Available online at www.CivileJournal.org

Civil Engineering Journal

(E-ISSN: 2476-3055; ISSN: 2676-6957)

Vol. 10, No. 10, October, 2024

Comparative Study of Different Classification Methods and Winner Takes All Approach

Khaled Mahmoud Abdel Aziz [1, 2*], Loutfia Elsonbaty [2]

[1] Department of Civil Engineering, College of Engineering in Al-Kharj, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia.

[2] Geomatics Engineering Department, Faculty of Engineering at Shoubra, Benha University, Cairo, Egypt.

Received 08 July 2024; Revised 18 September 2024; Accepted 24 September 2024; Published 01 October 2024

Abstract

One of the most popular methods in remote sensing for gathering and evaluating satellite data is the classification of images. Several categories exist for image classification techniques, including supervised and unsupervised classification, pixel-based, object-based, and rule-based approaches. Each type of technique has pros and cons of its own. Choosing the method that produces the best results is one of the issues with image classification. The "best" model for classifying images relies on the particular task and the dataset used. The ideal classification technique is a crucial component in increasing classification accuracy. The strengths and drawbacks of various models vary, so selecting one that is appropriate for the job is critical. The main objective of this research is to analyze and compare the results of each classifier used, including ISODATA, K-mean, Maximum likelihood, Minimum distance, Support vector machine, and Neural network then integrate these different types of classification using the winners-takes-all classification approach in order to try to improve the results. The classified images were assessed, and both the overall accuracy and kappa coefficient were calculated and gave 79.50%, 73.89%, 77.05%, and 84.98%, 86.53%, 87.18%, and 88.69% for ISODATA, K-means, Minimum distance (MD), Maximum likelihood (MXL), Support vector machine (SVM), Neural network (NNT), and winner takes all (WTA), respectively. From the results, the Winner takes all (WTA) presented a superior in terms of the overall accuracy and kappa coefficient.

Keywords: Classification; Kappa Coefficient; Error Matrix; Winner Takes All Classification.

1. Introduction

Image classification has many uses in many different domains. For several reasons, the classification of aerial and satellite images is crucial. It aids in the identification of various land cover types, including forests, aquatic bodies, urban areas, and agricultural fields. This information is essential in many other fields of applications, such as resource management, environmental monitoring, and urban planning. Image classification is rapidly evolving due to several key factors, and continuous research is leading to the development of more efficient and accurate algorithms. Techniques like transfer learning, data augmentation, and ensemble methods are enhancing model performance. One of the most important stages of information extraction in satellite image analysis is the classification of items that appear in an image. Agricultural challenges, natural disasters, environmental monitoring, and other remote sensing issues can all be solved with the use of the information gathered from the satellite image. When it comes to realistic classification with high accuracy, the majority of conventional classifiers used to categorize satellite images fall short. As a result, classifying the satellite images in a realistic and precise manner remains a difficult endeavor [1].

* Corresponding author: khaled.mahmoud@feng.bu.edu.eg

http://dx.doi.org/10.28991/CEJ-2024-010-10-016

Extracting and interpreting meaningful information from massive datasets of aerial or satellite images requires the classification of such images. Satellite imagery categorization is necessary for several sectors, such as spatial data mining [2], effective decision-making, field surveys, information extraction for applications, and disaster management. Classifying pixels in images involves organizing them into meaningful groups. Extracting information from an image is an alternative name for what is called image classification. Selecting the right classification technique for an image is not hard, but the analyst has to take into consideration many factors [3]. Image interpretation, examining both vegetation types and land use area, beside identifying land use in the region all come under the umbrella of the process of image classification [4].

Many pixel-based techniques for classifying land cover were applied to medium/low-resolution multi-spectral data, like Landsat, in the early 1970s. Pixel-based techniques handle each pixel as a separate unit and assign it to a class. This technique deals with each pixel as an individual unit and assigns it to a specific class. Pixels in the same class have similar spectral characteristics to those that exist in other classes. Conventional machine learning approaches, which may be classified into supervised and unsupervised categories, are beneficial for simple pixel-based procedures. Every category has advantages and disadvantages. The most commonly used classifiers in land cover classification are Maximum Likelihood (MXL), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Tree (DT), and Artificial Neural Network (ANN).

The majority of supervised classifiers, also known as parametric classifiers like Maximum Likelihood Classification (MXL), depend on assumptions about the distribution of data. MXL processes effortlessly and quickly and performs well in a variety of land cover classification scenarios [5]. MXL works effectively even with a limited number of training samples since it makes the assumption that the data is distributed [6]. The accuracy of the land cover classification is increased by using prior probabilities. The improvement is not statistically significant because of the characteristics of the remotely sensed dataset and the degree of spectral overlap of class pairs [7]. It comes to light that the most dominated classes may become over-fitted due to MXL. Despite their demonstrated usefulness, parametric classifiers have two significant shortcomings: (1) the distribution of land cover surfaces is highly unpredictable and cannot be well explained by data distribution alone; (2) data with high heterogeneous land covers are often not normally distributed [5]. On the other hand, non-parametric classifiers like SVM, ANN, and DT do not presuppose anything regarding the data's distribution. The goal of SVM is to locate the best feature space border between the classes. It can effectively control over-fitting issues and offer a good generalization. SVM performs better than other classifiers in the majority of cases, according to several studies [8-11], because it can handle complex features even with small training sample sizes. On the other hand, the quality of the training samples can affect SVM.

For supervised techniques, choosing training samples is crucial because, in addition to being as representative as feasible for the whole data set, the samples also need to ensure that there are enough of them to avoid over-fitting. Reportedly, supervised classification needs an adequate number of training data. [12]. For coarse-resolution data, a pixel may include more than one type of land cover—makes selecting training samples for data often challenging, which is named as the mixed pixel problem. Research has indicated that raising the number of training samples can enhance the precision of classification [8]. Several studies find that MLC, SVM, DT, and ANN are vulnerable to mislabeled training data. When training samples are insufficient, it is therefore less effective to classify land cover directly using traditional supervised methods.

Choosing the most suitable classifier algorithm depends on several factors, including the spatial resolution of the image, where the higher-resolution images may require more complex models, while lower-resolution images might be handled by simpler models [5]. The number and type of spectral bands available can affect the choice of the classifier. The feature complexity can also influence the choice of the classifier; the more complexity the features in the image, the more sophisticated a model is needed. It is crucial to understand how the model makes decisions; simpler models, such as decision trees, might be preferred for their interpretability over other complex models like deep neural networks.

Many factors can significantly affect the performance of image classification models; these factors have to be taken into consideration, like data size; larger datasets lead to better model performance as they provide more examples for the model to learn from. Data quality: noisy data can degrade model performance. Class imbalance, where some classes are underrepresented, can lead to biased models that perform poorly on minority classes. Tuning parameters such as learning rate, number of layers, and batch size can greatly affect model performance.

Sowmya et al. (2017) carried out four processing stages in this study: image preprocessing, enhancement, transformation, and classification. Geometric, radiometric, and atmospheric corrections are all included in preprocessing. The tone-mapping algorithm is used for enhancement purposes. The maximum likelihood approach is suggested for classification. This paper also explains object-oriented classification techniques [13]. Sathya & V. Baby Deepa (2017), four supervised classification models are analyzed by K-nearest neighbor, Parallelepiped, Minimum distance, and Maximum likelihood. The cornerstone of classification statistics is training samples [14].

Manohar et al. (2021), this study used effective automated classification of satellite images. Satellite image classification and feature extraction are done using convolutional neural networks. CNN will attain up to 91.3% categorization success rate. The classification accuracy is calculated using the confusion matrix [15]. Murtaza &

Romshoo (2014), in this work, the accuracy and suitability of various classification techniques are examined. The paper addresses three methods: minimum distance, Mahalanobis distance, and maximum likelihood. The thirteen classifications into which images are classified are agro, aqua, vegetation area, barren region, built-up areas, exposed rock region, forest region, horticulture area, grazing fields, plantation area, riverbed region, scrub region, snow region, and water area. IRS datasets are gathered [16].

Altaei & Mhaimeed (2017) conducted a two-phase investigation. The satellite image is encoded for the first phase, and in the second phase, an artificial neural network (ANN) is used to classify the image, which uses the first phase's produced codes as input materials and the identified image areas as the output. 99% classification accuracy is achieved [17]. Neware & Khan (2018), in this study, the researchers addressed about four classes: water lands, urban regions, non-vegetation, and vegetation. They are employed in parallelepiped classification, minimum distance, and spectral angle mapping. The NDVI method is also used, and it provides the best results when identifying vegetation [18]. Pandya & Science (2015), this paper discusses the use of satellite imagery in agriculture. The process is executed in all of MATLAB. The main goal is to develop a model of a vegetation classification system that can gather data from various sources for satellite forecasters, classify vegetation areas using image-processing techniques, and evaluate important forecast policies for decision-makers [19].

Vimala (2020), Unsupervised image classification has been used in this study to categorize the images into several land use categories. There are seven different land use classes: roads, vegetation, water bodies, mines, fallow land, and barren land. The assessment of classification accuracy is calculated using the field knowledge gathered from field surveys. Results: For each class, the resulting accuracy ranges from 83 to 86% [20]. Ouchra et al. (2023) used processing tools such as the Google Earth Engine and LANDSAT 8 OLI datasets in this research. The classifiers used in this study are as follows: Using the Landsat 8 OLI dataset and the GEE platform, the following classifiers were used to map the Moroccan territory: CART, RF, SVM, Gradient Tree Boost, DT, and MD. The zones mapped were Water, Forest, Built-up, Sand, Barren, and Cropped. Because of its extremely high classification accuracy in comparison to the other classifiers utilized, the study demonstrated that the Minimum Distance classifier is the best-performing classifier. This classifier's accuracy is 93.85%, and we were able to determine its kappa coefficient with 0.93%. [21].

Yasin & Kornel (2024), this research provided a thorough analysis of cutting-edge classification techniques, such as Support Vector Machines (SVMs), Classification Trees (CTs), and Artificial Neural Networks (ANNs). Additionally, it provided a comparison of these contemporary approaches with more established ones, highlighting how well each performed concerning specific performance measures when used with satellite data [22]. Ouchra et al. (2024) used Support Vector Machine (SVM), Random Forest (RF), Classification and Regression Trees (CART), Minimum Distance (MD), Decision Tree (DT), and Gradient Tree Boosting (GTB) to assess the effectiveness of six supervised machine learning algorithms in the classification of land cover in Casablanca, Morocco. The findings show that the accuracy of the Random Forest algorithm reaches 95.42%; it performs better than the Support Vector Machine technique, which trails behind with an accuracy of 83% [23].

Nigar et al. (2024), this study uses machine learning and deep learning models to compare how land is classified in District Sukkur, Pakistan. With a Kappa coefficient of 0.90 and an overall accuracy of 91.3%, the machine learning models, including Random Forest, were successful. It correctly identified 40.4% of the area as vegetation, 1.9% as water bodies, 54.8% as barren land, and 2.7% of the region as built-up area [24]. In Ahmadi (2024), the classification of urban LULC was done using optical satellite imagery, namely Landsat-8 and Sentinel-2 multispectral. On the Google Earth Engine (GEE) platform, it employed three distinct machine learning algorithms: random forest (RF), support vector machine (SVM), and classification and regression tree (CART). The findings showed that, when using optical satellite images from Landsat and Sentinel, RF was the best classifier for classifying urban areas [25].

This research aims to evaluate the efficiency and precision of six classifiers algorithms and determine which algorithms perform better and are more appropriate for similar conditions. In order to compare and ascertain the outcomes, the producer accuracy, user accuracy, overall accuracy, and Kappa coefficient are all assessed. Besides that, an ensemble method (WTA) is used to integrate different classifiers to strengthen the advantages of all classifiers.

This research provides a broad perspective on the performance of six classifiers algorithms. This comprehensive approach is valuable for understanding the strengths and weaknesses of each classifier in various scenarios. In this research, several matrices are used to evaluate the results; using the kappa coefficient as a matrix allows for a more robust evaluation of classifier performance, especially in the presence of imbalanced classes. Producer and user matrices offer insights into the reliability of the classification from both the producer’s and the user’s perspectives. Implementing WTA based on the best-performing classifier for each class is an innovative approach. This method leverages the strength of different classifiers, potentially leading to improved overall performance. In addition to that, emphasis on detecting and handling class imbalances ensures more reliable and fair performance. Compared to other recent studies that might focus on a single classifier or a limited set of metrics, this research stands out for its comprehensive and practical approach. By addressing key challenges such as error analysis and class imbalance, this work contributes significantly to the field and offers valuable insights for both researchers and practitioners.

2. Material and Methods

2.1. Study Area

The study area is situated on the University of New South Wales campus in Sydney, Australia, and is around 500 by 500 meters. This metropolitan region has residential buildings, sizable campus buildings, major and small roads, trees, and green spaces, as seen in Figure 1.

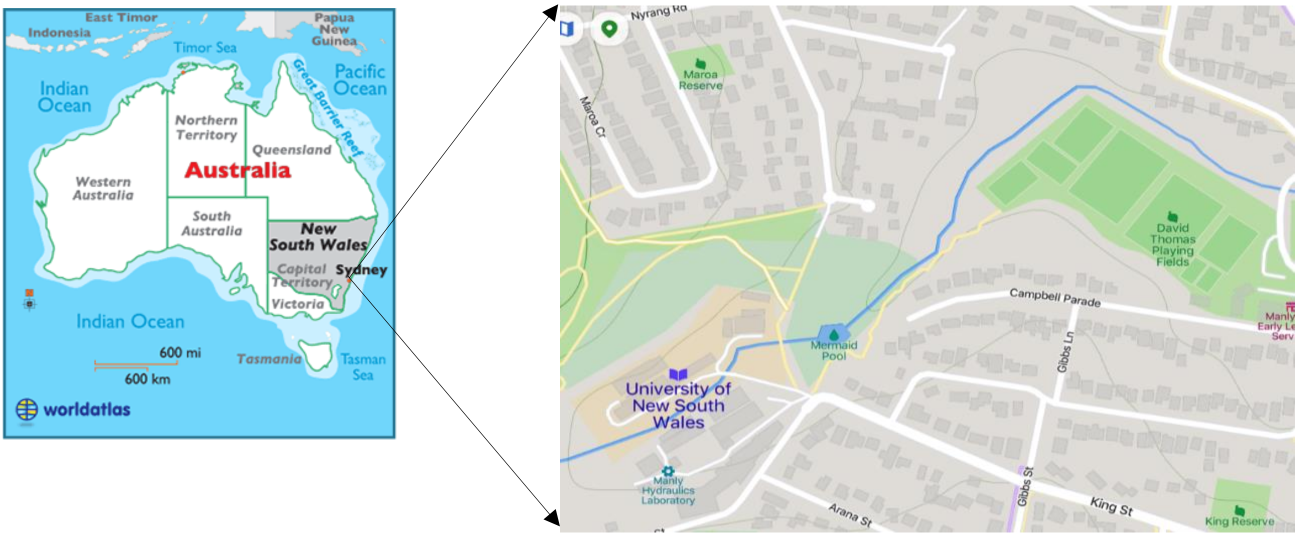


Figure 1. Location of the study area

2.2. Data Acquisition

In June 2005, AAMHatch captured multispectral imagery at a 1:6000 scale using a film camera. As shown in Figure 2, the film was scanned in TIFF format with three color bands (red, green, and blue), a pixel size of 15μm (GSD of 0.096m), and a 16-bit radiometric resolution.

***Test area**: UNSW*
***Size:** 0.5 × 0.5 km*
***Band:** RGB*
***Pixel size:** 9 cm*
***Long track:** ±30*
***Cross track:** ±30*

Figure 2. Aerial photo for study area

3. Research Methodology

Both supervised and unsupervised classification techniques were used in this study. The ISODATA and K-means classifiers were used for unsupervised classification as they don't require training data. The Maximum Likelihood and Minimum Distance classifiers were used in the supervised classification process. Additionally, neural networks and support vector machines were used to carry out per-pixel categorization. The accuracy of supervised categorization is highly dependent on the number and quality of training sites. Typically, supervised classifications follow a sequence of steps: Step 1 involves identifying the training sites, Step 2 focuses on extracting signatures, and Step 3 entails classifying the image. In unsupervised classification, pixels are grouped based on their reflectance characteristics, with each group known as a “cluster”. The analyst organizes the various clusters and selects the bands to utilize. They identify clusters that correspond to different land cover classes. Afterward, the analyst assigns meaningful labels to these clusters and supplies precise satellite images. The clusters are combined by the analyst into a category of land cover. Whether the classification was done under supervision or not, the outcomes were assessed. To maximize the advantages of both supervised and unsupervised methods while minimizing their use, Winner Takes All technique is applied, and the outcomes are then compared with those of the other techniques as shown in Figure 3.

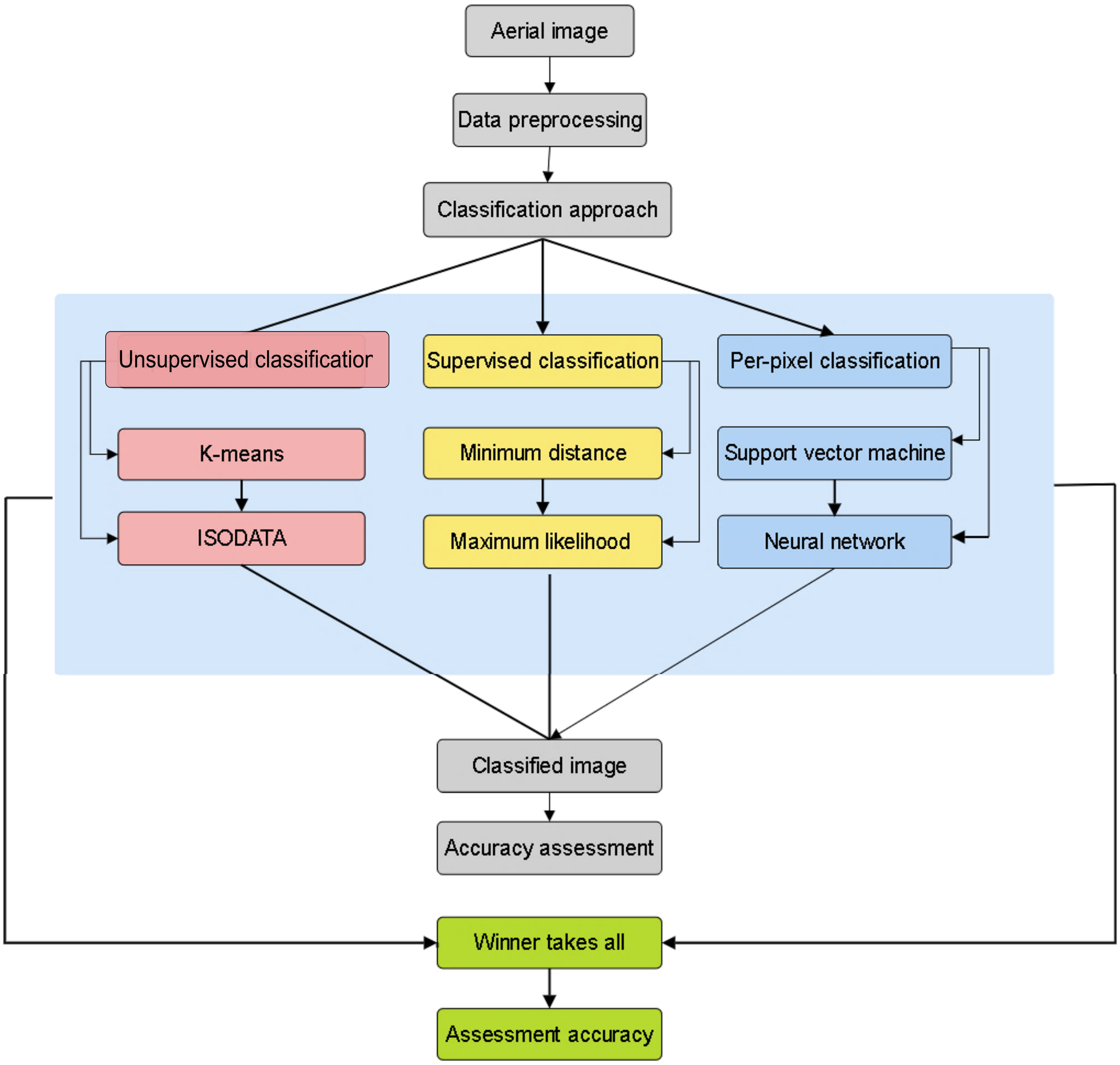


Figure 3. Flowchart for study methodology

3.1. Image Classification

In this study, classification was performed using various classifiers, including K-means, ISODATA, Minimum Distance (MD), Support Vector Machine (SVM), Neural Network (NNT), and Maximum Likelihood (MXL).

Iterative Self-Organizing Data Analysis (ISODATA) is one of the most widely used methods for unsupervised classification [20, 26]. It iteratively refines clusters based on the data’s characteristics without needing prior knowledge of the number of clusters. The mathematical basis of this classifier revolves around iterative clustering, where the algorithm refines clusters through splitting and merging. The ISODATA (Iterative Self-Organizing Data Analysis Technique) algorithm generates a set number of unlabeled clusters or classes in an image. These clusters are later assigned meaningful labels. ISODATA uses various parameters to control the number of clusters, and the number of iterations performed. Sometimes, clusters may contain pixels from different classes. In such cases, ISODATA employs a technique called cluster-busting to identify and separate complex classes [1]. The ISODATA algorithm uses the following mathematical concept; it calculates the Euclidean distance between a data point (x) and a cluster center (c) as the following equation: $d(c,x) = \sqrt{\sum_{i=1}^{n}(x_i - c_i)^2}$ where n is the number of dimensions. The a new cluster center (c) for a cluster with (k) points is computed from the following equation $c = \frac{1}{k}\sum_{i=1}^{k} x_i$ then a cluster is split in case where its standard deviation exceed the threshold this is follow by merging clusters if the distance between them is less than a threshold. The ISODATA algorithm continues to iterate until one of the following conditions is met: the average distance between cluster centers falls below a certain threshold, the change in this distance is less than a specified threshold, or the maximum number of iterations is reached.

The K-means algorithm is a widely utilized clustering method. Although it is an unsupervised technique for clustering in pattern recognition and machine learning, the initial selection of the number of clusters significantly influences the performance of the K-means algorithm and its variations. In other words, the k-means algorithm is not precisely an unsupervised technique for clustering [27]. The K-means algorithm divides (n) observations into (k) clusters using the Euclidean mean. Its benefits include ease of processing and quick execution. However, a drawback is that the analyst must know the number of clusters in advance [1]. In K-means clustering, each data point is assigned to the cluster whose mean is closest, with this mean acting as the cluster's centroid. Initially, centroids are chosen randomly, and each data point is allocated to the nearest centroid. The centroids are then recalculated as the average of all data points in each cluster using the given equation $\mu_i = \frac{1}{|c_j|}\sum_{x_i \in c_j} x_i$ where $(c-j)$ is the set of points in cluster (j). This assignment process continues until either the maximum number of iterations is reached or the centroids remain unchanged.

The Minimum Distance algorithm, also known as spectral distance, classifies each pixel in an image by calculating the Euclidean distance between the pixel's digital value (Dv) and the mean value (Mt) of each training data class. The pixel is assigned to the class with the smallest distance. This method is advantageous because it is fast and ensures that all pixels are classified, depending on the training set used [28]. This technique evaluates training data using two bands. The Minimum Distance classification method computes the Euclidean distance from each pixel in the image to each class:

$$D_i = \sqrt{(x-m_i)^J (x-m_i)} \tag{1}$$

where: D is Euclidean distance, i is the i^{th} class, x is n-dimensional data (where n is the number of bands), m_i is mean vector of a class [16].

Maximum likelihood classification is a popular method for classifying remote sensing images. It relies on two key principles: the assumption that class samples in a multidimensional space are normally distributed, and the application of Bayes' theorem for decision making. When assigning each cell to a class from the signature file, the algorithm considers both the variance and covariance of the class signatures [2]. The mean vector and covariance matrix can accurately characterize the class if the sample has a normal distribution. To ascertain a cell's membership in a class, the statistical probability of each class is computed using these two factors for each cell value [29]. For every pixel in the image, the discriminant functions are calculated using maximum likelihood classification:

$$gi(x) = 1np(\omega i) - {}^{1}\!/_{2}\, 1n|\textstyle\sum i| \dashv\; - {}^{1}\!/_{2}\,(x-mi)^J \sum_i^{-1}(X-m_i) \tag{2}$$

where i is the i^{th} class, x is n-dimensional data (where n is the number of bands), $p(\omega i)$ is probability that a class occurs in the image and is assumed the same for all classes, $|\sum i| \dashv$ is determinant of the covariance matrix of the data in a class, $\sum_i^{-1}$ is the inverse of the covariance matrix of a class, mi is mean vector of a class [16].

Support Vector Machine (SVM) is a robust machine learning technique applicable to binary classification, regression, and anomaly detection. It works by projecting data samples into high-dimensional spaces and constructing hyperplanes to separate them. In infinite dimensional space, a set of hyper-planes is created that divides the data into partitions so that data from the same class is placed in the same partition [30]. The support vector machine will generate the hyperplane iteratively in order to minimize error. The goal is to classify the datasets in order to find the maximum marginal hyperplane. It creates one or more hyperplanes in a high-dimensional space, with the optimal hyperplane being the one that maximizes the distance to the closest training data point from any class, ensuring effective separation between the two classes. The effectiveness of this algorithm largely depends on the kernel function employed, with the linear, Gaussian, and polynomial kernels being the most frequently utilized [31]. For binary classification problem, the hyperplane defined as $w.x + b = 0$; where w is the weight vector, x is the input vector and b is the bias term. The margin which defined as the distance between hyperplane and nearest data points can be expressed as $margin = \frac{2}{//w//}$ hence the objective of SVM is to maximize this margin, while ensuring that all data point are correctly classified [32, 33].

Artificial Neural Networks (ANNs) are sophisticated machine learning algorithms inspired by the architecture and functionality of biological neural systems. These networks are adept at discerning complex relationships between input variables and output categories, effectively handling non-linear data interactions. This capability renders them particularly valuable in the domain of remote sensing classification. ANNs are composed of multiple layers of interconnected neurons, each receiving input signals, performing nonlinear transformations, and transmitting the processed signals to subsequent layers. The output layer ultimately produces the classification labels [33]. The success of Artificial Neural Networks (ANNs) in classification can be attributed to several factors: they do not require any prior assumptions about data distribution, and they enable users to incorporate initial knowledge about classes and their potential boundaries [34].

By combining the outputs of several classifiers, ensemble-based approaches, or multiple classifiers, have been used to improve classification accuracy [35]. Some strategies for ensemble learning approaches are now known, including majority voting, fuzzy integrals, and Dempster-Shafer evidence theory. The most common method is majority voting, which collects the label outputs of each classifier for a given pixel and then assigns the pixel to the majority label. In this research, a classifier ensemble is implemented using the Winner-Takes-All (WTA) approach, coupled with a majority voting mechanism. For all classification methods implemented, WTA classification assigns each pixel to the corresponding class that has the majority [36].

The Winner-Takes-All (WTA) method differs from other aggregation techniques for classifier ensembles in several key ways. In WTA, the classifier with the highest confidence or score for a given input is selected as the final decision. This contrasts with methods like majority voting, where the class label that receives the most votes from all classifiers is chosen. WTA is relatively straightforward, as it relies on the output of a single classifier rather than aggregating results from multiple classifiers. This can make it faster and easier to implement. WTA can be more sensitive to outliers or noisy data, as the final decision is based on the confidence of one classifier. In contrast, techniques like majority voting or averaging can mitigate the impact of outliers by considering the consensus among multiple classifiers. For WTA to be effective, the ensemble needs to include classifiers that are diverse and complementary. Other methods, such as weighted averaging, can still perform well even if the classifiers are not highly diverse. WTA is often used in scenarios where one classifier is expected to be significantly more accurate than others for certain types of data. Majority voting and other aggregation methods are more commonly used when the goal is to leverage the collective strength of multiple classifiers.

3.2. Classification Accuracy Assessment

Accuracy evaluation, the last phase in the image classification process, is crucial for assessing classification strategies and detecting possible inaccuracies in the image. A square array of numbers called an error matrix also known as confusion matrix is used to evaluate the performance of a classification algorithm. It compares the actual class labels with the predicted class labels, providing a detailed breakdown of correct and incorrect predictions. The total of the main diagonal elements represents the classification's overall accuracy. While omission mistakes arise when a test region is mistakenly placed into a different category, omission errors happen when a test area is not classified into its right informative category. The accuracy of the user and the producer shed light on these mistakes. To determine the classification of a test area, a test area of the same size as the training areas was utilized [37]. The alignment between the classification and the ground truth was evaluated using overall accuracy and Kappa Statistics. These metrics can be computed using the formulas provided in Equations 3 to 6 [38].

$$\text{User's accuracy (precision)} = \frac{Number\ of\ Correctly\ Classified\ Pixels\ in\ each\ Category}{Total\ Number\ of\ Reference\ Pixels\ in\ that\ Category\ (The\ Row\ Total)} \times 100 \tag{3}$$

$$\text{Producer accuracy (recall)} = \frac{Number\ of\ Correctly\ Classified\ Pixels\ in\ each\ Category}{Total\ Number\ of\ Reference\ Pixels\ in\ that\ Category\ (The\ Column\ Total)} \times 100 \tag{4}$$

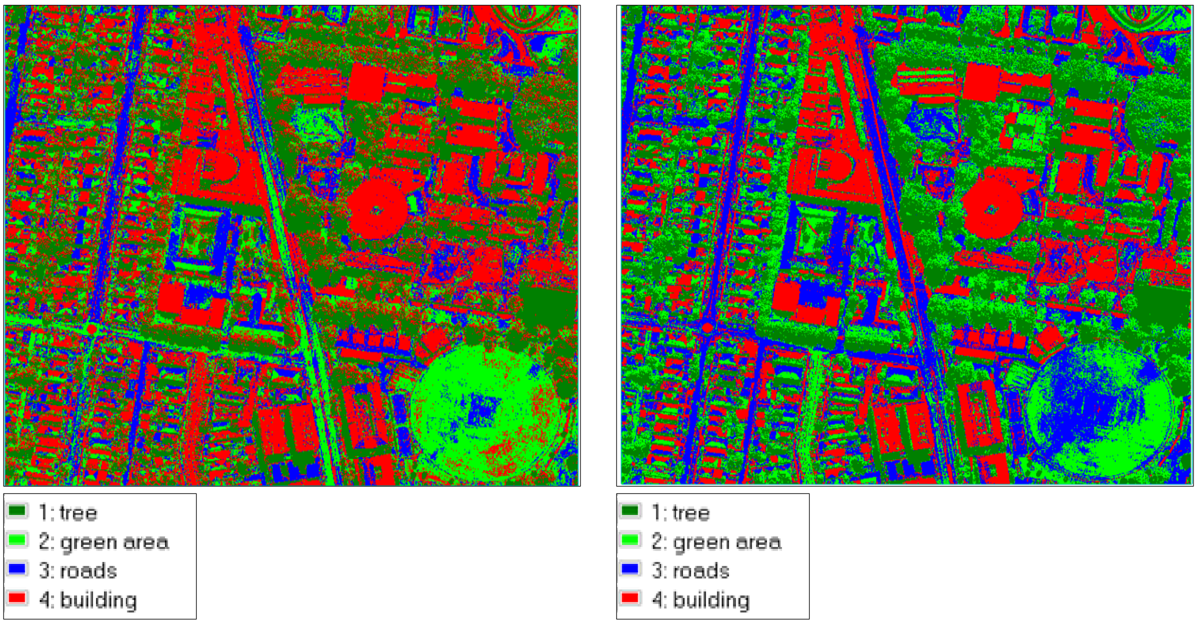
$$\text{overall accuracy} = \frac{Total\ Number\ of\ Correctly\ Classified\ Pixels\ (Diagonal)}{Total\ Number\ of\ Reference\ Pixels} \times 100 \tag{5}$$

$$\text{Kappa Coefficient (T)} = \frac{(TSxTCS) - \sum(Column\ Total \times Row\ Total)}{TS^2 - \sum(Column\ Total - Row\ Total)} \tag{6}$$

There are several factors have great influence on the classification accuracy. Addressing class imbalance is crucial for real-world application. Class imbalance occurs when the number of instances in different classes of a dataset is not evenly distributed. This can cause problem as they might become biased towards the majority class and perform poorly on the minority class. This problem can be handle through data augmentation, class weighting, and segmentation techniques. Quality and quantity of data can also affect the accuracy of classification model. Noisy, incomplete or insufficient data can reduce classification accuracy.

4. Results and Discussion

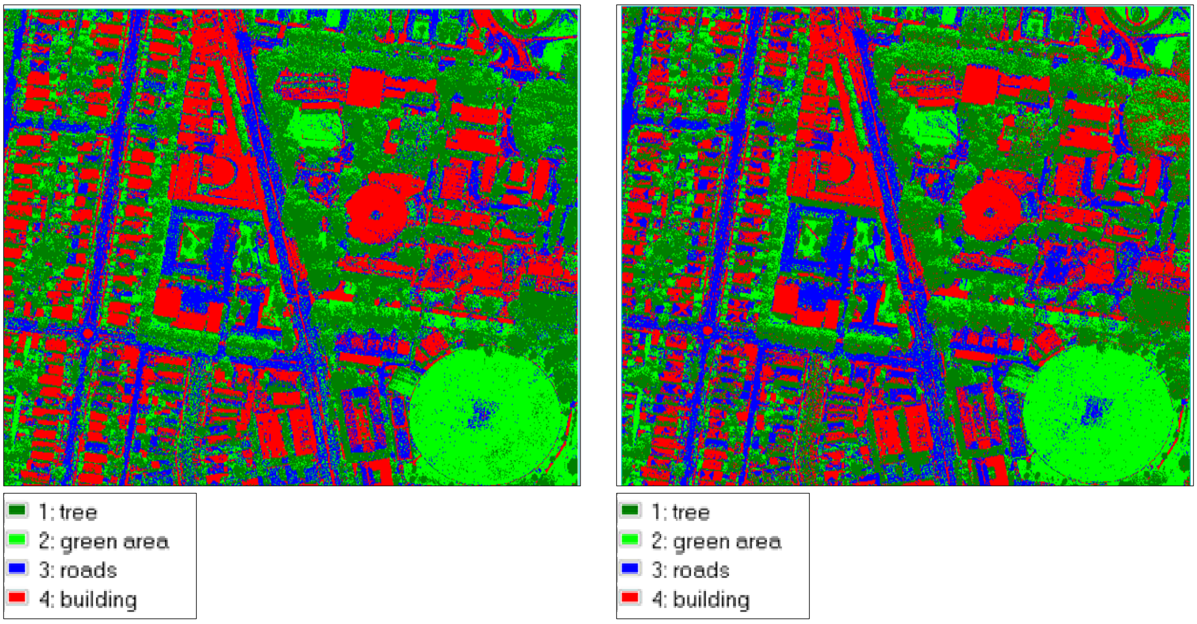
The results of the unsupervised classification using K-means and ISODATA methods showed that ISODATA outperformed K-means, even though ISODATA needs multiple parameters. K-means, on the other hand, is affected by the initial cluster centers chosen by the analyst, the geometric characteristics of the data, and the clustering parameters, as illustrated in Figure 4.



K-means classification ISODATA classification

Figure 4. Unsupervised classification

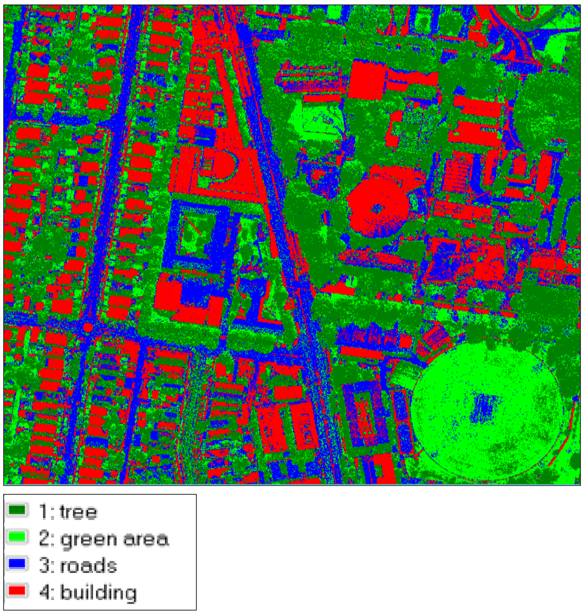
As shown in Figure 5, neural networks performed most effectively in per-pixel classification, and the overall accuracies and kappa coefficients of both methods were close, with differences in overall accuracy and kappa coefficients of 0.65 and 0.01 respectively.



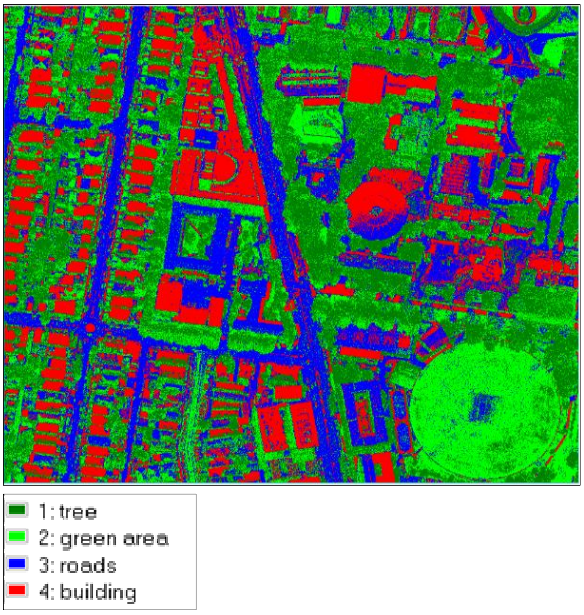
Support vector machine classification Neural network classification

Figure 5. Classification

In supervised classification method, Maximum likelihood classification gave 84.97% overall accuracy while minimum distance gave 77.05% overall accuracy as seen in Figure 6.

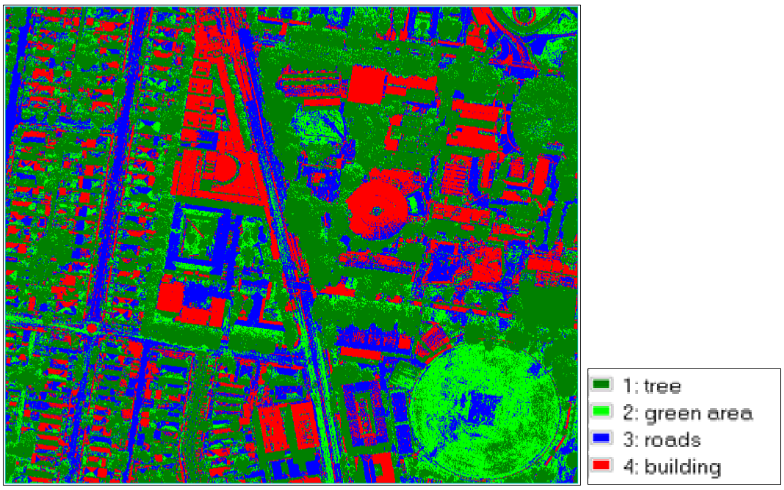


Maximum likelihood classification



Minimum distance classification

Figure 6. Supervised classified image



Winner takes all approach

Figure 7. Classified image using WTA

As seen in Figure 7, winner takes all gives the best result as the overall accuracy reached 88.69%.

Table 1 and Figure 8 demonstrated the classifiers' comparative performance in terms of overall accuracy and kappa coefficients. Tables 2 to 8 show the individual class accuracies for the classification methods assessed by producer and user accuracy.

Table 1. Accuracy assessment of the four classification techniques

Method	Overall accuracy (%)	Kappa coefficient
K-means classification	73.79%	0.58
ISODATA classification	79.50%	0.64
Maximum likelihood classification	84.97%	0.76
Minimum distance classification	77.05%	0.64
Support vector machine classification	86.53%	0.78
Neural network classification	87.18%	0.79
Winner takes all	88.69%	0.80

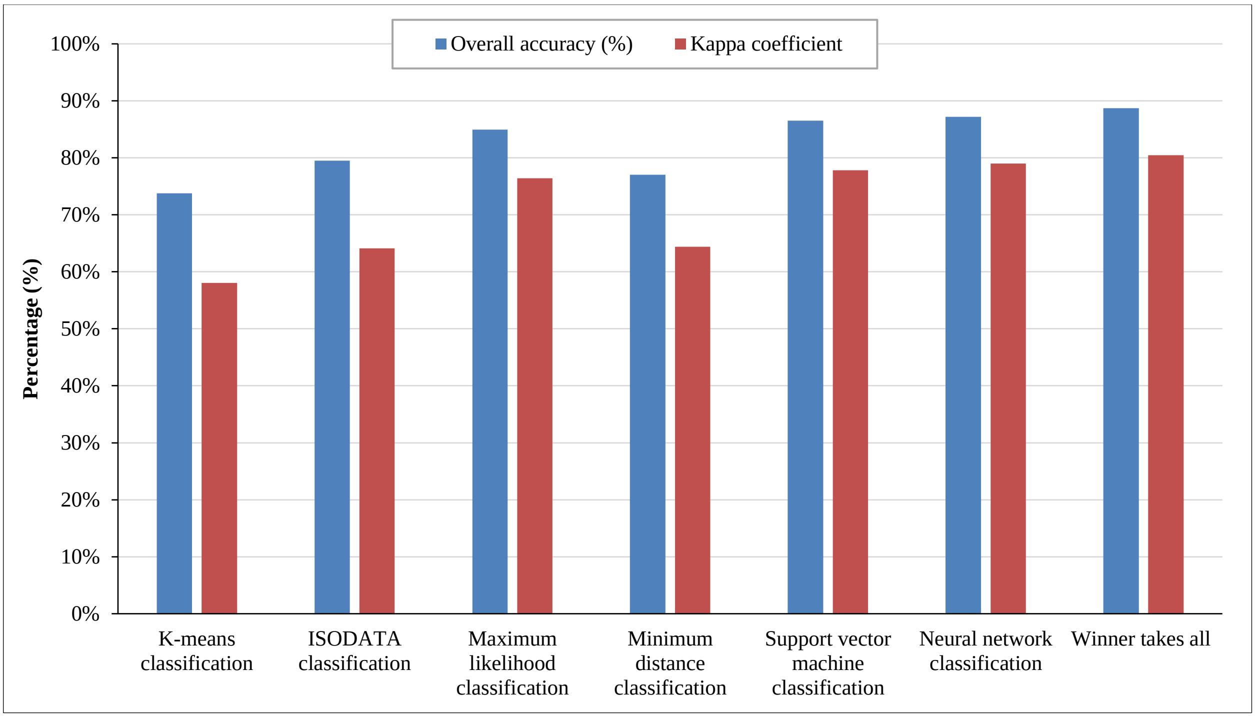


Figure 8. Overall accuracy and kappa coefficients for the used classification method

The three most accurate classification methods were SVM, NNT, and WTA, with overall accuracies of 86.53%, 87.18%, and 88.69%, respectively. Given the close accuracy rates of these methods, WTA emerged as the best performer with 88.69%, while K-means had the lowest performance at 73.79%. The Kappa values for MXL, SVM, NNT, and WTA were all above 0.75, signifying excellent agreement. In contrast, the Kappa coefficients for K-means, ISODATA, and MD ranged from 0.4 to 0.75, indicating moderate to good agreement.

As the overall accuracy indicates, the ensemble (WTA) method performed the best. This suggests that combining multiple classifiers can lead to better performance compared to individual classifiers. The neural network also performed well, indicating its strong capability in handling complex patterns in data. The kappa coefficient for WTA is the highest among all classifiers. This indicates a very high level of agreement between observed and expected classifications, suggesting that the ensemble method is highly reliable; the NNT also shows a strong kappa value, reinforcing its robustness.

Maximum likelihood classification achieves an overall accuracy of 84.97% and a Kappa coefficient of 0.76, indicating it performs fairly well. However, it is surpassed by more advanced methods such as SVM and NNT. Neural network with 87.18% overall accuracy and a kappa of 0.79, it shows excellent performance, though it requires more data and computational power. Support vector machines achieve 86.53% overall accuracy and a kappa of 0.78, indicating strong performance, especially in high-dimensional spaces. Minimum distance classifier, which is considered the simplest method with 77.05% overall accuracy and a kappa of 0.6. K-means achieves the worst overall accuracy and a kappa coefficient; it is affected by the initial placement of the cluster centers. ISODATA achieves higher accuracy than k-means due to its ability to adjust the number of clusters dynamically. Ensemble technique (WTA) is the best performer with 88.69% overall accuracy and a kappa of 0.80, demonstrating the power of combining multiple classifiers. By comparing the performance of these classifiers using user and producer accuracy matrices for four classes. These matrices provide insights into the reliability and effectiveness of each classifier in correctly identifying and classifying these classes.

Table 2. Accuracy assessment of ISODATA classification

Class	Producer accuracy (%)	User accuracy (%)
Building	92.91	96.58
Green area	75.68	52.86
Road	54.19	56.68
Trees	24.62	69.01

In ISODATA classification. Buildings have high user and producer accuracy, so most pixels classified as buildings are buildings, and most actual buildings are correctly classified. In green areas, producer accuracy is moderate; therefore, many green areas are correctly classified, while in roads and trees, producer accuracy is low and very low, respectively, hence most actual roads and trees are missed.

Table 3. Accuracy assessment of K-means classification

Class	Producer accuracy (%)	User accuracy (%)
Building	92.78	95.57
Green area	58.54	50.70
Road	67.90	37.67
Trees	23.91	73.21

In k-means classifications, most actual buildings are correctly classified, many green areas are missed, many roads are correctly classified, and most actual trees are missed.

Table 4. Accuracy assessment of Maximum likelihood classification

Class	Producer accuracy (%)	User accuracy (%)
Building	91.76	99.68
Green area	80.28	75.53
Road	89.71	57.52
Trees	61.22	73.68

In maximum likelihood classification, buildings have a high user and producer accuracy, which indicates that most pixels classified as buildings are indeed buildings, and most actual buildings are correctly classified. In green areas, user accuracy and producer accuracy are moderate, which means many green areas are correctly classified, but some non-green areas are also classified as green. In roads, user accuracy is low while producer accuracy is high. In trees, user and producer accuracy are moderate, so many trees are correctly classified, but some non-trees are also classified as trees.

Table 5. Accuracy assessment of Minimum distance classification

Class	Producer accuracy (%)	User accuracy (%)
Building	87.82	98.93
Green area	57.93	54.91
Road	79.84	50.00
Trees	59.48	62.77

In minimum distance classification, most actual buildings are correctly classified, while green areas are missed, and many non-green areas are classified as green. Many roads are correctly identified, but many non-road areas are also classified as roads. Many tree areas are correctly classified, but some non-tree areas are also classified as trees.

Table 6. Accuracy assessment of Support vector machine classification

Class	Producer accuracy (%)	User accuracy (%)
Building	99.12	96.66
Green area	86.99	69.82
Road	69.14	79.25
Trees	47.81	75.23

In support vector machines, high user accuracy and very high producer accuracy are resulted in buildings. User accuracy is moderate while producer accuracy is high in green areas. User and producer accuracy are moderate in roads. While in trees, user accuracy is moderate, and producer accuracy is low.

Table 7. Accuracy assessment of neural network classification

Class	Producer accuracy (%)	User accuracy (%)
Building	99.12	97.35
Green area	87.40	71.31
Road	79.42	75.37
Trees	44.61	78.14

In neural networks, user accuracy is high in buildings and moderate in roads, green areas, and trees. Buildings have a very high producer accuracy, only high in green areas, low in trees, and moderate in roads.

Table 8. Accuracy assessment of Winner takes all

Class	Producer accuracy (%)	User accuracy (%)
Building	96.86	98.08
Green area	83.74	72.79
Road	80.66	71.79
Trees	54.52	73.05

In WTA, user accuracy is very high in buildings and moderate in green areas, trees, and roads. Producer accuracy is very high in buildings, high in green areas and roads, whereas it is moderate in trees.

In Figure 9, the misclassification rate represents the percentage of instances that were incorrectly classified by each method. High misclassification rate in k-means indicates that it struggled significantly with the dataset. This could be due to its reliance on distance measures, which might not capture the complexity of the classes well. ISODATA represents a moderate misclassification rate better than k-means but still relatively high. The maximum likelihood classifier has a lower misclassification rate, which indicates better performance because this method follows a Gaussian distribution. Misclassification rate in minimum distance classification indicates that this method struggled with the data. In support vector machines, the misclassification rate indicates good performance. This is due to the fact that SVMs are powerful for high-dimensional spaces and can handle non-linear boundaries well. Also in neural networks, misclassification rate indicates very good performance. It can model complex relationships in the data. Winner takes all represents the lowest misclassification rate that indicates the best performance among all classifiers.

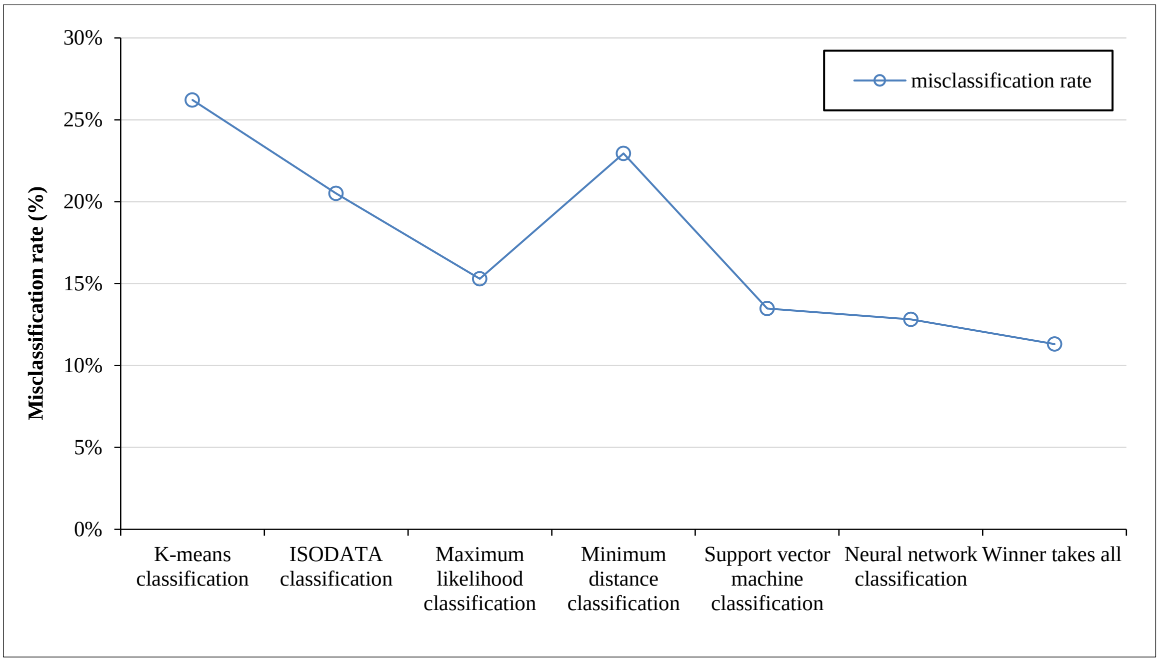


Figure 9. Misclassification rate for the used classification method

Figures 10 to 12, SVM and NNT show the greatest producer accuracy, while MXL has the highest user accuracy in the case of building, while in trees WTA has the greatest accuracy in both producer and user metrics. NNT shows the highest producer and user accuracy in the case of green areas, while in roads, MXL has the highest producer accuracy and WTA has the highest user accuracy.

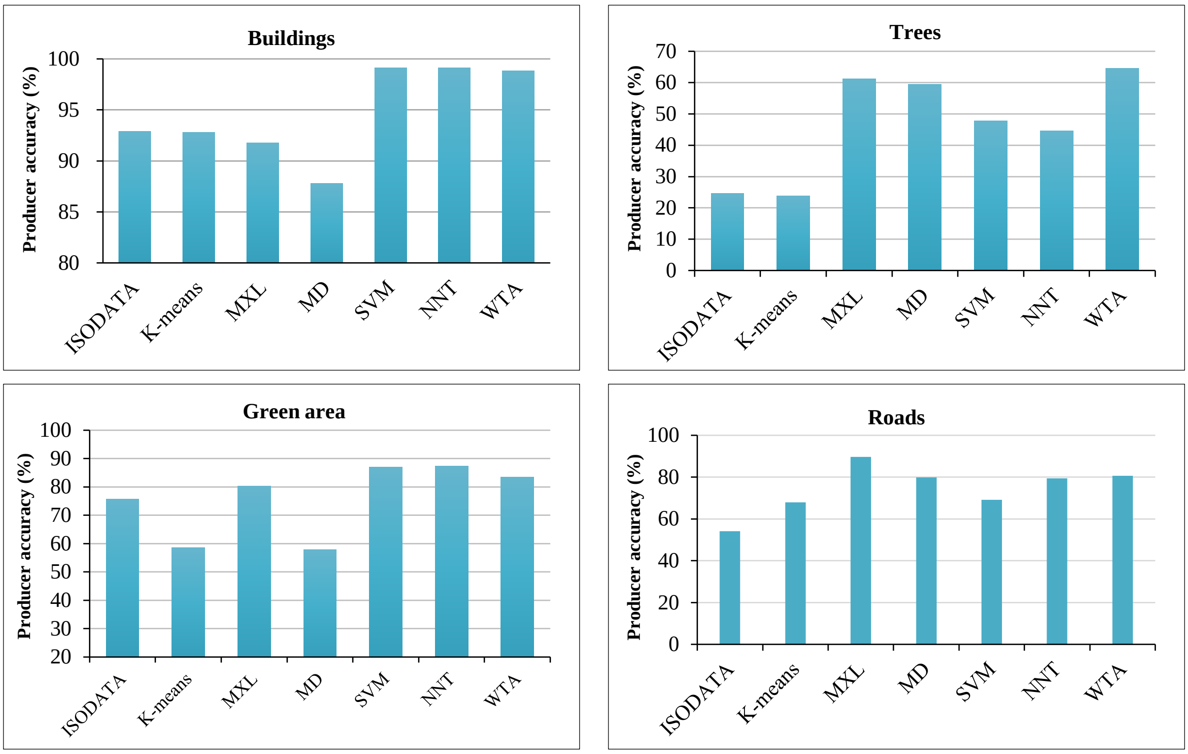


Figure 10. Accuracy of the four classes in each classification method in terms of producer accuracy

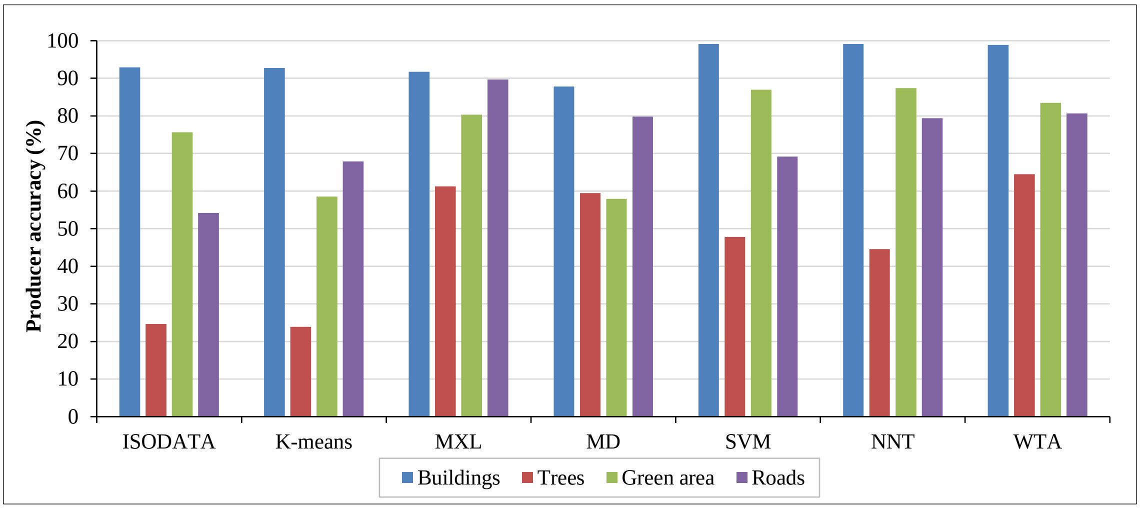


Figure 11. Comparison between different selected classification methods in the four classes in term of producer accuracy

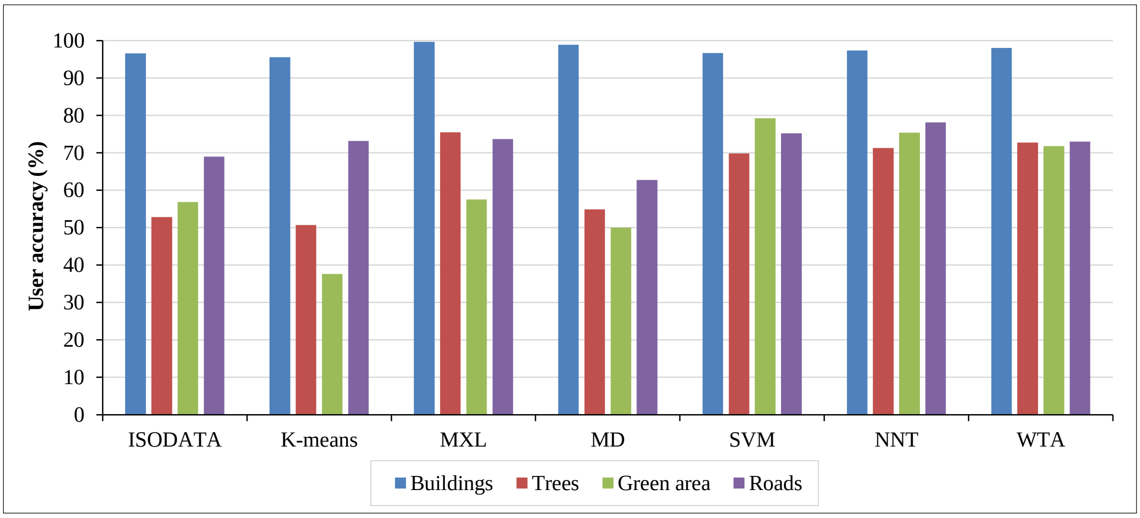


Figure 12. Comparison between different selected classification methods in the four classes in term of user accuracy

5. Conclusion

Six different classification techniques and a classifier ensemble were investigated in this research, and four different categories were classified. It would be reasonable to conclude that each type of classification method has different sensitivity and advantages on classes, and classifier ensembles could improve overall classification accuracy as well as individual class-based accuracy. The ISODATA classifier is superior to the K-means classifier. The results show that WTA classification outperformed the other methods by 88.69%, but it was still close to SVM and NNT, while K-means had the lowest classification accuracy of 73.79%. Furthermore, NNT provided the second highest overall classification accuracy of 87.18%, and no method can accurately classify all classes with high accuracy.

The best overall classifier is WTA; it stands out as the most balanced and effective classifier, showing high accuracy across all categories. NNT is a strong alternative; it also performs exceptionally well, particularly for buildings, green areas, and roads. For building, both SVM and NNT are excellent choices. For trees, WTA is the best option. For green areas, NNT shows the highest accuracy. For roads, both MXL and WTA are reliable. Based on a comprehensive analysis of various classifiers, it is recommended to prioritize the WTA classifier for classification tasks. The WTA classifier demonstrates consistently high accuracy across all categories. This balanced performance makes it a robust choice for diverse classification needs, ensuring reliable results in both producer and user accuracy matrices. The neural network also shows strong performance, making it a reliable choice. However, the selection of a classifier may be influenced by specific requirements, such as the need for interpretability, the characteristics of the data, and the computational resources available. In our future work, we aim to evaluate the classification performance of additional advanced techniques, including convolutional neural networks (CNN), recurrent neural networks (RNN), and object-based SVM.

6. Declarations

6.1. Author Contributions

Conceptualization, K.M.A. and L.E.; methodology, L.E.; software, K.M.A. and L.E.; validation, K.M.A. and L.E.; formal analysis, K.M.A. and L.E.; investigation, K.M.A. and L.E.; resources, K.M.A.; data curation, K.M.A. and L.E.; writing—original draft preparation, K.M.A. and L.E.; writing—review and editing, K.M.A. and L.E.; visualization, K.M.A. and L.E.; supervision, K.M.A. and L.E.; project administration, K.M.A. and L.E.; funding acquisition, K.M.A. and L.E. All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

Data sharing is not applicable to this article.

6.3. Funding

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2023/01/25799).

6.4. Conflicts of Interest

The authors declare no conflict of interest.

7. References

[1] Singh, M., & Tyagi, K. D. (2021). Pixel based classification for Landsat 8 OLI multispectral satellite images using deep learning neural network. Remote Sensing Applications: Society and Environment, 24, 100645. doi:10.1016/j.rsase.2021.100645.

[2] Soliman, O. S., & Mahmoud, A. S. (2012). A classification system for remote sensing satellite images using support vector machine with non-linear kernel functions. 2012 8th International Conference on Informatics and Systems (INFOS), 14-16 May, 2012, Giza, Egypt.

[3] Abburu, S., & Babu Golla, S. (2015). Satellite Image Classification Methods and Techniques: A Review. International Journal of Computer Applications, 119(8), 20–25. doi:10.5120/21088-3779.

[4] Briney, A. (2019). An Overview of Remote Sensing. ThoughtCo, New York, United States. Available online: https://www.thoughtco.com/an-overview-of-remote-sensing-1434624 (accessed on September 2024).

[5] Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing, 28(5), 823–870. doi:10.1080/01431160600746456.

[6] Foody, G. M., Pal, M., Rocchini, D., Garzon-Lopez, C. X., & Bastin, L. (2016). The sensitivity of mapping methods to reference data quality: Training supervised image classifications with imperfect reference data. ISPRS International Journal of Geo-Information, 5(11), 199. doi:10.3390/ijgi5110199.

[7] Shivakumar, B. R., & Rajashekararadhya, S. V. (2018). Investigation on land cover mapping capability of maximum likelihood classifier: A case study on North Canara, India. Procedia Computer Science, 143(579), 579–586. doi:10.1016/j.procs.2018.10.434.

[8] Huang, C., Davis, L. S., & Townshend, J. R. G. (2002). An assessment of support vector machines for land cover classification. International Journal of Remote Sensing, 23(4), 725–749. doi:10.1080/01431160110040323.

[9] Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. International Journal of Remote Sensing, 39(9), 2784–2817. doi:10.1080/01431161.2018.1433343.

[10] Pal, M., & Mather, P. M. (2005). Support vector machines for classification in remote sensing. International Journal of Remote Sensing, 26(5), 1007–1011. doi:10.1080/01431160512331314083.

[11] Mathur, A., & Foody, G. M. (2004). Land cover classification by support vector machine: Towards efficient training. International Geoscience and Remote Sensing Symposium (IGARSS), 2, 742–744. doi:10.1109/igarss.2004.1368508.

[12] Salah, M. (2017). A survey of modern classification techniques in remote sensing for improved image classification. Journal of Geomatics, 11(1), 21.

[13] D., S., Deepa, P., & K., V. (2017). Remote Sensing Satellite Image Processing Techniques for Image Classification: A Comprehensive Survey. International Journal of Computer Applications, 161(11), 24–37. doi:10.5120/ijca2017913306.

[14] Sathya, P., & Baby-Deepa, V. (2017). Analysis of Supervised Image Classification Method for Satellite Images. International Journal of Computer Science Research, 5(2), 16–19.

[15] Manohar, N., Pranav, M. A., Aksha, S., & Mytravarun, T. K. (2021). Classification of Satellite Images. Information and Communication Technology for Intelligent Systems. ICTIS 2020, Smart Innovation, Systems and Technologies, 195, Springer, Singapore. doi:10.1007/978-981-15-7078-0_70.

[16] Murtaza, K. O., & Romshoo, S. A. (2014). Determining the suitability and accuracy of various statistical algorithms for satellite data classification. International Journal of Geomatics and Geosciences, 4(4), 585-599.

[17] Altaei, M. S. M., & Mhaimeed, A. D. (2017). Satellite image classification using image encoding and artificial neural network. International Research Journal of Advanced Engineering and Science, 3(2), 149-154.

[18] Neware, R., & Khan, A. (2018). Survey on Classification Techniques Used in Remote Sensing for Satellite Images. 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), 1860–1863. doi:10.1109/iceca.2018.8474881.

[19] Pandya, A., & Science, C. (2015). Classification of Vegetation Area from Satellite. International Journal of Research in IT, Management and Engineering, 5(3), 48–56.

[20] Vimala, R, Marimuthu, A., Venkateswaran, S., & Poongodi, R. (2020). Unsupervised ISODATA algorithm classification used in the landsat image for predicting the expansion of Salem urban, Tamil Nadu. Indian Journal of Science and Technology, 13(16), 1619–1629. doi:10.17485/ijst/v13i16.271.

[21] Ouchra, H., Belangour, A., & Erraissi, A. (2023). Machine Learning Algorithms for Satellite Image Classification Using Google Earth Engine and Landsat Satellite Data: Morocco Case Study. IEEE Access, 11, 71127–71142. doi:10.1109/access.2023.3293828.

[22] Yasin, H. E., & Kornel, C. (2024). Evaluating Satellite Image Classification: Exploring Methods and Techniques. Geographic Information Systems - Data Science Approach. IntechOpen, London, United Kingdom. doi:10.5772/intechopen.1003196.

[23] Ouchra, H., Belangour, A., & Erraissi, A. (2024). Supervised Machine Learning Algorithms for Land Cover Classification in Casablanca, Morocco. Ingenierie Des Systemes d'Information, 29(1), 377–387. doi:10.18280/isi.290137.

[24] Nigar, A., Li, Y., Jat Baloch, M. Y., Alrefaei, A. F., & Almutairi, M. H. (2024). Comparison of machine and deep learning algorithms using Google Earth Engine and Python for land classifications. Frontiers in Environmental Science, 12, 10 3389 2024 1378443. doi:10.3389/fenvs.2024.1378443.

[25] Ahmadi, K. (2024). Assessment of the Accuracy of Various Machine Learning Algorithms for Classifying Urban Areas through Google Earth Engine: A Case Study of Kabul City, Afghanistan. European Journal of Artificial Intelligence and Machine Learning, 3(3), 1–8. doi:10.24018/ejai.2024.3.3.40.

[26] Vasquez, J., Acevedo-Barrios, R., Miranda-Castro, W., Guerrero, M., & Meneses-Ospina, L. (2024). Determining Changes in Mangrove Cover Using Remote Sensing with Landsat Images: a Review. Water, Air, & Soil Pollution, 235(1), 18. doi:10.1007/s11270-023-06788-6.

[27] Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-means clustering algorithm. IEEE Access, 8, 80716–80727. doi:10.1109/ACCESS.2020.2988796.

[28] Abinaya, V., & Poonkuntran, S. (2019). Classification of satellite image using Minimum Distance classification algorithm. International Journal of Computer Science and Engineering, 6(3), 15-18.

[29] Ahmad, A., & Quegan, S. (2012). Analysis of maximum likelihood classification on multispectral data. Applied Mathematical Sciences, 6(129), 6425-6436.

[30] Ibrahim, S. (2019). Comparative Analysis of Support Vector Machine (SVM) and Convolutional Neural Network (CNN) for White Blood Cells' Classification. International Journal of Advanced Trends in Computer Science and Engineering, 8(1.3), 394–399. doi:10.30534/ijatcse/2019/6981.32019.

[31] Sanghvi, K. (2020). Image Classification Techniques. Medium, San Francisco, United States. Available online: https://medium.com/analytics-vidhya/image-classification-techniques-83fd87011cac (accessed on September 2024).

[32] Cortes, C. (1995). Support-Vector Networks. Machine Learning, 20, 273-297.

[33] Kranjčić, N., Cetl, V., Matijević, H., & Markovinović, D. (2023). Comparing Different Machine Learning Options to Map Bark Beetle Infestations in Croatia. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-4/W7-2023, 83–88. doi:10.5194/isprs-archives-xlviii-4-w7-2023-83-2023.

[34] Hasan, M., Ullah, S., Khan, M. J., & Khurshid, K. (2019). Comparative Analysis of SVM, ANN and CNN for Classifying Vegetation Species Using Hyperspectral Thermal Infrared Data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W13, 1861–1868. doi:10.5194/isprs-archives-xlii-2-w13-1861-2019.

[35] Du, P., Xia, J., Zhang, W., Tan, K., Liu, Y., & Liu, S. (2012). Multiple classifier system for remote sensing image classification: A review. Sensors, 12(4), 4764–4792. doi:10.3390/s120404764.

[36] Ustuner, M., Esetlili, M. T., Sanli, F. B., Abdikan, S., & Kurucu, Y. (2016). Comparison of crop classification methods for the sustainable agriculture management. Journal of Environmental Protection and Ecology, 17(2), 648–655.

[37] Patil, M. B., Desai, C. G., & Umrikar, B. N. (2012). Image Classification Tool for Land Use / Land Cover Analysis: A Comparative Study of Maximum Likelihood. International Journal of Geology, Earth, and Environmental Sciences, 2(3), 189–196.

[38] Islami, F. A., Tarigan, S. D., Wahjunie, E. D., & Dasanto, B. D. (2022). Accuracy Assessment of Land Use Change Analysis Using Google Earth in Sadar Watershed Mojokerto Regency. IOP Conference Series: Earth and Environmental Science, 950(1), 012091. doi:10.1088/1755-1315/950/1/012091.