Bulletin of Electrical Engineering and Informatics
Vol. 10, No. 6, December 2021, pp. 3178~3190
ISSN: 2302-9285, DOI: 10.11591/eei.v10i6.3199

Assessing mangrove deforestation using pixel-based image: a machine learning approach

Ahmad Yahya Dawod[1], Mohammed Ali Sharafuddin[2]
[1]International College of Digital Innovation, Chiang Mai University, Thailand
[2]College of Maritime Studies and Management, Chiang Mai University, Thailand

Article Info

Article history:

Received Aug 3, 2021
Revised Oct 3, 2021
Accepted Nov 1, 2021

Keywords:

Mangrove forests
Multiclass classifier
Object-based image classification
Random forest
Support vector machine
Time series

ABSTRACT

Mangrove is one of the most productive global forest ecosystems and unique in linking terrestrial and marine environment. This study aims to clarify and understand artificial intelligence (AI) adoption in remote sensing mangrove forests. The performance of machine learning algorithms such as random forest (RF), support vector machine (SVM), decision tree (DT), and object-based nearest neighbors (NN) algorithms were used in this study to automatically classify mangrove forests using orthophotography and applying an object-based approach to examine three features (tree cover loss, above-ground carbon dioxide (CO_2) emissions, and above-ground biomass loss). SVM with a radial basis function was used to classify the remainder of the images, resulting in an overall accuracy of 96.83%. Precision and recall reached 93.33 and 96%, respectively. RF performed better than other algorithms where there is no orthophotography.

Corresponding Author:

Mohammed Ali Sharafuddin
College of Maritime Studies and Management
Chiang Mai University
Samut Sakhon, 74000 Thailand
Email: mohammedali.s@cmu.ac.th

1. INTRODUCTION

Mangroves offer numerous benefits such as carbon sequestration, soil erosion prevention, and rich biodiversity. Thus monitoring deforestation and managing mangrove forest sustainability are of crucial concern, and satellite remote sensing is used to map and assess the changes over time [1]. However, the traditional remote sensing approaches lack accuracy due to the coarse spatial resolution of the satellite image, resulting in spectral species being confused with vegetation in landward areas [2]. A previous study [3] reported 75 to 90% accuracy in detecting mangrove species, although the research only considered primary forest areas overlapping mangrove tree cover. Light detection and ranging (LiDAR) and optical remote sensing combined with the SVM algorithm to map and create a mangrove inventory was already adopted and tested [4]. However, to the best of author's knowledge, the applicability of LiDAR and optical images, together with RF, is not tested for mapping mangrove forests. Furthermore, no comparison between random forest (RF) and support vector machine (SVM) in mangrove mapping has yet been performed [5]. Changes in mangroves forests have been previously studied using freely available time-series data of satellite-based imagery [6]. Such remotely sensed time series data have proven to be effective for monitoring mangrove ecosystem changes, using both aerial photographs (APs) as well as optical and synthetic-aperture radar (SAR) data [7]. Thus, numerous studies have recently applied time series data to map mangrove changes. In those, APs and visual interpretation are the most commonly applied techniques for detecting changes in mangrove forests [8], [9]. This study similar to Turubanova *et al*. [10] defines primary forests as tropical

woodland cover that has not remained totally cleared and regrown in later history. Landsat images was classified into primary forest data, using a separate algorithm for each region and defined tree cover as all vegetation higher than three meters as of 2000 [10]. To date, to the best of the author's knowledge, there is no published research available on the investigation of mangrove of all the 18 provinces within the Gulf of Thailand. Hence, this research provides the first record of 21st century mangrove history for seven different provinces along with the mangrove changes around the Gulf of Thailand from 2001-2020. The growth and distribution of the mangrove species for the year 2020 is predicted using datasets from 2001-2019, consisting of the sea level rises, climate scenarios, and the influence of terrain change.

2. RELATED WORK

Recently, in Thailand, Palaeoecological analysis was conducted to assess mangrove species evolution and changes over a period of one thousand five hundred years [11]. The study area covered the Salak Phet Bay of Koh Chang Island in the Eastern Gulf of Thailand. The study findings revealed that mangroves have survived throughout the past 1500 years, but the growing sea level raise challenges their existence. Hence, satellite-based time series analysis could help to quantify and assess the extent of mangrove forest evolution, as well as the decreasing or increasing trends over time. This data could be utilized as a marker for structural features changes of the forests [12]. Therefore, sensing the satellite-based spatio-temporal failures and success dynamics of naturally disturbed forests or rehabilitation projects and their recovery, along with field surveys, are essential for assessing and maintaining the long-term sustainability of mangrove forests. In addition, earth observation data can assist to fill the data gaps arising from the using only the field data [13]. Hence, using Landsat time series data to analyze and monitor the forest changes is reliable [14]. However, the authors noted several challenges to expanding the same approach in a global context. A precise study on mangrove forests in Thailand, classifying them into two: the mature and naturally present mangroves in the western coast of Thailand, and the younger rehabilitated mangroves in the Gulf of Thailand [15]. The study reported serious mangrove deterioration throughout the last century and mentioned various reasons for it, such as land encroachment, aquaculture, and pollution. On the positive side, several community-driven [16] restoration projects have helped to restore the mangrove forests in the Gulf of Thailand. However, no studies have been conducted on structural development, species variety, and tree density in adjacent natural and young reforested or rehabilitated mangrove. Hence, the seven provinces in the Gulf of Thailand undergoing mangrove restoration are included in this study.

3. OBJECTIVES

The first objective of this study is to examine time series data of seven provinces in the Gulf of Thailand using statistical measurement approaches and selected modern and robust machine learning algorithms to classify mangrove species using high spatial resolution imagery. The second objective is to compare the performance and accuracy of four machine-learning algorithms, namely SVM, RF, decision tree (DT), and object-based NN algorithms to classify mangroves using datasets from the global land analysis and discovery laboratory of the University of Maryland, Google, USGS, and NASA, and by applying an object-based approach. Finally, this study aims to provide three alternative features: tree cover loss, aboveground CO_2 emissions, and aboveground biomass loss for the years 2001-2019 with positive values.

4. DATA AND METHODS

4.1. Study area

Mangrove forests survive in intertidal zones, so they tend to be immersed in periodic seawater. They also grow close and thick over time, making access difficult for extensive field surveys, and sampling. Consequently, remote detected data are widely utilized in mangrove mapping, ranging from mid-high-resolution satellite imagery to high-resolution drone based imagery. The classification of mangroves from remote sensing can involve general image analysis and object-based image classification using pixels. The representation used to characterize the results of the machine learning is a hypothesis for four machine-learning methods. Thus, SVM, RF, DT, and NN were selected to train the dataset. The hypotheses are described in the forms of specific relation between four statistical methods to involve the raw data by using three features such as tree cover loss, aboveground CO_2 emissions, and aboveground biomass loss.

Thailand is a Southeast Asian country with 77 provinces. The Gulf of Thailand covers roughly 320,000 km^2 and is the busiest marine region in the country [17], [18]. Although 17 provinces within the Gulf of Thailand have reported increased forest loss, this study focuses only on seven provinces namely: Samut Sakhon, Samut Songkram, Surat Thani, Chon Buri, Rayong, and Nakhon Si Thammarat Figure 1.

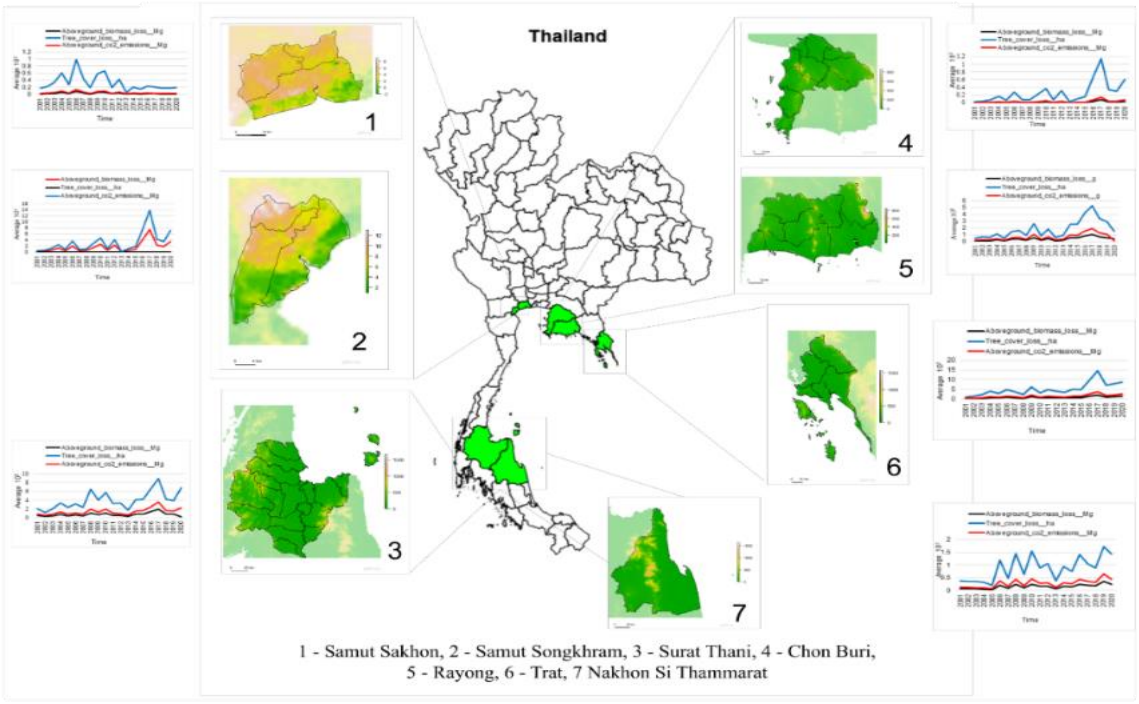

Figure 1. Seven Provinces included in the study and their time-series plot of tree cover loss, above ground CO_2 loss and above ground bio-mass loss data source: [17]; map source: GADM

4.2. Dataset selection

The dataset (2001-2020) used in this study was obtained from a collaboration project between the GLAD laboratory at the University of Maryland, Google, USGS, and NASA [19], Tree cover regions measure decrease in resolution of around 30 × 30 m^2. Data were obtained from the Landsat 5 Maper (TM), the Landsat 7 map plus (ETM+), and the Landsat 8 operating landscaper (OLI) sensors utilizing multi-spectral satellite imaging. More than one million satellite images, including around 600,000 Landsat 7 for 2000-2012, and over 400,000 Landsat 5, 7, and 8 images for the period 2011-2019 were processed and analyzed, providing updates to the satellite. There is obvious land area monitoring and a supervised learning method in satellite images. Song *et al*. [20] the pixel tree cover loss has been utilized to identify. The tree cover is described in this data set as any vegetation greater than 5 m in height that may be planted through a range of canopy densities in the shape of nature's forests. Tree cover loss is described as a booth replacement disturbance or the entire tree cover evacuation canopy within the pixel range of Landsat. Tree cover loss may be due to human actions including forest collection or deforestation (natural forest transformation into various land uses) and also natural conditions like storm damage or illness. Tree cover losses may also be the result of human activity. Therefore, deforestation is not comparable to loss.

4.3. Normalization

4.3.1. Normalize data

Normalization is a preprocessing step in the scaling method and used to find new information from the existing array, and can aid the prediction process. There are several approaches for predicting data. To maintain a large variation in prediction and forecasting, the normalization method is required to make them closer [21]. Standard deviation can still be used with normalised data because both translation and linear scaling have no effect on it. The classification stage entails determining the correlation values for each pair of graphs and applying several rule-based techniques to categorize the characteristics of mangrove forests into various categories [22]. The proposed normalization method is given as shown in with an explanation.

$$B = \frac{(|A|)-(10^{n-1})\times(|x|)}{10^{n-1}} \quad (1)$$

Where A is the particular data component, $n$is the number of digits in component A, x is the first digit of a data component, $B$is the scaled value between 0 and 1.

4.3.2. Mangrove forest normalization

The Landsat images were normalized utilizing image normalization to decrease spectral variations between Landsat, as described as shown in. This approach of the mangrove forest pixels applies the average values for linear correction to each spectral band [7]. The variation in mean sea level over the 20 years of monitoring showed negative correlation with normalized soil pore water salinity, such that years of high salinity with low mean sea level are related with [23], indicative of lower levels of tidal inundation in several regions in the Gulf of Thailand, thus impeding mangrove growth as.

$$\varphi_{adj(\lambda)} = \varphi_{ori(\lambda)} - (M_{man(\lambda)} + M_{ref(\lambda)}) \quad (2)$$

Where $\varphi_{adj(\lambda)}$ is the normalized median for the band λ, $M_{man(\lambda)}$ is the median value of the dense mangrove forest of the sample size for band λ , and $M_{ref(\lambda)}$ is the reference dense mangrove forest value for band λ as computed for the well-known, stable locations identified using the Landsat images across the study area.

4.4. Machine learning algorithms

Machine learning (ML) discipline is closely connected to the database discipline [6] therefore machine learning involves the integration of additional questions about the computational architectures and algorithms which can be most effectively used to pull-in, index, merge, recover, and store these data; how different learning subtasks can be arranged in a bigger system, and respond to the questions of computational tractability. Thus, the model is considered as an approximation of the process that machines are required to mimic. In such a situation, some input errors may be obtained, but mostly, the model provides correct answers. Hence, another measure of performance (besides that involving the memory usage and speed metrics) of a machine learning algorithm is the accuracy of results. In this study, four machine-learning algorithms are chosen, consisting of SVM, RF, DT, and the object based NN algorithm.

4.4.1. Support vector machine

Machine learning is defined as the science of getting computers to learn the thoughts and actions of humans to improve the performance of computers in analyzing data and information from the human command [24]. One of the machine learning algorithms that may be utilized in the Google Earth Engine is the SVM. The SVM, being a supervised non-parametric statistical learning method, may be used for both classification and regression, making it appropriate for categorizing mangrove and non-mangrove forests based on pixel reflectance [25]. The SVM is responsible for finding the decision boundary from several different classes and maximizing the margin. An illustration of this could be found in the research conducted by Heumann [1] which classifies mangrove cover using the SVM algorithm on the results of segmentation. The SVM is employed in the training process by entering the training data into the vector space. The nearest pattern from the training data is called a support vector. Only one previous study has applied the SVM [26] in the analysis of mangroves as part of a combination methodology to map mangroves using spectral and image texture data, while the performance of SVM for the multispectral classification of mangroves stays untested. The mathematical equation can be annotated as.

$$f(x) = w^t x + a \quad (3)$$

Considering the linear classifier for a binary classification problem with labels y and features$x$, $y \in \{-1,1\}$ is used to denote the class labels and parameters w, a as$w$: normal to the line,a: bias.

4.4.2. Random forest

The segmented training samples from the dataset were exported as comma-separated values (CSV) with all seven classes input into the random forest classification and implemented in Weka 3.9 [27]. Preprocessing was performed on the training and testing sets (remaining non-ground objects) using CSV which is the file structure utilized by Weka. The classification model was trained applying RF with 500 iterations and then utilized to classify the validation tests as well as the continuing test delineated as.

$$\rho\sigma^2 + \frac{1-\rho}{B}\sigma^2 \quad (4)$$

Where the expectation of an average B value of such trees is the same as the expectation of any one of them. An average of B variables randomly, each with the variance σ^2, has a variance of $\frac{1}{B}\sigma^2$, positive correlation in pairs ρ, the average variance.

4.4.3. Decision tree

The machine learning techniques are better than traditional approaches and remote sensing applications are increasingly used to monitor the change of mangrove forest from series data in distant wetland research. Recent research has demonstrated the effective usage of DT learning in remote sensing. It is one of the most popular techniques in mangrove studies. It promotes various machine learning approaches and has various advantages such as being very accurate, cost-effective, time-efficient in land-cover classification, based on remote saving, and providing long-term data access [28]. Therefore, the purpose of this research is to develop straight from the training classification criteria for researching theDT learning data's capacity without human participation. In addition, DT does not use multi-temporary Landsat data ranges like the max-forest cover in seven provinces, unlike other analytical statistical techniques for classification and detection of changes DT [29] as:

$$LM(x) = \sum w_i \, v_i(x) \tag{5}$$

Where $v_i(x)$ is the value of attribute A_i; x and w_i are weights. The attributes used in the path of the tree can be ignored.

4.4.4. Object-based NN classification

In object-based classification, segmentation is the process of finding pixel clusters with comparable properties and can generate objects of varying numbers and dimensions based on spectral seamlessness and compactness thresholds. There can be a hierarchy of segmentation levels extending from a few large items to numerous little objects with each object belonging to a huge object at a higher segmentation level and reduced objects on a lower level. The spectral attributes of each object include not only statistical values such as maximum, minimum, and standard variation of each band but also median value for each band from the pixels participating in the object. All these variables can be utilized during the classification process to support the discrimination of objects and their correct position to the land-use cover classes [30]. The classes are constructed according to a tree-like hierarchy in order to inherit the higher-level super-class features in the classification tree in a similar way to the object connection. the lower structural levels. In this study, the image was segmented using the cluster merging method, which starts with pixel-sized objects and iteratively matures to combine tiny objects into a bigger one of the image objects until the homogeneous threshold is exceeded. The homogeneous threshold is concluded by user-defined parameters such as shape, scale, color, compactness, smoothness, and image layer weights, with the scale being the most essential parameter, directly controlling the measure of object images. The object-based NN classification method is utilized in this study to distinguish three mangrove varieties and surrounding land-use types, with the NN algorithm achieving an overall accuracy more than 93.22%. The criteria or attributes mentioned above were used to label the objects and for further object-based nearest neighbor (NN) classification. This supervised classification methods allows all objects to be classified in a whole picture based on the selectable samples and statistics [31].

$$f = \sum_{i=1}^{i} W_i \, ({}^{n}M\sigma M - ({}^{n}obj1\sigma obj1 + {}^{n}obj2\sigma obj2)) \tag{6}$$

Where n is the number of bands and W_i is the weight of the current band, ${}^{n}M$, ${}^{n}obj1$, and ${}^{n}obj2$ are the number of pixels within the merged object, initial object 1, and initial object 2, respectively. Symbols σM, $\sigma obj1$, and $\sigma obj2$ are the variances of the merged object, initial object 1, and initial object 2, respectively, derived from the local tone heterogeneity and weighted by the size of the object images and summed over n image bands. After the segment of an image is done, it is subsequently classified as an object-based classification at the segment level.

5. RESULTS AND DISCUSSION

Both the USGS dataset and time-series Landsat sensor data were applied in this study. From 2001, the main cause of the mangrove forests was the settlement of the mudflats. Mapping was used to quantify these areas, where tree loss may have occurred at any point throughout the period of the time series, by considering three features: tree cover loss, aboveground CO_2 emissions, and aboveground biomass loss for the years 2001-2020. A rising rule was then applied to categorize the remaining tree loss locations which were not detected in the tree cover. The expanding rule barrier was raised. However, objects were only involved in the iterative process of tree cover loss if they shared a border with those already classified as trees. The growth rule was applied to finalize the tree cover classification and the findings were then utilized to determine changes of mangrove cover, with three stated characteristics to make it clearer.

5.1. Training data selection

The satellite imagery was used in this research to map land use in the study region by means of training data. First, the classification key was formulated using a high-resolution imagery map acquired (3-5 km spatial resolution) from the seven provinces inside the Gulf of Thailand following a stratified random sampling strategy conducted with datasets. Training and validation data collected from [32] the University of Maryland, Google, USGS, and NASA (2001-2019), for this study using a predictive model to obtain the dataset for 2020. A large training set of 17 classes based on 17 provinces in the Gulf of Thailand was validated using 10-fold cross-validation, with 99.6% accuracy achieved. However, we then reduced the number of provinces to seven to reduce the training data while making the study more challenging, achieving an accuracy of 96.86%. The training data consisted of three features, namely tree cover loss, aboveground CO_2 emissions, and aboveground biomass loss over 20 years from 2001-2020. For classification and validation over 20 years, the training and testing datasets were divided into 80% and 20%, respectively [33]. Finally, four machine-learning methods (SVM, RF, DT, and NN) were selected to train the dataset. The results of the training were then tested for model development. Figure 2 illustrates the architecture of the classification process for mangrove forest models using the method proposed in this study.

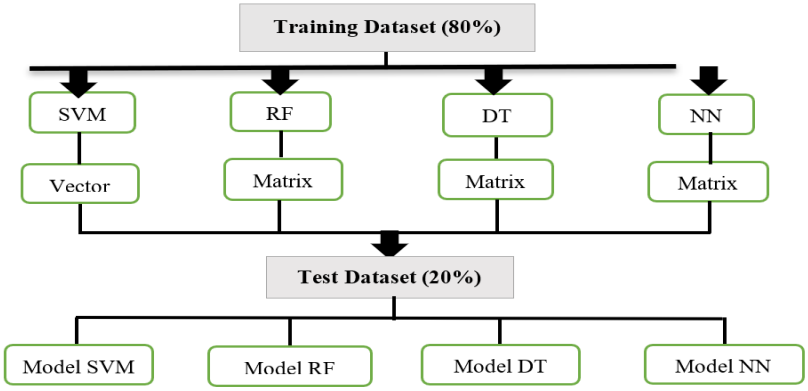

Figure 2. The architecture of classification process for mangrove forest models

5.2. Evaluation and measurement

In the area ratio of seven classes (provinces) to data training and testing, 700 samples for three characteristics were taken in seven provinces, 80% were randomly picked for classifier modelling and 20% were randomly selected for independent accuracy checks. Due to inaccessible circumstances and difficulties in reaching each province, as well as complications in accessing and covering all the information due to COVID 19 pandemic situation, we focused on this research to cover all deForest ation on the coast of the Gulf of Thailand in the study. Changes in the extent of mangrove forest s were observed across all provinces under study due to both natural and anthropogenic drivers of change. During the period 2001-2020, no region had an intact mangrove extent, with anthropogenic disturbance/removal occurring in all of the sites. The conversion of mangrove to aquaculture and agriculture (especially notable in the Gulf of Thailand provinces) was the most prevalent source of anthropogenic-induced change, which was virtually exclusively restricted to Thailand. Whereas this distinguishes the topographical dispersion of forest loss, it does not convey the extent to which the training happened. Several areas in Southern Thailand exhibited a localized loss. Land conversion for large-scale agriculture such as palm oil; shrimp framing, illegal logging is some of the primary drivers of deForest ation in Thailand and road and infrastructure development acted as an indirect driver, opening up new areas for timber harvesting. The transformation of mangroves into commercial shapes of nourishment and asset generation has been widely observed. The foremost common cause of anthropologically initiated alter was the transformation of mangroves into aquaculture/agriculture; the majority of which occurred in the areas of the gulf. Natural mangrove loss and advancement processes were regularly observed and widely distributed, arising in all areas. This study has recognized locales of seriously alter, both in mangrove gain and loss, and increased future monitoring is recommended. Where TP is true positive (positive issue of correct segmentation), TN is the true negative (negative effect of appropriate segmentation), FP is false positive (incorrect segmentation of a positive issue), and FN is false negative (negative point of incorrect segmentation). In evaluating the three features under study, namely tree cover loss, aboveground CO_2 emissions, and aboveground biomass loss for the years 2001-2020, seven out of the 17 provinces in the Gulf of Thailand exhibit positive values.

$$TNR = \frac{TN}{TN+FP} \times 100\% \quad (7)$$

$$TPR = \frac{TP}{TP+FN} \times 100\% \quad (8)$$

$$F_1 = \frac{2TP}{2TP+FP+FN} \times 100\% \quad (9)$$

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (10)$$

We calculate the classifiers performance concerning various execution measurements such as recall, precision, F-measure, accuracy region under ROC curve, and gamma measurement. The mean absolute error E_n of a particular a variable n is calculated by (11):

$$E_n = \frac{1}{s}\sum_{m=1}^{s}\left|P_{(nm)} - T_m\right| \quad (11)$$

Where $P_{(nm)}$ is the value predicted by the specific variable n for a sample issue$m$, and s is sample instances; T_m is the target value for a sample issue$m$. For a great fitting, $P_{(nm)} = T_m$ then $E_n = 0$ so, the E_n value is index ranges from 0to ∞, with 0 equivalent to the model [34]. There are distinct errors related to a simple predictor such as the relative absolute error is relative to a simple predictor, which is the average of the real values. In this issue of mangrove forest, though, the error is the overall absolute error instep of the total squared error. Consequently, the relative absolute error requires the overall absolute error and normalizes it by separating by the overall absolute error of the predictor. We calculate the relative absolute error E_n (12).

$$E_n = \sum_{m=1}^{s}\left|P_{(nm)} - T_m\right| / \sum_{m=1}^{s}\left|T_m - \bar{T}\right| \quad (12)$$

Where we calculate the value of $\bar{T}$ to the left side from (12) to (13), which s is sample instances.

$$\bar{T} = \frac{1}{s}\sum_{m=1}^{s} T_m \quad (13)$$

We compute the root mean squared error, which is the square root from (13), and root relative squared error is square root is calculated on (14).

$$E_n = \sqrt{\frac{\sum_{m=1}^{s}(p_{(nm)} - T_m)^2}{\sum_{m=1}^{s}(T_m - \bar{T})^2}} \quad (14)$$

Ten-fold cross-validation was applied to test the feasibility of distinct models. In the cross-validation procedure, a portion of data is kept aside, and it is being arranged by the remaining data for three features and 7 Provinces. Moreover, the process is common for various parts of the data by starting with 100 iterations and base learner. The portions are selected based on the value of k. The 10-fold cross-validation was used here implies data is divided into 10 parts [32]. Kappa statistic coefficient was developed to validate the nearness of the subjects that were displayed. The Kappa coefficient is a statistical measure of inter-rater unwavering quality or assertion that is utilized to evaluate subjective reports and decide assertion between two raters. We calculated by using the statistic formula defined by [35].

$$Kappa\ statistic\ coefficient = \frac{k_a - k_c}{1 - k_c} \quad (15)$$

$$k_a = \frac{TP+TN}{TP+TN+FP+FN} = Accuracy\% \quad (16)$$

$$k_c = \frac{(TP+FN)\times(TP+FP)+(FP+TN)\times(FN+TN)}{(TP+TN+FP+FN)^2} \quad (17)$$

Where k_a is a likelihood of success classification, and k_c is a likelihood of success due to the chance to estimate the kappa statistic coefficient. The cross-validation accuracy of machine learning methods such as FR, HMM, and RBM for applied Kappa statistics coefficient algorithm is greater than 9 so it is excellent results shown in Table 1.

Error rate (RE): ER predicted is an assessment of the likelihood of an error happening amid the completion of an assignment. Error rate predicted is utilized to decrease the probability of mistakes happening within the future [35].

$$ErrorRate = \frac{FN+FP}{TP+TN+FN+FP} \quad (18)$$

Table 1. Comparison of the error between four methods for CO_2

Errors	RF	SVM	DT	NN
Kappa statistic coefficient (k)	0.94	0.979	0.985	0.991
Mean Absolute Error (MAE)	0.001	0.002	0.007	0.006
Root Mean Squared Error (RMSE)	0.030	0.038	0.040	0.035
Relative Absolute Error (RAE)	11.327	12.738	9.049	7.232
Root Relative Squared Error (RSE)	15.114	18.731	19.839	17.620

Matthew's correlation coefficient (MCC): MCC is employed in machine learning as a measure of the value of binary (two-class) classifications. Its true positives, false negatives division with, false negatives, and false positives is commonly considered as acknowledged as an adjusted degree that may be used even when the classes are of extremely different sizes [36].

$$MCC = \frac{(TP*TN)-(FP+FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (19)$$

After comparing the performance of four machine learning algorithms (SVM, RF, DT, NN), classified on the basis of the accuracy obtained, the results for the seven provinces out of 17 in the Gulf of Thailand showed better mangrove class as well as other classes. The RF classifier received higher values for precision, recall, F-score, while also achieving a higher overall accuracy of 97.96%. This could be because characteristically, the SVM has difficulty in training a good model if there are many training samples and instances of 10-fold cross-validation. In mangrove mapping, many training samples are required to differentiate classes, specifically mangroves from other trees. On the other hand, due to their construction, RF, DT, and NN can handle a large amount of training data. This study applied a confusion matrix involving four machine-learning models: SVM, RF, DT, and object based NN classification. Tables 2 (a)-(d) shows the SVM, RF, DF, and NN algorithm in the confusion matrix of CO_2 with an accuracy of 93.17%. Carbon emissions reflect the carbon dioxide emitted to the atmosphere because of aboveground live-forest ed biomass loss. All biomass loss is "committed" radiations to the atmosphere upon the clearing, despite the fact that there are periods of low activity due to a few reasons of tree death. Emissions are "gross" estimates rather than "net" estimates, which means that due to a present lack of accurate data, data on the fate of the land after clearance, as well as its carbon value, is not incorporated. Emissions related with other carbon pools, such as shown in ground biomass, deadwood, litter, and soil carbon, are excluded from the files. Loss of biomass, like loss of tree cover, may arise for numerous reasons, involving deForest ation, fire, and monitoring within the course of forest ry operations. Table 2 (b) shows the confusion matrix value by accuracy 96.80% using the dataset for classes of classification random forest (RF) method. As shown in Figure 3, the performance of the four machine learning methods was assessed using precision recall and F-score parameters, which were categorized according to the seven provinces for CO_2 emissions.

Table 2. These tables are, (a) CO_2 emissions (SVM model), (b) CO_2 emissions (RF model), (c) CO_2 emissions (DT model), (d) CO_2 emissions (NN model)

(a)

True Label	*a*	*b*	*c*	*d*	*e*	*f*	*g*
a	48	0	1	0	0	0	0
b	0	53	2	0	1	0	0
c	0	0	70	0	0	0	0
d	0	0	0	21	0	0	0
e	0	0	0	1	20	0	0
f	1	0	0	1	0	129	2
g	0	0	0	1	0	0	160

(b)

True Label	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*Class*
a	48	0	1	0	0	0	0	Trat
b	0	52	2	0	0	0	2	Rayong
c	0	0	68	0	0	0	2	Chanthaburi
d	0	0	0	21	0	0	0	Samut Sakhon
e	0	0	0	0	21	0	0	Samut Songkhram
f	0	0	0	0	0	132	1	Surat Thani
g	1	0	0	0	0	1	159	Nakhon Si Thammarat

Predicated Label

(c)

True Label	*a*	*b*	*c*	*d*	*e*	*f*	*g*
a	46	1	2	0	0	0	0
b	0	52	2	1	0	0	1
c	0	2	66	2	0	0	0
d	0	0	0	20	0	0	1
e	0	0	0	1	20	0	0
f	0	0	0	0	0	130	3
g	0	0	0	1	1	2	157

(d)

True Label	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*Classified as*
a	40	3	2	2	0	2	0	Trat
b	0	49	4	1	2	0	0	Rayong
c	1	4	61	1	0	0	3	Chonburi
d	0	0	0	20	0	0	1	Samut Sakhon
e	0	0	1	1	18	0	1	Samut Songkhram
f	0	1	0	0	0	123	9	Surat Thani
g	1	1	2	1	1	5	150	Nakhon Si Thammarat

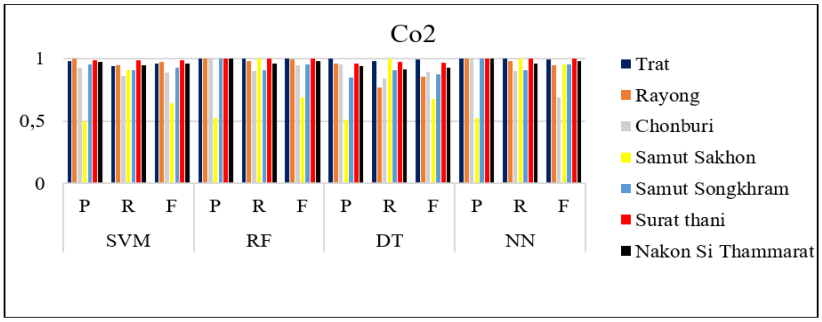

Figure 3. The performance of our four statistical measurement algorithms for seven provinces of CO_2 to show precision, recall, and F-score methods

Four classification algorithms were used in the confusion matrix for seven provinces. Tables 3 (a)-(d) present the aboveground biomass classifications for SVM, RF, DT, and NN, respectively. Figure 4 shows the performance results of the four machine learning algorithms for above-ground biomass, with SVM, RT, DT, and NN achieving 95.84%, 98.96%, 96.92%, 94.88% accuracy, respectively. The comparative error values of K, MAE, RMSE, RAE, & RSE for all the four methods for aboveground biomass is provided in Table 4.

Table 3. These tables are, (a) aboveground biomass (SVM), (b) aboveground biomass (RF), (c) aboveground biomass (DT), (d) aboveground biomass (NN)

(a)

True Label	a	b	c	d	e	f	g
a	7	0	0	0	0	0	0
b	0	6	0	0	1	0	0
c	1	0	6	0	0	0	0
d	0	0	0	7	0	0	0
e	0	0	0	0	7	0	0
h	0	0	0	0	0	7	0
f	0	0	0	0	0	0	7

(b)

True Label	a	b	c	d	e	f	g	classified as
a	7	0	0	0	0	0	0	Trat
b	2	5	0	0	0	0	0	Rayong
c	1	0	6	0	0	0	0	Chanthaburi
d	0	0	0	7	0	0	0	Samut Sakhon
e	0	0	0	0	7	0	0	Samut Songkhram
f	0	0	0	0	0	7	0	Surat Thani
g	0	0	0	0	0	0	7	Nakhon Si Thammarat

Predicated Table

(c)

True Label	a	b	c	d	e	f	g
a	6	1	0	0	0	0	0
b	0	7	0	0	0	0	0
c	0	0	7	0	0	0	0
d	0	0	0	7	0	0	0
e	0	0	0	0	7	0	0
f	0	0	0	0	0	7	0
g	0	0	0	0	0	0	7

(d)

True Label	a	b	c	d	e	f	g	classified as
a	7	0	0	0	0	0	0	Trat
b	0	6	0	0	1	0	0	Rayong
c	1	0	6	0	0	0	0	Chanthaburi
d	0	0	0	7	0	0	0	Samut Sakhon
e	0	0	0	0	7	0	0	Samut Songkhram
f	0	0	0	0	0	7	0	Surat Thani
g	0	0	0	0	0	0	7	Nakhon Si Thammarat

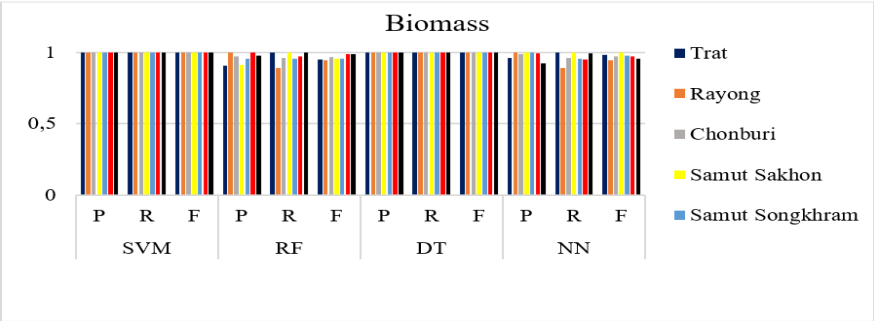

Figure 4. The performance of precision, recall and F-score methods by using four statistical measurement algorithms for seven provinces of aboveground biomass

Table 4. Comparison of the error between four methods for aboveground biomass

Errors	RF	SVM	DT	NN
Kappa statistic coefficient (k)	0.976	0.954	0.991	0.935
Mean Absolute Error (MAE)	0.005	0.023	0.002	0.014
Root Mean Squared Error (RMSE)	0.053	0.062	0.038	0.114
Relative Absolute Error (RAE)	2.298	10.735	16.292	6.4355
Root Relative Squared Error (RSE)	16.168	18.840	11.502	10.934

The tree cover can take the shape of natural woods or pasture in various densities surrounding a canopy, as the whole vegetation more than 5 m is defined. Loss is the removal or death of the covering of the tree and can be due to a number of reasons, such as mechanical fawning, disease, fire, and storm damage. The loss is therefore not like deForest ation. Pixels of loss are covered corresponding to the concentration of loss at around 30 x 30 m scale. The darker Pixels are shading embody regions with a greater thickness of tree cover loss, but brighter pixels shading indicate a smaller intensity of tree cover loss. Once the information is at a full resolution since not have any change in pixel shading.

In this study, four statistical measurements are involved for determining the tree loss in seven provinces, and this study attempts to modify these parameters to provide high accuracy for SVM, RF, DT, and NN of 94.27, 97.34, 96.20, 95.55%, respectively. Tables 5 (a)-(d) show the tree cover loss classifications confusion matrix for SVM, RF, DT, and NN, respectively. Figure 5 presents the performance results of the four machine learning methods for tree cover loss. The comparative error values of K, MAE, RMSE, RAE, & RSE for all the four methods for treecover loss is provided in Table 6.

Table 5. These tables are, (a) classification of tree cover loss (SVM), (b) classification of tree cover loss (RF), (c) classification of tree cover loss (DT), (d) classification of tree cover loss (NN)

(a)

True Label	a	b	c	d	e	f	g
a	48	0	1	0	0	0	0
b	0	53	2	0	1	0	0
c	0	0	70	0	0	0	0
d	0	0	0	21	0	0	0
e	0	0	0	1	20	0	0
f	1	0	0	1	0	129	2
g	0	0	0	1	0	0	160

(b)

True Label	a	b	c	d	e	f	g	Class
a	48	0	1	0	0	0	0	Trat
b	0	52	2	0	0	0	2	Rayong
c	0	0	68	0	0	0	2	Chonburi
d	0	0	0	21	0	0	0	Samut Sakhon
e	0	0	0	0	21	0	0	Samut Songkhram
f	0	0	0	0	0	132	1	Surat Thani
g	1	0	0	0	0	1	159	Nakhon Si Thammarat

Predicated Label

(c)

True Label	a	b	c	d	e	f	g
a	46	1	2	0	0	0	0
b	0	52	2	1	0	0	1
c	0	2	66	2	0	0	0
d	0	0	0	20	0	0	1
e	0	0	0	1	20	0	0
f	0	0	0	0	0	130	3
g	0	0	0	1	1	2	157

(d)

Ture Label	a	b	c	d	e	f	g	Class
a	40	3	2	2	0	2	0	Trat
b	0	49	4	1	2	0	0	Rayong
c	1	4	61	1	0	0	3	Chonburi
d	0	0	0	20	0	0	1	Samut Sakhon
e	0	0	1	1	18	0	1	Samut Songkhram
f	0	1	0	0	0	123	9	Surat Thani
g	1	1	2	1	1	5	150	Nakhon Si Thammarat

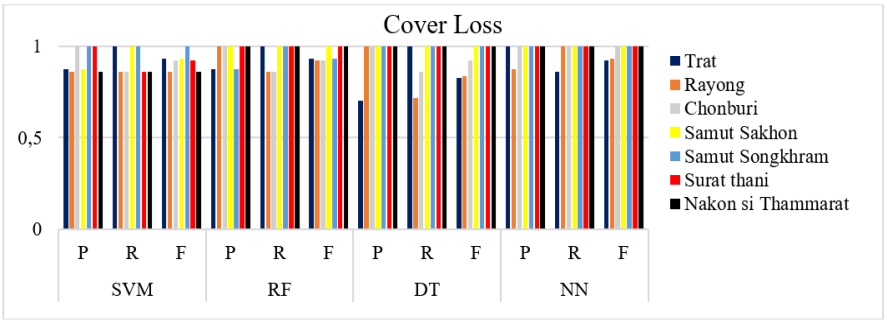

Figure 5. Tree cover loss feature involved the four classification methods for showing the performance of precision recall and F-score

Table 6. Comparison of the error between four methods for tree cover loss

Errors	RF	SVM	DT	NN
Kappa statistic coefficient	0.922	0.934	0.991	0.935
Mean Absolute Error	0.003	0.004	0.002	0.005
Root Mean Squared Error	0.041	0.045	0.029	0.034
Relative Absolute Error	13.324	11.657	16.301	9.608
Root Relative Squared Error	17.081	19.550	12.301	11.234

6. CONCLUSION

This study offers the initial assessment of changes in the mangrove forest range over the period 2001-2020 using SVM, RF, DT, and NN. In this paper, only seven provinces are presented to provide time series data on the influence of terrain change and climate scenarios with various levels of accuracy obtained by the algorithms for each province. This study proposes an alternative approach by using three features: tree cover loss, aboveground CO2 emissions, and aboveground biomass loss from 2001-2020 with different values for each feature. This approach can be tested using an alternative location to check its appropriateness for diverse data. Moreover, based on the data and classifier performance, the SVM, RF, DT, and NN performed well in mangrove classification achieving 95.84%, 98.96%, 96.92%, 94.88%, but it should be noted that RF performed better than the other three algorithms in areas with no orthophotography. In addition, RF Better performance in terms of accuracy since it achieved higher values in precision, recall, F-score and overall accuracy when comparing the three features. Future work should involve the selection of all seventeen provinces, involving several mangrove features in the Gulf of Thailand, as well as more machine learning approaches for higher precision and efficiency, such as Neural Network models and deep learning.

ACKNOWLEDGEMENTS

We thank the anonymous reviewers for their valuable comments and suggestions for improving the paper.

REFERENCES

[1] B. W. Heumann, "An object-based classification of Mangroves using a hybrid decision tree-support vector machine approach," *Remote Sensing*, vol. 3, no. 11, pp. 2440-2460, 2011, doi: 10.3390/rs3112440.
[2] M. Li, L. Ma, T. Blaschke, L. Cheng and D. Tiede, "A systematic comparison of different object-based classification techniques using high spatial resolution imagery in agricultural environments," *International Journal of Applied Earth Observation and Geoinformation*, vol. 49, pp. 87-98, 2016, doi: 10.1016/j.jag.2016.01.011.
[3] J. Cao, W. Leng, K. Liu, L. Liu, Z. He and Y. Zhu, "Object-Based Mangrove species classification using unmanned aerial vehicle hyperspectral images and digital surface models," *Remote Sensing*, vol. 10, no. 1, p. 89, 2018, doi: 10.3390/rs10010089.
[4] G. P. Petropoulos, C. Kalaitzidis and K. P. Vadrevu, "Support vector machines and object-based classification for obtaining land-use/cover cartography from Hyperion hyperspectral imagery," *Computers and Geosciences*, vol. 41, pp. 99-107, April 2012, doi: 10.1016/j.cageo.2011.08.019.
[5] A. V. Pada, J. Silapan, M. A. Cabanlit, F. Campomanes and J. J. Garcia, "Mangrove Forest cover extraction of the coastal areas of negros occidental, western visayas, Philippines using lidar data," in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences-ISPRS Archives*, vol. 41, pp. 73-79, July 2016, doi: 10.5194/isprsarchives-XLI-B1-73-2016.
[6] H. Wehle, "Machine Learning, Deep Learning and AI: What's the Difference?," *Conference: Data Scientist Innovation Day*, July 2017.
[7] U. Pimple *et al.*, "Google Earth Engine Based Three Decadal Landsat Imagery Analysis for Mapping of Mangrove Forest s and Its Surroundings in the Trat Province of Thailand," *Journal of Computer and Communications*, vol. 6, pp. 247-264, 2018, doi: 10.4236/jcc.2018.61025.
[8] K. Rogers, L. Lymburner, R. Salum, B. P. Brooke and C. D. Woodroffe, "Mapping of Mangrove extent and zonation using high and low tide composites of Landsat data," *Hydrobiologia*, vol. 803, no. 1, pp. 49-68, 2017, doi: 10.1007/s10750-017-3257-5.
[9] T. D. Pham, N. Yokoya, D. T. Bui, K. Yoshino and D. A. Friess, "Remote sensing approaches for monitoring Mangrove species, structure and biomass: Opportunities and challenges," *Remote Sensing*, vol. 11, no. 3. p. 230, January 2019, doi: 10.3390/rs11030230.
[10] S. Turubanova, P. V. Potapov, A. Tyukavina and M. C. Hansen, "Ongoing primary Forest loss in Brazil, Democratic Republic of the Congo and Indonesia," *Environmental Research Letters*, vol. 13, no. 7, p. 074028, July 2018, doi: 10.1088/1748-9326/aacd1c.
[11] A. Englong, P. Punwong, K. Selby, R. Marchant, P. Traiperm and N. Pumijumnong, "Mangrove dynamics and environmental changes on Koh Chang, Thailand during the last millennium," *Quaternary International*, vol. 500, pp. 128-138, 2019, doi: 10.1016/j.quaint.2019.05.011.
[12] C. Boisvenue, B. P. Smiley, J. C. White, W. A. Kurz and M. A. Wulder, "Integration of Landsat time series and

field plots for Forest productivity estimates in decision support models," *Forest Ecology and Management*, vol. 376, pp. 284-297, 2016, doi: 10.1016/j.foreco.2016.06.022.
[13] A. Banskota, N. Kayastha, M. J. Falkowski, M. A. Wulder, R. E. Froese and J. C. White, "Forest Monitoring Using Landsat Time Series Data: A Review," *Canadian Journal of Remote Sensing*, vol. 40, no. 5, pp. 362-384, 2014, doi: 10.1080/07038992.2014.987376.
[14] B. DeVries, "Monitoring tropical Forest dynamics using Landsat time series and community-based data," *PhD Thesis*, Wageningen University, 2015.
[15] N. Pumijumnong, "Mangrove Forest s in Thailand," in *Mangrove Ecosystems of Asia: Status, Challenges and Management Strategies*, pp. 61-79, 2014, doi: 10.1007/978-1-4614-8582-7_4.
[16] M. A. Sharafuddin and M. Madhavan, "Measurement Model for Assessing Community Based Wellness Tourism Needs," *e-Review of Tourism Research*, vol. 18, no. 2, pp. 311-336, 2020.
[17] M. A. Sharafuddin, "Types of Tourism in Thailand," *e-Review of Tourism Research (eRTR)*, vol. 12, no. 3, 2015.
[18] M. A. Sharafuddin and M. Madhavan, 'Thematic Evolution of Blue Tourism: A Scientometric Analysis and Systematic Review', *Glob. Bus. Rev.*, 2020, doi: 10.1177/0972150920966885.
[19] P. Potapov *et al.*, "Landsat analysis ready data for global land cover and land cover change mapping," *Remote Sensing*, vol. 12, no. 3, p. 426, 2020, doi: 10.3390/rs12030426.
[20] X. P. Song, C. Huang, J. O. Sexton, S. Channan and J. R. Townshend, "Annual detection of Forest cover loss using time series satellite measurements of percent tree cover," *Remote Sensing*, vol. 6, no. 9, pp. 8878-8903, 2014, doi: 10.3390/rs6098878.
[21] M. Madhavan, M. Ali Sharafuddin, P. Piboonrungroj and C. C. Yang, "Short-term Forecasting for Airline Industry: The Case of Indian Air Passenger and Air Cargo," *Global Business Review*, 2020, doi: 10.1177/0972150920923316.
[22] S. Patro and K. K. Sahu, "Normalization: A preprocessing stage," *arXiv Prepr. arXiv1503.06462*, 2015.
[23] K. McKee, K. Rogers and N. Saintilan, "Response of Salt Marsh and Mangrove Wetlands to Changes in Atmospheric CO_2, Climate and Sea Level," in *Global Change and the Function and Distribution of Wetlands*, Dordrecht: Springer Netherlands, 2012, pp. 63-96, doi: 10.1007/978-94-007-4494-3_2.
[24] N. S. A. Yasmin, N. A. Wahab and A. N. Anuar, "Improved support vector machine using optimization techniques for an aerobic granular sludge," *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 5, pp. 1835-1843, October 2020, doi: 10.11591/eei.v9i5.2264.
[25] F. Campomanes, A. V. Pada and J. Silapan, "Mangrove classification using support vector machines and random Forest.. algorithm: a comparative study," 2016, doi: 10.3990/2.385.
[26] Y. Jiang *et al.*, "High-resolution Mangrove Forest s classification with machine learning using worldview and uav hyperspectral data," *Remote Sensing*, vol. 13, no. 8, p. 1529, 2021, doi: 10.3390/rs13081529.
[27] A. Cutler, D. R. Cutler and J. R. Stevens, "Random Forests," in *Ensemble Machine Learning: Methods and Applications*, pp. 157-175, 2012, doi: 10.1007/978-1-4419-9326-7_5.
[28] E. B. B. Palad, M. J. F. Burden, C. R. Dela Torre and R. B. C. Uy, "Performance evaluation of decision tree classification algorithms using fraud datasets," *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 6, pp. 2518-2525, December 2020, doi: 10.11591/eei.v9i6.2630.
[29] L. T. Hauser, N. A. Binh, P. V. Hoa, N. H. Quan and J. Timmermans, "Gap-free monitoring of annual Mangrove Forest dynamics in ca mau province, vietnamese mekong delta, using the landsat-7-8 archives and post-classification temporal optimization," *Remote Sensing*, vol. 12, no. 22, p. 3729, 2020, doi: 10.3390/rs12223729.
[30] E. O. Makinde, A. T. Salami, J. B. Olaleye and O. C. Okewusi, "Object Based and Pixel Based Classification Using Rapideye Satellite Imager of ETI-OSA, Lagos, Nigeria," *Geoinformatics FCE CTU*, vol. 15, no. 2, pp. 59-70, 2016, doi: 10.14311/gi.15.2.5.
[31] G. P. Petropoulos, K. P. Vadrevu and C. Kalaitzidis, "Spectral angle mapper and object-based classification combined with hyperspectral remote sensing imagery for obtaining land use/cover mapping in a Mediterranean region," *Geocarto international*, vol. 28, no. 2, pp. 114-129, 2013, doi: 10.1080/10106049.2012.668950.
[32] M. C. Hansen *et al.*, "High-Resolution Global Maps of 21st-Century Forest Cover Change," *Science*, vol. 342, no. 6160, pp. 850-853, November 2013, doi: 10.1126/SCIENCE.1244693.
[33] A. Y. Dawod and N. Chakpitak, "Novel Technique for Isolated Sign Language Based on Fingerspelling Recognition," *13th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)*, pp. 1-8. 2019, doi: 10.1109/SKIMA47702.2019.8982452.
[34] M. J. Nordin, A. Y. Dawod and J. Abdullah, "An Adaptive Fingertips Detection Based on Skeletonization Technique for Invariant Hand Gestures," *International Conference on Engineering Research & Applications (ICERA)*, pp. 208-216, 2017, doi: 10.15242/DIRPUB.DIR0517046.
[35] C. Vasilakos, D. Kavroudakis and A. Georganta, "Machine learning classification ensemble of multitemporal Sentinel-2 images: The case of a mixed mediterranean ecosystem," *Remote Sensing*, vol. 12, no. 12, p. 3015, June 2020, doi: 10.3390/rs12122005.
[36] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, pp. 1-13, 2020, doi: 10.1186/s12864-019-6413-7.

BIOGRAPHIES OF AUTHORS

Ahmad Yahya Dawod, he is currently a lecturer at International College of Digital Innovation Chiang Mai University, Chiang Mai, Thailand. He received his BSc from the Faculty of Computer Science at Al-Mustansiriya University, Iraq in 2006. He completes his Master degree from the Faculty of Computing and informatics at Multimedia University - MMU Cyberjaya, Malaysia in 2012. He completes his Ph.D. from the Faculty of information science and technology at National University of Malaysia in 2018. His research interests include Artificial Intelligent, Machine Learning, Pattern Recognition, computer vision, Image Processing, and Medical Image Analysis.

Mohamed Ali Sharafuddin currently works as a lecturer at the College of Maritime Studies and Management, Chiang Mai University. His research interest is in “Blue economy” with special focus on Coastal, marine and Maritime areas and wellness Tourism. His most recent and working projects are, "Coastal, Marine and Maritime Tourism" and “electronic images”. He also actively works on projects in "Maritime marketing". His other interests include scientometrics, Structural Equation Modeling, prediction modeling and econometrics for both maritime and tourism business.