Global Scientific JOURNALS

GSJ: Volume 11, Issue 12, December 2023, Online: ISSN 2320-9186

www.globalscientificjournal.com

COMPARATIVE ANALYSIS OF FEATURES EXTRACTION METHODS IN REMOTE SENSING

JONAH Iyowuna Benjamin, Department of Surveying &Geomatics, Rivers State University, Port Harcourt. Email: jonah.iyowuna1@ust.edu.ng

OKEKE Francis Ifeanyi, Department of Geoinformatics and Surveying, University of Nigeria, Enugu Campus

OKWERE Chiziandu Enyinda, Department of Surveying &Geomatics, Rivers State University, Port Harcourt.

AKPOMRERE Oghenesuohwo Rufus, Department of Surveying and Geoinformatics, Delta State University of Science and Technology, Ozoro, Delta State. Email: akpomrerer@dsust.edu.ng

ABTRACT

The classification of satellite images entails dividing the pixel values of the images into relevant groups. This means assigning corresponding levels with respect to groups of homogeneity characters, with the aim of differentiating multiple objects from each other from the image. In most cases, this process became very difficult to understand. Comparatively analysis of features extraction methods was aimed to be the study. The materials covered ENVI 4.5, Idrisi 16.0, ArcGIS 10.7.1, Landsat imagery and administrative map of Rivers State. The KA results showed SVM (94.72%), BE (82.25%) and MAH (88.22%). Conclusively, a review of the literature was done on the more sophisticated approaches to classification, including binary encoding, support vector machines, and artificial neural networks. As a result, this study suggested developing classification systems that use a variety of classification algorithms.

KEY WORDS: Artificial Intelligence, Features Extraction and Remote Sensing

1.0 INTRODUCTION

The process of classifying satellite images entails dividing the pixel values of the images into relevant groups. This means assigning corresponding levels with respect to groups of homogeneity characters, with the aim of differentiating multiple objects from each other from the image (David, 2015). It is the process of classifying every pixel in an image or raw remotely sensed satellite data to produce a predetermined set of labels or land cover themes, according to Lillesand and Keifer (1994). It is essential for providing geographic information, which generates both quantitative and qualitative data and lessens the complexity of field work. Interpreting remote sensing imagery, conducting spatial data mining, and researching different vegetation types—such as forests and agriculture—all fall under the category of image classification (Sunitha & Suresh, 2015; Campbell, 2002).

In a multispectral image, a raster image is automatically classified using colours and spectral patterns, which classify all of the cells into a predetermined number of spectral classes. Recall that following appropriate classification, the relationship between the spectral classes and land surface materials will be known (Randall, 2011; Julius & Raphael, 1974).

The process of classifying satellite images, which is the act of allocating each pixel or group of pixels to thematic classes, is the primary method used to determine the spatial distribution of land cover and land use (Richards & Xia, 1999). The techniques used for classification are the supervised and unsupervised classification (Lillesand & Kiefer, 2000; Anderson et al.,1976; Jensen, 1996). The fields of spatial data mining, information extraction for applications, thematic map creation, field surveys, effective decision making, and disaster management are just a few of the applications that can benefit from spatial classification (Sunitha & Suresh, 2015; Butera, 1983; Earnst & Hoffer, 1979; Lo & Watson, 1998; Ozesmi & Bauer, 2002, Dean &Smith, 2003; Pal & Mather, 2003; Liu et al., 2002).

Several authors have different application modules to address a particular issue at a time, their results sometimes vary in their position or expectation. This is the problem of the research to see how close the various results with respect to the problem at hand. The goal of the study is to clarify the sophisticated classification techniques used in Geographic Information Systems

(GIS) and Remote Sensing.

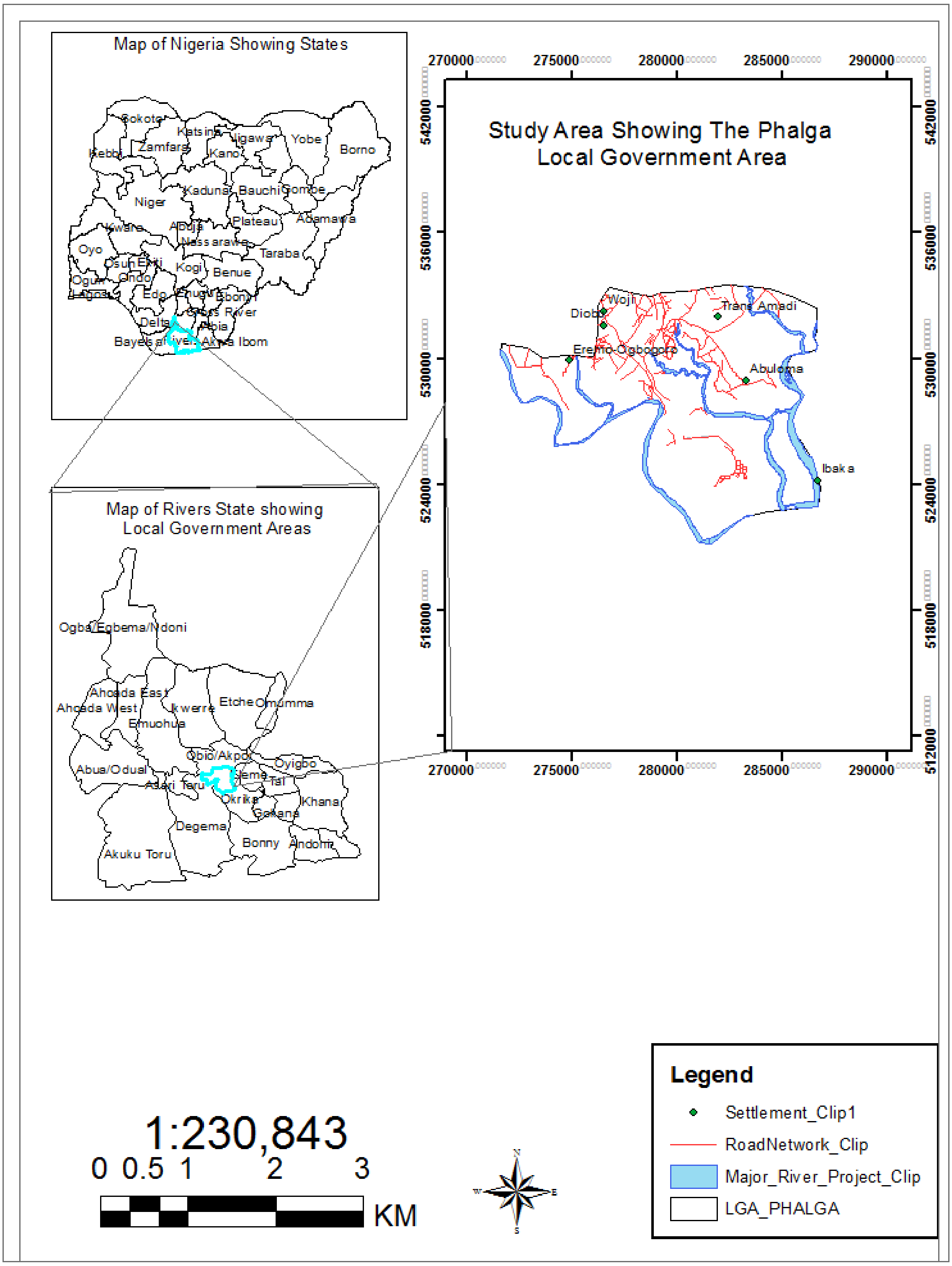


Figure 1: The Study Area

2.0 LITERATURE REVIEW

2.1 **Advance Image Classification Methods**

Advanced classification methods are grouped under the following and they are:

1.Artificial Neural Network (ANN)

2. Support Vector Machine (SVM)

3. Binary Encoding (BE)

2.2 Biological Neuron

Human brain is interconnected with over 100 billion of neurons. The interconnectivity of network of neuron is basically used to transfer information with each other through chemical and electrical signals. Synapse is developed to distant two neurons which are interconnected. Conversely, if two neurons are connected strongly, it means that the information it passes to each other will be strong too and this acts vis-visa. The four normal components of a typical neuron are the soma, axon, synapse, and dendrites. After receiving inputs from other neurons, dendrites produce a non-linear response and send a signal to other neurons through the axon when a threshold is crossed (Mirza, 2017; Fiona, 2001).

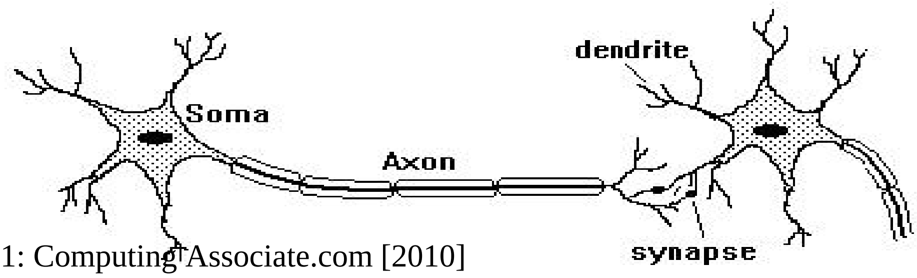


Figure 1: Computing Associate.com [2010]

2.2.1 **Artificial Neural Network (ANN)**

According to a historical event, the field of neurophysiology was the first to employ the term connectionism in the early 1940s to describe classical connections found in science philosophy, which enabled the majority of their computational capacity. The significant of computation intensified in 1950s to 1970s during the birth of computer. Old connectionism was recalled in that era which marked the beginning of theory artificial neural networks. This aspect of learning is linked to cognitive science revolution. Research showed that after the first two periods, new connectionism came to stay in 1986 called parallel distributed processing best described as the sub-symbolic processing, multi-layer neural network NN models. Late 1990s, newer connectionism was introduced to the field of Academy which vigorously heighted multilayer generative models, probabilistic approach to solve ontology problems (Igor, 2012; Kandel *e*t al., 1991; Kuffler et al., 1984; Thompson, 1993).

The area of artificial intelligence (AI) known as neural network models is commonly called artificial neural networks (ANNs). The architecture of this process defines how its several neurons are arranged, or position to each other. These arrangements are structured essentially by directing the synaptic connections of the neurons. It means that the association of the homologous material and paternal chromosomes during the initial part of the meiosis is well defined. The topology of the system is the bedrock of any neural network, within a particular architecture, which can be assumed with different structural compositions (Saravanan & Sasithra, 2014; SIPS, 2017).

A school of thought bemoaned the fact that statistical analyses carried out with conventional algorithms in the majority of biological sciences did not always produce very good results, especially when it came to classification analysis. According to studies by Sparling and Williams (1987), Martindale (1980), Figueredo et al. (1992), Le Pape and Chevalet (1992), Terhune et al. (1993), and other researchers, the majority of researchers' current classification techniques solely rely on parametric or non-parametric multivariate analyses: discriminant analysis. These methods are often rather inefficient when the data are nonlinearly distributed, even after variable transformation. Consequently, the need for effective and efficient method of classification was considered namely artificial neural networks (ANN).

Sten and Tambet (2015) discussed the statistical learning algorithms that are communicated by biological neural networks are mirrored by artificial neural networks (ANNs). They are employed in a plethora of tasks, including computer computations, speech recognition, and basic classification issues. (Maravall et al., 1991; Gemello & Mana, 1991; Vincent & Kevin, 2002) and character or image recognition (Belliustin et al., 1991; Van Allen et al., 1990; Tirakis et al., 1990; Omatu et al., 1990; Fukushima and Wake, 1990; Iwata et al., 1990). In hydrobiology models, comparison to multiple regressions, ANN clearly improved prediction performance (Lek et al., 1994, 1995). For classification purposes, ANN has been used for the analysis of protein structure (Qian & Sejnowski, 1988; Andreassen et al., 1990). Nevertheless, it also used in the classification of seaweeds and the recognition of impulsive noises in marine mammals (Nicolas et al., 1989; Smits et al., 1992**;** Yu-guo &Hua-pen, 2010; Furferi et al., 2011). Antennas are similar to biological neural networks in that they contain elements known as nodes, which are functionally similar to biological neurons when they are interconnected. The numerical values between two nodes are called weights, and they differ from one node to another. Alterations of these values in the network analysis readjust system to its desired function or result. Every node in the network receives numerous inputs from other nodes and uses the connection weights and inputs to calculate a single output. Usually, another neuron receives this output, repeating the process. With the knowledge from the previous sentence, it is simple to imagine the internal hierarchical structure of the artificial neural network, which is represented by the diagram below and consists of neurons arranged into various layers. The input layer is responsible for receiving inputs, while the output layer generates outputs.

Intelligent machines were discovered to complement human reasoning (McCullock & Pitts, 1943). Muller et al., (1995) state that there are two primary motivations for studying neural networks: the first is to learn more about the functioning of the human brain, and the second is to create machines that can solve complicated problems that computers that operate in sequence were unable to solve. The statement further stated that neural network has the capability of solving traditional program better in terms of error analysis. Artificial Neural Network (ANN) algorithms mimic the way humans learn by assigning meaningful labels to individual pixels in an image. The ability to easily incorporate additional data into the classification process and increase classification accuracy are two benefits of using ANN-based satellite image classification algorithms (Sunitha & Suresh, 2015).

2.2.2 **Architecture of Neural Network**

There three architectural properties which simply explain art and science of designing network system, they are:

1. Feed Forward Networks: The feed forward network is a collection of input, hidden and output layers. The input layer makes input signal to the neuronal network which are called passive nodes but do not partake in the signal modification. Hence, the active layers are called hidden layers because they carry arbitrary number of neurons and nodes in these layers are always involved in signal modification. Thirdly, the nodes in these layers are active ones, this is to say that the number of neurons in the output layers also correspond to numbers of output values of neural network (Hornik et al.,1989).

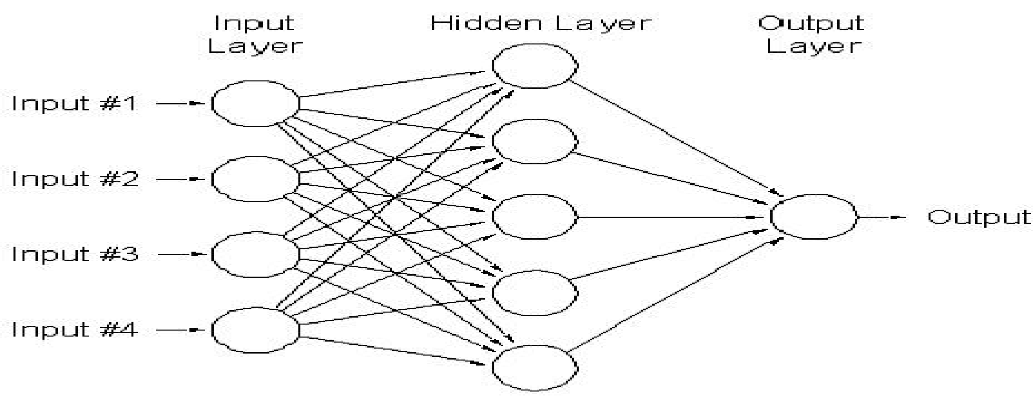


Figure 2: Simple Neural Network

2. The Backpropagation Algorithm

Artificial neural network is based on the principle of layered feed forward which backpropagation algorithm inculcates to address network problem (Rumelhart & McClelland, 1986; Guoqiang, 2000). In contrast, artificial neurons are arranged in layers, transmitting signals "forward" before errors spread backward. Neurons are supplied to the network by the input layer, and the opposite is true for the output layers, which receive neurons from the output network.

The backpropagation training algorithm is defined as half the square of the output error vector's Euclidean norm and is based on the gradient descent principle. The function known as the backpropagation algorithm is used to create elaborate artificial neural networks using the weighted sum, which is the total of the inputs (xi) multiplied by their individual weights w ji):

$$A_j\left(\overline{x}, \overline{w}\right) = \sum_{i=0}^{n} x_i w_{ji} \quad (1)$$

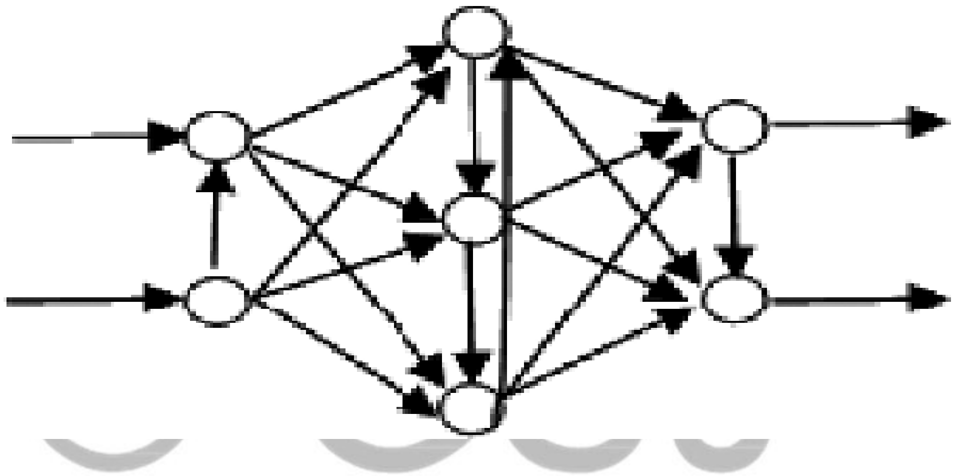
The function that solves the equation is sigmoidal, meaning that for large positive numbers, it is very close to one, for large negative numbers, it is very close to 0.5, and so on. This enables a seamless change from the neuron's low to high output (near zero or near one). It is evident that the activation alone determines the output, which is reliant on the input values and weights. It is evident that the process of training aims to produce a desired result given a set of inputs. The discrepancy between the actual and intended output is known as the error; it is contingent upon the weights and must be minimized by adjusting the weights.

$$E_j\left(\overline{x}, \overline{w}, d\right) = \left(O_j\left(\overline{x}, \overline{w}\right) - d_j\right)^2 \quad (3)$$

Here, the sigmoid function explains the discrepancy between the desired function and the output square. As a result, the backpropagation algorithm now determines how the inputs, weights, and outputs affect the error. Once this is determined, we can use the gradient descendent method to modify the weights. The activation function is applied in logistic function to calculate derivatives in neural network. Other application areas are hyperbolic tangent, arctangent weight determination, supervised training set, learning rate, hidden neuron, low pass filter and tends to aid convergence (Vincent &Kevin, 2002).

3. Lateral Networks

The lateral network combines feedback and forward networking. Different neurons appear to be coupling within a single layer, and there isn't a clear feedback path connecting the various layers (Vincent & Kevin, 2002).



Example of lateral network

2.2.3 **Types of Neural Network**

2.2.3.1 **Single-layer ANNs,**

As the name suggests, a single layer consists of a single input layer, a single neural layer, and a single output layer. Data from the Hopfield network showed that signals flowed from the input layer to the output layer in a single direction, or unidirectionally. Networks architecture belonging to the single layer proved that the number of network outputs will always coincide with its number of neurons. These networks are usually employed in pattern classification and linear filtering problems. Among the main network types belonging to feedforward architecture are the perceptron and the Adaline, whose learning algorithms used in their training processes are based respectively on Hebb's rule and Delta rule (Silva et al., 2017).

2.2.3.2 **Multilayer Feedforward ANNs,**

One typical backpropagation, functional link, and product unit is multilayer feedforward. A feed-forward network with a single hidden layer and a finite number of neurons can approximate continuous functions on compact subsets of the universal set, according to a universal approximation in mathematical theory of artificial networks, provided that the activation function is mildly assumed. Thus, the theorem asserts that, given suitable parameters, simple neural networks can represent a wide range of interesting functions; it makes no mention of the algorithmic learnability of those parameters (Hornik et al., 1989).

Hornik et al., (1991) demonstrated that the multilayer feedforward architecture itself, rather than the activation function selection, is what allows neural networks to potentially serve as universal approximators. It is always assumed that the output units are linear. Just the one output case will be displayed for notational ease. From the single output case, the general case can be inferred with ease. According to a synopsis of the theory, it called universal approximation bounds for sigmoidal function super-positions sensibly (Barron & Cover, 1993). Achieved squared error of approximation is obtained by integrating an artificial neural network with one layer of sigmoid nodes on a bounded subset of variables.

2.2.3.3 Theoretical Properties of Multilayer Feedforward Networks

1. Standard multilayer feedforward networks can approximate any measurable function to any desired level of accuracy, making them universal approximators.
2. Theoretically, feedforward networks can succeed in any situation.
3. Inadequate learning, a deficiency in hidden units, or an absence of a deterministic relationship between input and target are the reasons for failure.
4. Convergence rate as the number of hidden units increases

5. Rate of increase of the number of hidden units for a fixed accuracy as the input dimension grows (Hornik et al., 1991).

2.3 **Application of Neural Network**

Numerous fields were covered by the application, including content addressable memory, clustering/categorization, function approximation, prediction/forecasting, optimization, matching/control, completion of patterns, massive parallelism, distributed representation, learning/generalization ability, data mining, noise reduction, fault tolerance, and times series modeling (Vincent & Kevin, 2002).

2.4 Support Vector Machine (SVM)

The generalization of data modeling caused a setback for neural networks historically (Hadamard, 1923) and Vapni et al., 1995) laid the groundwork for Support Vector Machines (SVM). The structural risk minimization (SRM) principle, which is embodied in SVM, is a new classification technique that is gaining traction in empirical evidence. Research has demonstrated that SRM is superior to the traditional empirical risk minimization (ERM) principle used by conventional neural networks (Gunn et al., 1997). Another benefit of optimization algorithms is the selection of goodness of fit otherwise called the best model. SVM prevents both underfitting (a tiny training error) and overfitting (a small testing error), which helps to clarify this. In order to achieve a goal in statistical learning and generalization, vector training is in fact another one of its characteristics (Hadamard, 1923).

Regression problems have been added to the domain of support vector machines (SVMs) recently. According to Vapnik et al. (1997), it is a sophisticated computational technique for resolving numerical issues where basic formulas are unable to generalize. A linear or nonlinear separating surface in the dataset's input space is used to classify data. The quadratic program solving strategy is the foundation of the algorithm (Mangasarian & David, 2017).

The algorithm needs to be applied to split the dual variable into basic and non-basic variables. Zero-valued variables are regarded as non-basic variables. The values of the variables are obtained by finding the gradient of the objective function to zero and solving the resulting linear equations for the fundamental variables. Remember that when you solve the linear equations, a basic variable gets set to zero and loses its basic status if it has a negative value. This is basically the linearization algorithm. The More-Toraldo finite equation is introduced to make the algorithm computational and use the SMW formula in order to make it converge and terminable.

2.4.1 **Architecture of SVMs**

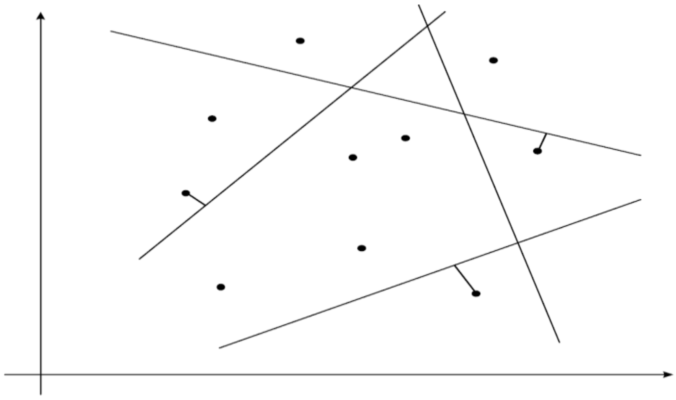
Support vector machine architecture consists of the following's conditions; thus, they are:

1. Formulated for two classification problems
2. Map the input space into feature space
3. Determine the optimal hyper plane in the feature space

A vector diagram that divides two training vectors from related classes is called an optimal hyper plane. Vapnik (1995) states that the set of vectors is separated by the hyper plane optimally if and only if the distance between the vectors that are closest to the hyper plane is maximal and the separation is error-free. Some of the equations were found to be redundant by observations, and it is proposed that they have a conical hyper plane with constrained parameters that maintains generality. Direct parameterization constraints should be preferred over simple formulation problems.and apply the SMW formula.

2.4.2 **Types of SVMs**

Two distinct types of Support Vector Machines are identified in the study. They are classified under hard and soft margin. The hard margin SVMs are linearly separable in feature space. It maximizes problem through generalization ability while the soft margin is not separable in the feature space, minimize classification error and maximize generalization ability (Steve, 1998).



Canonical Hyperplanes

2.4.3 Application of SVMs

1. Initially, the binary classes are separated using the support vector machine (SVM) and the maximized margin criterion. It would be more accurate to say that discrimination for more than two categories is necessary when solving real-world problems. However, there are many applications for multi-class recognition, including bioinformatics, optical character recognition, handwriting recognition, intrusion detection, and speech recognition.
2. It is also applicable to error correcting output codes (ECOC), where a good ECOC matrix is rebuilt to eliminate issues with one versus rest and one versus one.
3. Lastly, the use of a single optimization processing is applied to multi-class problems. According to Szedmak et al. (2004), these models simultaneously accomplish the classification of multiple classes by combining multiple binary-class optimization problems into a single objective function.

2.5 Binary **Encoding**

Depending on whether a band is above or below the spectrum mean, the binary encoding classification method encodes the data and endmember spectra as ones or zeros. It goes without saying that the function creates a classification image by comparing each encoded reference spectrum with the encoded data spectra. Unless you specify a minimum match threshold, in which case some pixels may be unclassified if they do not meet the requirements, all pixels are classified to the endmember with the greatest number of bands that match (Mazer et al., 1988). Rastegari et al., (2016) discussed it as a unified framework of matrix factorization and integer decomposition. Three integer values (-1, 0, +1) are associated with the weight matrix when factorized to exploit low rank signatures. Binary encoding is however described as non-parametric classification where objects are described by binary and nominal attributes. The principle guiding non-parametric (instance-based) learnings to keep the whole learning set as a basic for decision making against metric rule (Duda et al., 2001). A new, efficient, and low-computation method is developed to simplify image classification problems. Mazer et al., (1988) also used this analytical technique in their classification process to find similar spectra and use hyperspectral images to identify the different mineral components. In 1979, Viterbi and Omura conducted a similar study to quantify spectral signature matches through distance decay.

2.5.1 Applications of Binary Encoding

1. It is used for scalable speech recognition or rapid speech content retrieval.
2. Prediction and normalization for component patches

3.0 Materials and Method

3.1 Material used

The study employed the various software such as ENVI 4.5, Idrisi 16.0 and ArcGIS 10.7.1 for mapping, training of dataset and map documentation.

Table 1: Dataset used

Data Name	Date	Property	Source	Purpose
Landsat	12/01/2000	28.5m	Global Land Cover Facility (GLCF)	It is used for the extraction Land cover
Administrative map Rivers State	2011	1:1000	Office of the Surveyor General Rivers State	To create a shapefile of the study area

Source: **(Author, 2023)**

3.2 METHODOLOGY

The methods involved the creation of shapefiles to extract administrative information from Rivers State administrative map. Landsat imagery was used to create land cover and land use map to understand the features within the study area after supervised classification such as Binary Encoding, Support Vector Machine, and Mahalanobis Distance.

3.3 Study Area

The Port Harcourt City Local Government Area in Rivers State is the study area. The study area, which encompasses 32,781 Ha of land, is the most densely populated area in the state with high water and sanitation stress. Situated in Nigeria's Niger Delta, the Port Harcourt Shoreline runs parallel to the Bonny River. The Bonny River experiences semi-diurnal tides. It is situated between longitudes 6º 57' 30'' E and 7º 09' 10'' E and latitudes 4º43' 20'' N and 4º56' 40'' N. The capital of Rivers State, Port Harcourt serves as the hub of the Nigerian oil and gas industry, the home of the Port Harcourt City Local Government Area, and the treasure base of Nigeria. Built in 1912, the Sea Port in Port Harcourt was not given its current name until 1913, when Lewis Vernon Harcourt, the colonies' Secretary of State at the time, was honoured with the name Port

Harcourt by Sir Fredrick Lugard.

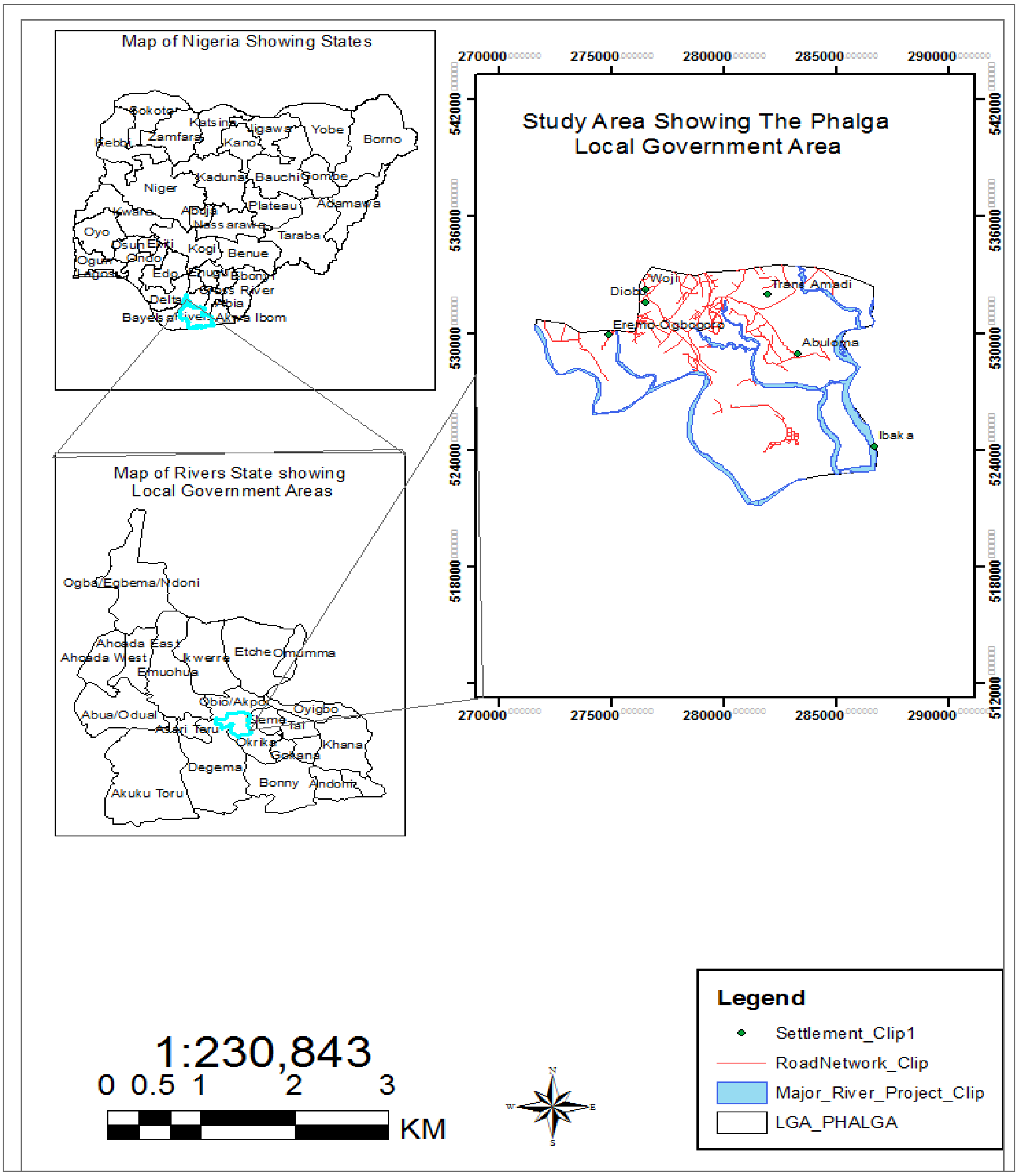


Figure 1: The Study Area

4.0 RESULTS AND DISCUSSION

Table 2 explains the class statistics for support vector operation where five classes were trained. They are River, light Vegetation, Mangrove, Built up and unclassed area. River body contained

15,733,282.50 meters square which give rise to 8.277 percent. Light vegetation and mangrove also appropriated 7.126% and 17.771% respectively. Built up in the same manner have 23.876% with a total square meter of 45,385,281.00. This result became necessary as a result of the various band analyses. Table 3 illuminated 3.8609 mean values while 1.271941 goes to the standard deviation for band 1.

Table 2: Class Statistics for Support Vector Machine

CLASS	NPTS	AREA (M^2)	PERCENTAGE
River	19,370	15,733,282.50	8.277
Light Vegetation	16,676	13,545,081.00	7.126
Mangrove	41,589	33,780,665.25	17.771
Built up	55,876	45,385,281.00	23.876
Un-classed	100,519	81,646,557.75	42.951

Source: **(Author, 2023)**

Table 3: Basic Statistic for Band formation

Basic stat	Minimum	Maximum	Mean	St Deviation
Band 1	1	5	3.860992	1.271941

Source: **(Author, 2023)**

The statistical categorization of support vector analyses, also referred to as an error matrix, is specifically shown in Table 4. The algorithm's performance—usually a supervised learning algorithm—can be seen in the table below. The predicted class was represented by each column in the matrix, and the actual class was represented by each row. The name of the classes makes it confusing as it appears on both predicted and actual in the contingency table. The algorithm was so trained to distinguish the given classes. The overall accuracy showed 94.72% while kappa coefficient (92.55%).

Table 4: Error Analysis for SVM classification

Class	Un-classified	River	Light Vege	Mangrove	Built up
Un-classified	0	0	0	0	0
River	0	1966	0	40	19
Light Vege	0	0	2944	26	135
Mangrove	0	90	50	3259	8
Built up	0	17	392	6	5882
Total		2073	3386	3331	6044

Source: (Author, 2023)

Overall Accuracy = (14051/14834) = 94.7216%

Kappa Coefficient = 0.9255

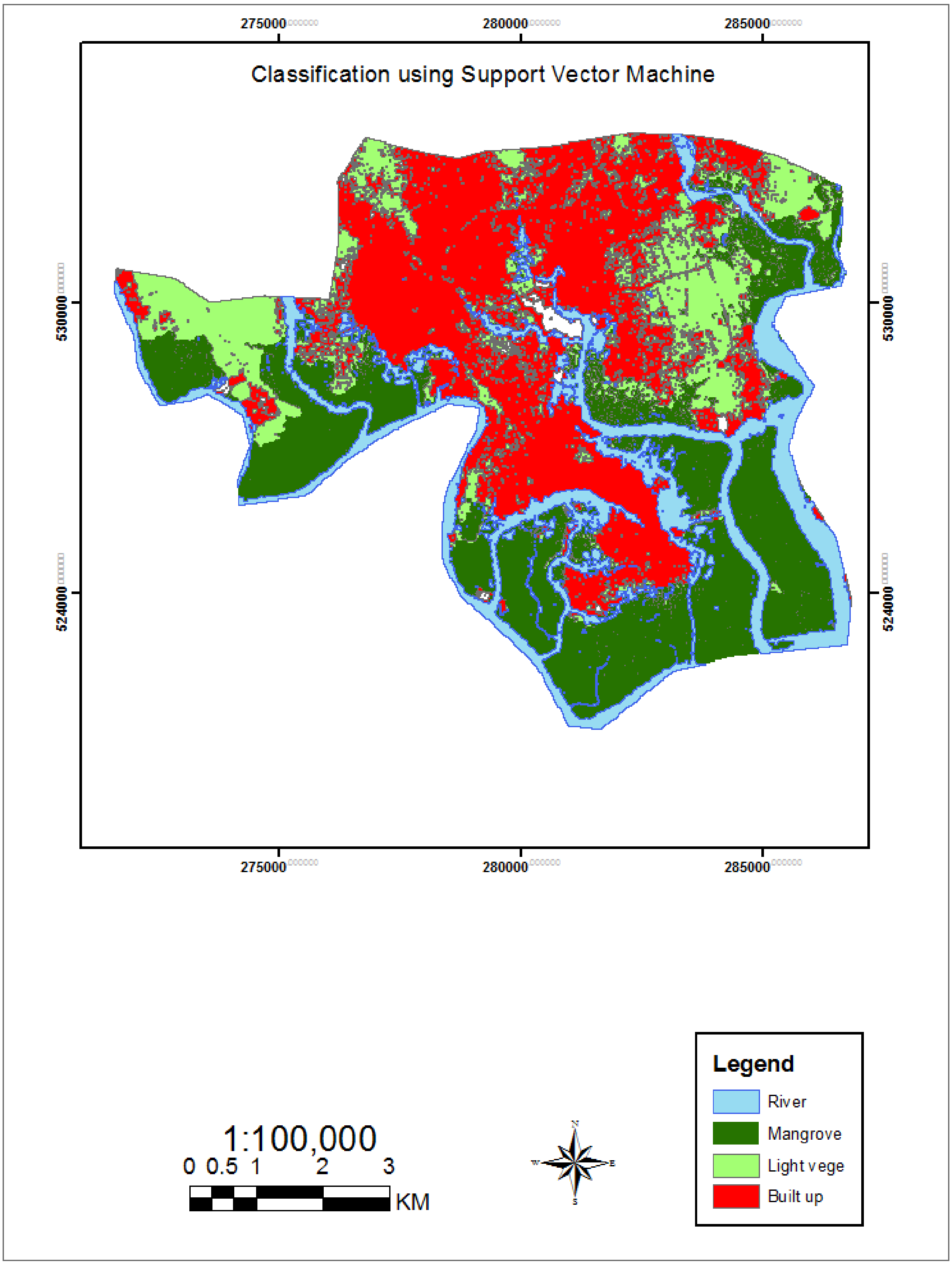


Figure 2: Classified Image using Support Vector Machine

Table 5 is a Binary Encoding (BE) where different colours were assigned to different class components to perform supervised classification. These colours enables expert to know what is what in the class distribution. River is red, light vegetation (green), mangrove takes blue, built up (yellow) and the unclassed unavoidably was juxtaposed cyan. These contravene cartographic symbolism because BE does not vectorised. Mean and standard deviation (SD) were calculated in (Table 6). The mean revealed 5.2193 while (SD) showed an increment of 8.8338.

Table 5: Colour Code for Binary Encoding

CLASS	**COLOUR**
River	Red
Light Vegetation	Green
Mangrove	Blue
Built up	Yellow
Un-classed	Cyan

Source: **(Author 2017)**

Table 6: Basic Band Statistic for Binary Encoding

Basic stat	Minimum	Maximum	Mean	St Deviation
Band 1	0	100	5.219356	8.833883

Source: **(Author 2017)**

Binary Encoding in Table 7 gives the class summary where number of pixel points (NPTS), area and percentage were used to define the classes appropriately. River (10.226%), light vegetation (7.117%), mangrove (17%), built up (21.149) and un-classed (41.176%). An error matrix comparing a classification with ground truth data is shown in Table 8 to assess how well the classification captures reality. The values of the un-leading diagonal were used to calculate error of omission. Results indicated that overall accuracy (81.259%) and the corresponding kappa coefficient (73.73%).

Table 7: Class Statistics for Binary Encoding

CLASS	NPTS	AREA (M^2)	PERCENTAGE
River	23,932	19,438,767.00	10.226
Light Vegetation	16,657	13,529,648.25	7.117
Mangrove	40,014	32,501,371.50	17.098
Built up	49,494	40,201,501.50	21.149
Un-classed	96,365	78,272,471.25	41.176

Source: **(Author, 2023)**

Table 8: Error Analysis for Binary Encoding Classification

Class	Un-classified	River	Light Vege	Mangrove	Built up
Un-classified	0	0	0	0	0
River	0	1929	66	25	290
Light Vege	0	8	2320	1041	53
Mangrove	0	128	515	2265	159
Built up	0	8	489	0	5549
Total	0	2073	3390	3331	6051

Source: **(Author, 2023)**

Overall Accuracy = (12063/14845) = 81.2597%

Kappa Coefficient = 0.7373

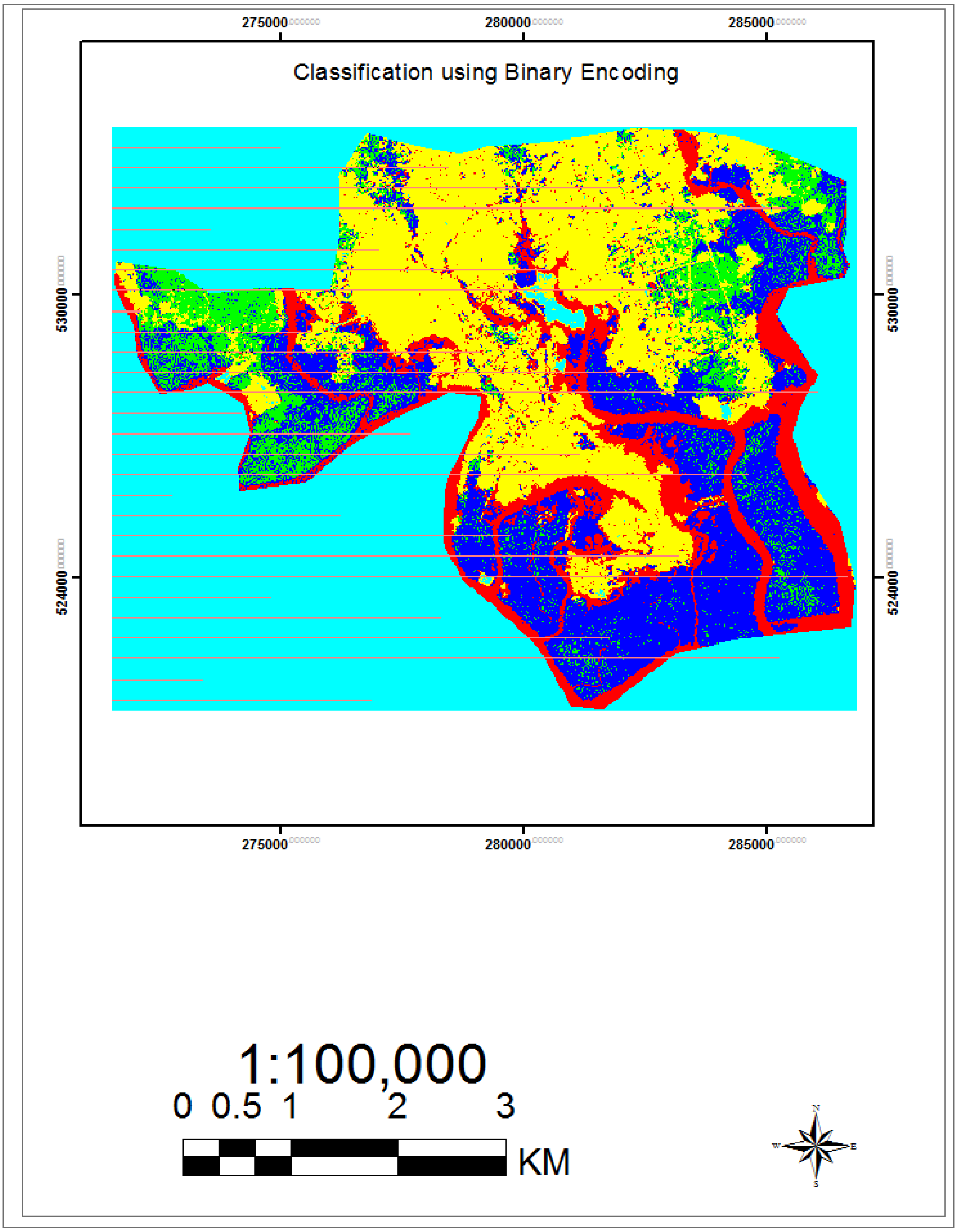


Figure 3: Classified Image using Binary Encoding

In a bit to ascertain the level of accuracy in the classification, mahalanobis distance supervised operation was further carried out to compare with other classification techniques. River reflects

(10.549%), light vegetation (8.731%), mangrove signify 14.798%, others are built up and un-classed equally have 18.913% and 47% respectively.

Machine learning, linear algebra, and theoretical computer science all use matrices. It was adopted to decompose five classification processes as contained in the tables (9 &10). The reason behind this was that kernel-based statistical learning computations improved the interpretability of data analysis methods. Therefore, the overall accuracy assessment of result was 88.22% while kappa index recorded 83.20%. Table 11 showed a comparative analysis of the study.

Table 9: Basic Statistics for Mahalanobis Distance

Class	**NPTS**	**AREA (M^2)**	**PERCENTAGE**
River	24688	20,052,828.00	10.549
Light Vegetation	20,433	16,596,704.25	8.731
Mangrove	34,631	28,129,029.75	14.798
Built up	44,263	35,952,621.75	18.913
Un-classed	110,015	89,359,683.75	47.009

Source: **(Author, 2023)**

Table 10: Error Analysis for Mahalanobis Distance

Class	Un-classified	River	Light Vege	Mangrove	Built up
Un-classified	0	0	0	0	0
River	0	2118	1	23	24
Light Vege	0	7	1817	1007	304
Mangrove	0	23	478	3883	5
Built up	0	3	54	1	6646
Total	0	2151	2350	4914	6979

Source: (Author, 2023)

Overall Accuracy = (14464/16394) = 88.2274%

Kappa Coefficient = 0.8320

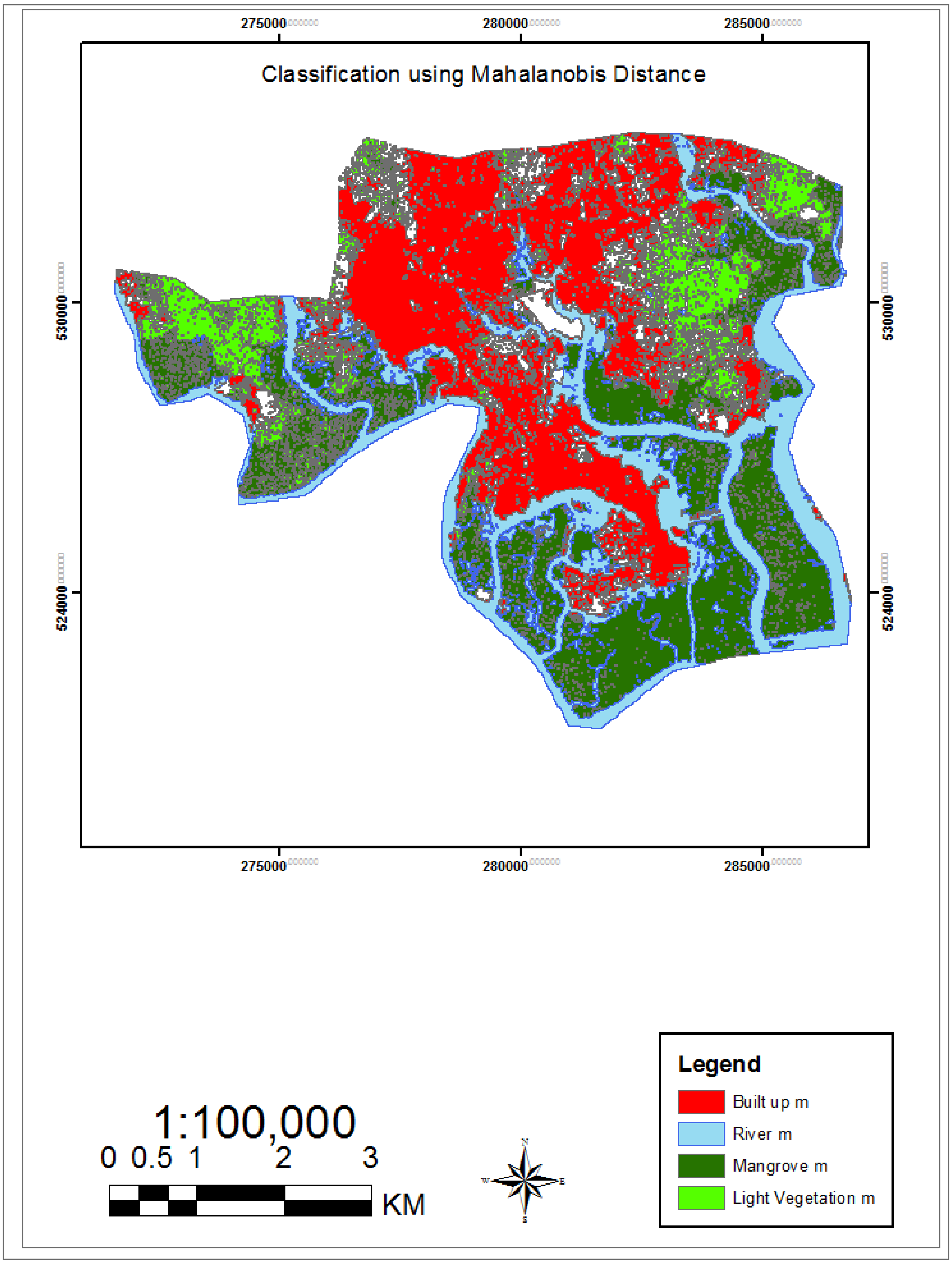


Figure 3: Classified Image using Mahalanobis Distance

Table 11: Comparative Analysis of the three Selected Methods

CLASS	SVM (%)	BE (%)	MAH (%)
River	8.277	10.226	10.549
Light Vegetation	7.126	7.117	8.731
Mangrove	17.771	17.098	14.798
Built up	23.876	21.149	18.913
Un-class	42.951	41.176	47.00
KA	94.72	82.259	88.22

Source: **(Author, 2023)**

5.0 Recommendations

The study has shown that there is no single method that clearly defined all methods of classification.

1. As a result, this study suggests developing classification systems that use a variety of classification algorithms
2. Secondly, Support Vector Machine and Binary Encoding are efficient and effective methods for image classification and should be widely used

6.0 Conclusions

A review of the literature was done on the more sophisticated approaches to classification, including binary encoding, support vector machines, and artificial neural networks. SVM is most appealing property where sigmoid function was applied to make high generalization classification. These methods and mahalanobis distance were widely applied for image classification process. The outcomes demonstrated that every classification technique has advantages and disadvantages. Table 10 deduces the information extracted from the classification. The study showed that Support vector machine and binary Encoding were consistent in their results. It would be recalled that River in SVM (8.277%), Binary Encoding (10.226%), Mahalanobis Distance (10.549). Light Vegetation depicted 7.126% in (SVM), BE

(7.117%) and MAH (8.731%). The mangrove species also recorded 17.771% and 14.798% respectively as contained in the table. Built up for SVM and BE are 23.876% and 21.149% while MAH is 18.913%. SVM and BE are valuable tool for un-class classification, both recorded 42.951% and 41.176% against 47% for MAH. Kappa Coefficient is measure of accuracy where SVM (94.72%), BE (82.259%) and MAH have 88.22%. The analytical study does not extend to artificial neural networks because software extension was supported in Envi 4.5 to justify real time kinematics study.

REFERENCES

Anderson, J.R., Hardy, E.E., Roach, J.T. and Witmer, R.E. (1976). A Land Use and Land Cover Classification System for use with Remote Sensor Data, U.S. Geological Survey Professional Paper, No. 964, USGS, Washington, D.C.

Andreassen, H., Bohr, H., Bohr, J., Brunak, S., Bugge, T., Cotterill, M.J., Jacobsen, C., Kusk, P., Lautrup, B., Petersen, S.B., Saermark, T. & Ulrich, K. (1990). Analysis of the Secondary Structure of the Human Immuno Deficiency virus (HIV) proteins p17, gp 20, and gp41 by computer modelling based on neural network methods. J. Acq. Immun. Defic. Syndrome, 3: 615–622.

Belliustin, N.S., Kuznetsov, S.O., Nuidel, I.V. &Yakhno, V.G. (1991). Neural Networks with Close Nonlocal Coupling for Analysing Composite Image. Neurocomputing, 3: 231–246.

Barron, A. R., & Cover, T. M. (1991). Minimum Complexity Density Estimation. IEEE Transactions on Information Theory, 37, 1034-1054.

Butera, M.K. (1983). Remote Sensing of Wetlands, IEEE Transactions on Geoscience and Remote Sensing GE-21, pp.383-392.

Campbell, L. (2000). Introduction to Remote Sensing, CORINE Land Cover Technical Guide, European Commission, Luxemburg, pp.21-53.

Campbell, L. (2001). Flow Cytometric Analysis of Autotrophic Picoplankton, Methods in Microbiology: Marine Microbial Ecology, 30, Academic Press, 317–343,

Cybenko, G. (1989). Approximations by Superpositions of a Sigmoidal Function. Mathematics of Control, Signals, and Systems, 2, 303-314.

Dean, A.M and Smith, G.M. (2003). An Evaluation of Per – parcel Land Cover Mapping using Maximum Likelihood Class Probabilities, International Journal of Remote Sensing, 24 (14), pp. 2905-2920.

David, P. L. (2015). Digital Image Processing and Analysis

Duda, R.O., Hart, P.E., &Stork, D.G. (2001). Pattern Classification. John Wiley, New York,

Ernst, C.L. and Hoffer, R.M. (1979). Digital processing of remotely sensed data for mapping wetland communities. LARS Technical Report 122079. Laboratory for Applications of Remote Sensing. Purdue University, West Lafayette, 119 pp.

Figueredo, A.J., Ross, D.M. and Petrinovich, L. (1992). The Quantitative Ethology of the Zebra Finch: a study in comparative psychometrics. Multiv. Behav. Res., 27 (3): 435–458.

Fiona, L. (2011). Neural Networks – algorithms and applications.

Furferi, R., Governi, L., & Y. Volpe, Y. (2011). Neural Network Based Classification of Car Seat fabrics. International journal of mathematical models and methods in applied sciences, Vol.3, Issue.5,

Fukushima, K. and Wake, N. (1990). Alphanumeric Character Recognition by the Neocognitron. In: R. Eckmiller (Editor) Advanced Neural Computers, Elsevier Science, Amsterdam, pp. 263–270.

Gemello, R. and Mana, F. (1991). A Neural Approach to Speaker Independent Isoled

Gunn, S.R., Brown, M & Bossley, K.M. (1997). Network Performance Assessment for Neuro Fuzzy Data Modelling. In X. Liu, P. Cohen, and M. Berthold, editors, Intelligent Data Analysis, volume 1208 of Lecture Notes in Computer Science, pages 313–323.

Guoqiang, P.Z. (2000). “Neural Networks for Classification: A Survey “,IEEE transactions on systems, man, and cybernetics , Vol.30, Issue.

Hadamard, J. (1923). Lectures on the Cauchy Problem in Linear Partial Differential Equations.Yale University Press, 1923.

Hornik, K., Stinchcombe, M., White, H. (1989). Multilayer Feed Forward Networks are Universal Approximators. Neural Networks **2,** 359‑66.

Igor, F. (2012). Image Classification using Artificial Neural Networks.

Iwata, A., Tohma, T., Matsuo, H. and Suzumura, N. (1990). A Large-scale Neural Network ‘CombNET’ and its application to Chinese character recognition. In: Proc. Int. Neural Networks Conf., Paris July 9–13, 1990, Kluwer Academic Publishers, Dordrecht, and Vol. 1, pp. 83–86.

Jensen, J. R. (1996). Introductory Digital Image Processing, A Remote Sensing Perspective, 2nd, Prinice Hall, 318 p.

Julius T. and Raphael C. G. (1974). Pattern Recognition Principles. Addison-Wesley Publishing Co.

Kandel, E.F., Schwartz, J.H, Jessell, T.J. (1999). Principles of Neural Science, 3rd edn. Amsterdam: Elsevier.

Kuffler, S.W., Nicholls, J. G., & Martin, A.R. (1984). From Neuron to Brain: a cellular approach to the function of the nervous system, 2nd edn. Sunderland, MA: Sinauer Associates.

Lek, S., Lauga, J., Giraudel, J.L., Baran, P. & Delacoste, M. (1994). Example of an Application of Formal Neural Networks in the Environmental Sciences. In: V. Bre´tagnolle and F. Beninel (Editors), Proc. Int. Meeting 'Ecology and statistical methods', C.N.R.S., Chize´, pp. 73–82.

Lek, S., Belaud, A., Lauga, J., Dimopoulos, I. & Moreau, J. (1995). An Improved Estimation of the Food Consumption of Fish Populations using Neural Networks. Aust. J. Mar. Freshwater Res., 46 (8): 1229–1236.

Le Pape, G. & Chevalet, P. (1992). Description des donne´es d'observation continues du comportement par une technique d'analyse de texte. 2: comparaisons des conduits maternelles dans trois souches de Souris. Behav. Proc., 26(1): 23–30.

Lillesand, T.M. & Keifer, R.W. (1994). Remote Sensing and Image Interpretation.

Lillesand, T.M & Keifer, R.W. (2000). Remote Sensing and Digital Image Interpretation, Wiley, New York, 724p

Liu, X.H., Skidmore, A.K. & Osten, V.H. (2002). Integration of Classification Methods for Improvement of Land-cover Map Accuracy, ISPRS Journal of Photogrammetry & Remote Sensing, 56, pp 257-268.

Lo, C.P., & Watson, L.J. (1998). The Influence of Geographic Sampling Methods on Vegetation Map Accuracy Evaluation in a Swampy Environment. Photogrammetric Engineering and Remote Sensing 64:1189-1200.

Mangasarian, O.L & David, R.M. (2017). Active Support Vector Machine Classification
Martindale, S. (1980). Multivariate analysis of avian vocalizations. J. Theor. Biol., 83: 107–110.

Maravall, D., Rfos, J., Pe´rez-Castellano, M., Carpintero, A. & Gı¨mez-Calcerrada, J. (1991). Comparison of Neural Networks and Conventional Techniques for Automatic Recognition of a Multilingual Speech Database. In: A. Prieto (Editor), Artificial Neural Networks, Proc. Int.Workshop IWANN'91, Granada (SPAIN), Sep-91, pp. 377–384.

Mazer, A.S., Martin, M., Lee, M., & Solomon, J.E. (1988). Image Processing Software for Imaging Spectrometry Data Analysis, Remote Sensing of Environment v. 24, pp. 201-210.

Mirza, C. (2017). Neural Networks and Back Propagation Algorithm. Institute of Technology Blanchardstown, Blanchardstown Road North.

Muller, J., Reinhardt, & Strickland, M.T. (1995). Neural Networks: An introduction. Number v. 1 in Physics of neural networks. Springer, 1995. ISBN 9783540602071.

Nicolas, J.M., Lemer, A. & Legitimus, D. (1989). Identification Automatique De bruit Impulsifs en Acoustique Sous-marine Par Re´seaux Multicouches. In: Neuro-Nıˆmes'89, Int. Workshop Neural Networks and their applications, Nıˆmes (FRANCE) November 13–16, 1989, EC2, Nanterre, pp. 269–278.

Omatu, S., Fukumi, M. & Teranisi, M. (1990). Neural Network Model for Alphabetical Letter Recognition. In: Proc. Int. Neural Networks Conf., Paris July 9–13, 1990, Kluwer Academic Publishers, Dordrecht, and Vol. 1, pp. 19–22.

Ozesmi, S.L & Bauer, M. (2002). Satellite Remote Sensing of Wetlands, Wetlands Ecology and Management, 10, pp. 381-402

Pal, M. & Mather, P. (2005). Support Vector Machines for Classification in Remote Sensing. International Journal of Remote Sensing, vol. 26, no. 5, pp. 1007-1011.

Qian, N., & Sejnowski, T.J. (1988). Prediction of the Secondary Structure of Globular Proteins using Neural Networks Models. J. Mol. Biol., 202: 865–884.

Rastegari, M., Vicente, O., Joseph, R., & Ali, F. (2016). ImageNet Classification using Binary Convolutional Neural Networks. ECCV, pp. 525–542

Randall, B. S. (2011). Image Classification. Micro-Images.

Richards, J.A & Jia, X. (1999). Remote Sensing Digital Image Analysis. Springer-Verlag.

Rumelhart, D.E., & McCllelland, J.L. (1986b). The PDP Research Group. Parallel Distributed Processing, vol. 1, ch. 9. Cambridge, MA: MIT Press (Bradford Books).

Saravanan, K.1 & Sasithra, S. (2014). Review on Classification Based on Artificial Neural Networks

Silver, I.N., Harnane, S.D., Andrade, F.R., Liboni, L.H.B., & Reis, A.S.F. (2017). Artificial Neural Network a Practical Course

Sten, S & Tambet, M. (2015). Artificial Neural Network for Image Classification

Smits, J.R.M., Breedveld, L.W., Derksen, M.W.J., Kateman, G., Balfoort, H.W., Snoek,

J. & Hofstraat, J.W. (1992). Pattern Classification with Artificial Neural Networks: Classification of Algae based upon Flow Cytometer Data. Anal. Chim. Acta, 258: 11–25.

Sunitha, A & Suresh, B. G. (2015). Satellite Image Classification Methods and Techniques: A review. International Journal of Computer Applications (0975 – 8887).

Szedmak, S., Shawe-Taylor, J., Saunders, C., Hardoon, D. (2004). Multiclass Classification by l1 Norm Support Vector Machine. In: Pattern Recognition and Machine Learning in Computer Vision

Terhune, J.M., Burton, H., & Green, K. (1993). Classification of Diverse Call Types using Cluster Analysis Techniques. Bioacoustics, 4: 245–258.

Tirakis, A., Sukissian, L., & Kollias, S. (1990). An Adaptive Technique for Segmentation and Classification of Textured Images. In: Proc. Int. Neural Networks Conf., Paris, Vol. 1, pp. 31–34.

Thompson, R.F. (1993). The Brain: A neuroscience primer. New York: Freeman.

Van Allen, E., Menon, M.M., & Dicaprio, N. (1990). A Modular Architecture for Object Recongnition using Neural Networks. In: Proc. Int. Neural Networks Confer., Paris July 9–13, 1990, Kluwer Academic Publishers, Dordrecht, Vol. 1, pp. 35– 37.

Vapnik, V. (1995). The Nature of Statistical Learning Theory. Springer, N.Y., ISBN 0-387-94559-8.

Vapnik, V., Golowich, S & Smola, A. (1997). Support Vector Method for Function Approximation, Regression Estimation, and Signal Processing. In M. Mozer, M. Jordan, and T. Petsche, editors, Advances in Neural Information Processing Systems 9, pages 281– 287, Cambridge, MA. MIT Press.

Vincent, C & Kevin, C. (2002). An Introduction to Neural Networks.

Williams, R.J. (1987). Reinforcement-learning Connectionist Systems. Technical Report NU CCS-87–3, Northeastern University, Boston.

Yu-guo, W & Hua-peng, L. (2010). Remote Sensing Image Classification Based on Artificial Neural Network. International Conference on Computer, Mechatronics, Control and Electronic Engineering (CMCE), Vol.1, Issue.2