testing sets of each class were generated by matching the segments generated from the image to the original training and testing sets.

Before classification was performed on the segments created by segmentation algorithms, feature extraction process was applied to provide segment representation. Specifically, feature extraction encompasses the analysis and extraction of machine-readable information from satellite images in order to obtain a concise, compact description. Naturally, the context of a specific task or domain determines the type of features to be extracted. In this study, the features extracted were largely statistical descriptors, such as (i) min, (ii) max, (iii) mean, (iv) variance, (v) kurtosis, and (vi) skewness, which helped improve the final classification process.

Admittedly, in view of the many techniques available, it was quite challenging to choose the appropriate ML classifiers for the OBIA in such problem domain. Nevertheless, the researchers managed to select 15 ML classifiers consisting of 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 were AdaBoost (ADB), Gradient Boosting (GDB), ET, and Random Forest (RF). After the model had learned part of the image regions of the training segments, the whole image segments were then classified with additional extracted features to improve the classification results. In addition, a lookup table (LUT) consisting of all colour classes was used to render the classified image with appropriate colours. In this study, brown and blue were used to represent land and water, respectively.

3.5 Accuracy assessment

The goal of the accuracy assessment phase was to measure the accuracy of the classification process carried out in the OBIA. Specifically, the accuracy of extracted satellite images was measured in terms of the percentage of an area of a map that was correctly classified according to the reference data. In this phase of the study, such percentage was treated as the overall accuracy of the ML classifiers. In fact, the overall accuracy has been used as the primary performance indicator in many studies of the evaluation of classification methods of satellite images. The accuracy of the classification is calculated by dividing the sum of entries that forms the major diagonal of the confusion matrix by the total number of samples taken as expressed in Eq. (3) and Eq. (4).

$$n = \sum_{i=1}^{k} \sum_{j=1}^{k} n_{ij} \qquad (3)$$

$$OA = \frac{\sum_{i=1}^{k} n_{ii}}{n} \qquad (4)$$

In Eq. (3) and (4), n and n_{ij} are the total number of samples and the diagonal elements of the confusion matrix, respectively.

Assessing the accuracy of ML classifiers in fitting a model usually entails the use of cross

validation techniques or rotation estimation. The k-fold cross-validation method helps partition an original sample into k subsamples of equal size. A single subsample will be treated as validation data to test the model, while the remaining (k-1) subsamples will be used as training data [16]. For this study, the 10-fold cross-validation method was used to deal with overfitting and class imbalance problems.

3.6 Post-processing

After assessing and verifying the classification accuracy, the resultant classified image was saved to a hard drive for further post-processing operations, such as sieve and clump. First, sieve was used to remove isolated pixels based on a specific threshold value (number of pixels). Then, clump was applied to add spatial coherency to existing classes by combining similar classified areas that were adjacent with one another. Later, a raster-to-vector conversion process was performed to convert the resultant classified image into GIS vector format. Subsequently, polygon-to-line conversion process was carried out to convert the image into vector lines format, which was then followed by line editing to remove unwanted data (i.e., lines that were not part of the shoreline). Such a process was laborious and prone to committing errors, such as the unintentional removal of important data (e.g., lines). Finally, line smoothing was applied to the cleaned shoreline to smooth straight edges and corners of angular features that helped ensure curves would be seamlessly connected at their vertices.

4. Results and discussions

The analysis of the experimental data was initially 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 28 GB RAM, running on Microsoft Windows 7 (a 64-bit operating system). However, the system memory could only support a few ML classifiers, such as DT, NB, LDA, LR, and SGD. Accordingly, the workstation's RAM was upgraded to 40GB to support all the 15 ML classifiers.

As acknowledged, our approach was different from those used in the previous studies carried out by [1-3, 6]. Actually, our work focused on examining the accuracy of OBIA of satellite images based on the hybridisation of several segmentation algorithms and machine learning classifiers. In contrast, [1-3, 6] used commercial tools in their studies, such as eCognition and ENVI, with which the OBIA process was carried out with unknown segmentation algorithms and
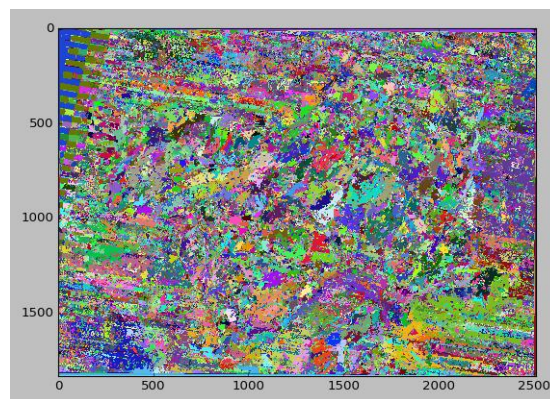
(a)
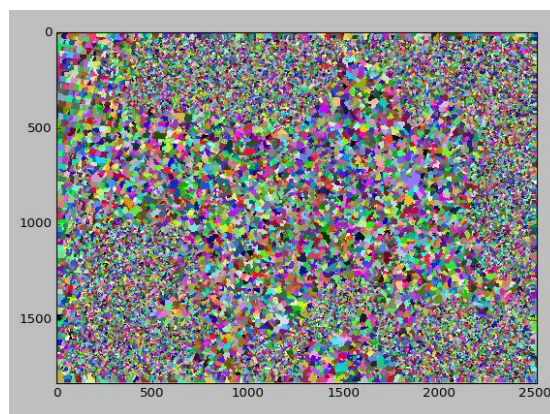
(b)
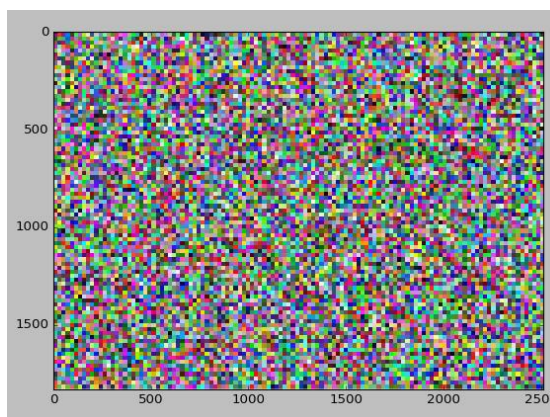
(c)

Figure.3 Image segmentation output: (a) Felz, (b) QS, and (c) SLIC

limited number of classifiers. Moreover, their studies did not involve accuracy assessments, which had been dealt with in our study. In view of the above differences, this present study helped address the void of research on hybridisation of techniques and assessments of accuracy.

Figures 3a, 3b, and 3c show the results of the segmentation outputs of Felz, QS and SLIC, respectively. The results showed that the different segmentation algorithms produced different types of segments, which were highly visible to the reader.

Table 2. Results of OBIA using Felz, QS and SLIC segmentation algorithms

Input	Configuration Parameters													
Segmentation Algorithm	Felz				QS					SLIC				
Scale	50	100	1,000	85	Kernel size	25	12	12	10	No of segments	4,000	10,000	40,000	80,000
Sigma	0.4	0.9	0.5	0.25	Max Distance	10	5	4	3	Compact	10	10	10	8
Min size	300	50	1000	100	Ratio	0.8	0.8	0.2	0.5	Sigma	0.25	0.8	1	0.8
Process	Segments Generated													
No of Segments	3,412	9,502	15,847	63,386	No of Segments	3,417	26,322	50,055	84,419	No of segments	3,996	10,440	37,848	71,906
Time to segment (s)	8,600	7,101	9,200	5,801	Time to segment (s)	150,327	67,169	67,204	5,699	Time to segment (s)	238	260	250	276
Training Segments for Class 1	91	206	85	478	Training Segments for Class 1	151	368	471	630	Training Segments for Class 1	204	297	475	633
Training Segments for Class 2	116	11	342	456	Training Segments for Class 2	94	294	491	694	Training Segments for Class 2	141	201	394	586
Output	Overall Accuracy (%)													
DT	90.4	98.5	98.8	99.11	DT	95.54	98.97	99.08	98.41	DT	99.9	99.81	98.74	98.36
kNN	90.4	74.9	99.09	98.37	kNN	99.38	97.95	98.07	98.87	kNN	98.8	98.7	99.57	99.67
NB	91	96.9	98.41	97.59	NB	92.97	99.21	98.23	97.99	NB	98.9	98.78	98.75	99
LDA	79.3	99.3	99.86	99.02	LDA	99.94	98.46	99.32	99.3	DA	99.2	99.24	99.13	99.2
QDA	89.8	98.3	98.62	96.93	QDA	98.86	99.3	98.07	97.82	QDA	98.3	98.52	98.39	98.46
LR	74.1	100	97.76	99.51	LR	95.52	99.15	98.82	98.95	LR	99.2	99.2	98.73	99.32
SGD	89.4	100	97.53	96.99	SGD	99.21	98.61	98.28	97.34	SGD	98	98.99	99.09	99.18
MLP	29	71.4	28.62	72.26	MLP	74.08	72.09	28.57	28.57	MLP	71.9	72.19	71.83	71.74
SVM-L	82.2	100	97.74	99.5	SVM-L	95.71	98.97	99.08	99.74	SVM-L	98.7	98.74	98.66	99.31
SVM-R	40.4	98.3	87.5	79.19	SVM-R	79.89	73.72	36.74	34.88	SVM-R	74.5	72.57	72.23	71.79
SVM-P	90.2	98.8	98.68	99.25	SVM-P	95.78	99.15	98.64	99.74	SVM-P	98.4	98.58	99.74	99.9
ADB	55	99.4	99.59	98.89	ADB	93.77	98.97	99.08	98.41	ADB	99.4	99.42	100	99.76
ET	90.8	100	100	99.74	ET	98.07	100	100	99.74	ET	**99.8**	**100**	**100**	**100**
GDB	90.9	95.7	98.78	99.59	GDB	98.07	98.97	99.08	100	GDB	98.4	99.62	100	99.95
RF	80.2	99.8	99.64	99.64	RF	98.07	100	100	99.7	RF	99.9	99.31	99.74	99.71

Table 2 summarizes the results of OBIA of the 15 classifiers using Felz, QS, and SLIC segmentation algorithms. As highlighted, the segmentation time was an important performance criterion in OBIA. Evidently, SLIC was the most efficient method, using 200 to 300 seconds to process 4,000 to 80,000 segments. In contrast, the least efficient method was QS, taking 6,000 to 150,000 seconds to process the same number of segments. Furthermore, it was observed that the time used to segment the satellite image in QS was directly proportional to the size of kernel, suggesting that the larger the kernel size the longer it would take to segment the image. In between the above two methods, Felz performed moderately well, needing 5,000 to 10,000 seconds to segment 4,000 to 70,000 segments.

Three parameters, namely the *scale*, *sigma*, and *min size*, of the Felz segmentation settings were used to determine the classification accuracy. The results clearly showed Felz-ET was the most accurate method, attaining extremely high accuracy in three out of four settings. More impressively, it managed to achieve 100% overall accuracy in two of the configuration settings. In addition, an analysis to determine which configuration settings had significantly contributed to high overall accuracy was carried out. For the first configuration setting (scale = 50, sigma = 0.4, and min size = 300), Felz-NB achieved the highest accuracy, which was closely followed by Felz-GDB and Felz-ET. Interestingly, as the number of segments increased further to 9,502 segments and beyond, Felz-ET was observed to be the most efficient method in all the remaining configuration settings.

Three parameters of QS segmentation settings were used to determine the classification accuracy, namely the *kernel size*, *max distance*, and *ratio*. The results showed QS-ET was the most accurate method as evidenced by its ability to achieve 100% accuracy in two of the configuration settings. For the remaining settings, it managed to achieve relatively high accuracy, except for the first configuration setting (kernel size = 25, max distance = 10 and ratio = 0.8), in which the recorded overall accuracy was 98.07%. The same analysis also focused on determining the best configuration settings that helped generate the number of segments and the respective training segments for each class. For the first configuration setting (kernel size=25, max distance=10 and ratio=0.8), QS-LDA was the most accurate method, which was followed by QS-kNN and QS-SGD. For the final configuration setting (kernel size = 10, max distance = 3, ratio = 0.5), the most accurate method was QS-GDB, which

impressively attained 100% accuracy. Almost equally impressive were QS-ET, QS-SVM-P, and QS-SVM-L, all of which managed to achieve 99.74% accuracy.

Three parameters of the SLIC segmentation settings were used in assessing the classification accuracy, namely the *number of segments*, *compactness*, and *sigma*. The results showed SLIC-ET was the most accurate method, attaining 100% accuracy in three out of four settings. For the remaining settings, this method also performed exceptionally well, recording an accuracy of 99.77%. In addition, the same analysis was performed to determine the best configuration settings that helped generate the number of segments and the respective training segments for each class. For the first configuration setting (number of segments = 4,000, compact = 10, and sigma = 0.25), the most accurate method was SLIC-RF. For the same setting, the second and third most accurate methods were SLIC-DT and SLIC-ET, respectively. For the remaining configuration settings involving 10,000 to 80,000 segments, SLIC-ET was clearly the best method in terms of accuracy, comfortably securing 100% accuracy.

The classification process produced an image classification map, which clearly demarcates the boundary separating the land class and the water class. Later, post-classification processes were performed on the image classification map to extract the shoreline from the satellite image. As demonstrated, the line editing process was easier to perform using the proposed OBIA approach compared to the pixel-based approach, as the raster-to-vector conversion process generated less unwanted data. Finally, the base map was overlaid with the extracted shoreline output to delimitate its boundary for further shoreline map analysis. Fig. 4 shows the shoreline (rendered in red) that distinctively outlines the base map of the original Landsat TM image.

Figure.4 The extracted shoreline overlaid the base map

As demonstrated in this study, the use of OBIA with hybridised SLIC-ET was the most efficient and accurate method, having attained the fastest segmentation time and highest classification accuracy. Such finding is hardly surprising given that SLIC can produce compact superpixels which effectively lowers the computational overhead of classification process. Additionally, the research findings showed Tree-based classifiers were able to achieve fast segmentation time and high classification accuracy, with ET being the best among such classifiers. Essentially, ET is quite similar to RF, but the former splits nodes randomly without having to comply with any selection criteria. This kind of randomness that makes ET better than RF.

5. Conclusion

In this study, the researchers proposed a hybrid SLIC-ET algorithm for OBIA involving two stages, namely fast segmentation and object-based classification. In the first stage, three segmentation algorithms were used, namely Felz, QS, and SLIC, to segment a Landsat TM image into superpixels. For the second stage, feature extraction was applied to the segmented images to extract several important features. Then, a series of experiments involving 15 ML techniques was carried out to classify the segmented image. The results showed that ET classifier in combination with the three segmentation algorithms (i.e., Felz, QS, and SLIC) was the most accurate classifier for image segmentation based on OBIA, having achieved the highest overall accuracy. Equally revealingly was that ET classifier outperformed the other methods in all settings of the three segmentation algorithms. RF was the second most accurate classifier, falling slightly short of the performance of ET. In contrast, MLP was the least accurate method among those tested in this study.

Interestingly, SLIC-ET was observed to be the best method among the combinations of algorithms for OBIA in extracting shoreline given its extremely high accuracy and exceedingly fast segmentation. Following closely were Felz-ET and QS-ET, which were observed to be the second and third most accurate method, respectively. Together, these findings suggest that SLIC-ET is the best hybrid for OBIA compared to other hybrids of segmentation algorithms and ML classifiers. Moreover, the findings also suggest that choosing a particular OBIA method for shoreline extraction depends on a number of factors, including the parameters of segmentation algorithms, extracted features of segmented objects, and type of classifiers.

Clearly, further research should focus on other critical features, such as spectral indexes, NDVI, and NDWI layers, to improve the classification accuracy. In addition, a comparative analysis can be carried out to compare the differences between extracted shorelines and reference shorelines.

Acknowledgments

References

[1] K. Kalkan, B. Bayram, D. Maktav, and F. Sunar, "Comparison of support vector machine and object based classification methods for coastline detection", *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.*, Vol. 40, No. 7W2, pp. 125–127, 2013.

[2] M. B. Giannini and C. Parente, "An object based approach for coastline extraction from Quickbird multispectral images", *Int. J. Eng. Technol.*, Vol. 6, No. 6, pp. 2698–2704, 2015.

[3] J. A. Urbanski, "The extraction of coastline using OBIA and GIS", *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2010, Vol. 46, No. 81, p. 378.

[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] O. Usmerjena, A. Podatkov, and D. Zaznavanja, "Object-based image analysis of remote sensing", *Geod. Vestn.*, Vol. 55, No. 4, pp. 665-688, 2011.

[6] X. Ge, X. Sun, and Z. Liu, "Object-oriented coastline classification and extraction from remote sensing imagery", In *Proc. of SPIE - The International Society for Optical Engineering*, Vol. 9158, p. 91580M, 2014.

[7] S. Szabó, Z. Gácsi, and B. Balázs, "Specific features of NDVI, NDWI and MNDWI as reflected in land cover categories", *Landsc. Environ.*, Vol. 10, No. 3-4, pp. 194-202, 2016.

[8] Y. J. Choung and M. H. 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*, Vol. 2017, No. 2009, 2017.

[9] S. A. Manaf, N. Mustapha, M. N. Sulaiman, N. A. Husin, and M. R. 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. 23, No. 12, pp. 13001-13006, 2017.

[10] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient Graph-Based Image Segmentation", *Int. J. Comput. Vis.*, pp. 1–26, 2004.

[11] A. Vedaldi and S. Soatto, "Quick shift and kernel methods for mode seeking", *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 5305 LNCS, No. PART 4, pp. 705–718, 2008.

[12] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 34, No. 11, pp. 2274–2281, 2012.

[13] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees", *Mach. Learn.*, Vol. 63, No. 1, pp. 3–42, 2006.

[14] 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", In: *CEUR Workshop Proceedings*, Vol. 1323, pp. 100–108, 2015.

[15] T. Schroeder, W. Cohen, C. Song, M. Canty, and Z. Yang, "Radiometric correction of multi-temporal Landsat data for characterization of early successional forest patterns in western Oregon", *Remote Sensing of Environment*, Vol. 103, No. 1. pp. 16–26, 2006.

[16] 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.