ARQII Publication
http://arqiipubl.com/ams

APPLICATIONS OF
MODELLING AND SIMULATION

eISSN 2600-8084 VOL 9, 2025, 122-138

An Iterative Pixel-Based Dimensional Voting Model for High Spatial-Resolution Image Classification

Muhamad Asyraf Mohd Pouzi[1], Haza Nuzly Abdull Hamed[1*], Muhammad Razib Othman[1], Hishammuddin Asmuni[1], Mohd Adham Isa[1] and Husni Ruslai[2]

[1]Faculty of Computing, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[2]GATES IT Solution Sdn. Bhd., Malaysia

[*]Corresponding author: haza@utm.my

Submitted 18 December 2024; Revised 11 March 2025; Accepted 12 March 2025; Available online 03 April 2025.

Abstract: Understanding land use and land cover changes is crucial for effective environmental management, particularly in mixed-land zones and urban areas, which exhibit distinct characteristics influenced by diverse factors. Rapid urbanization, land utilization patterns, green space preservation, clouds, shadows, pollution, and dynamic human activities pose significant challenges in accurate classification. Traditional classification methods often struggle due to ineffective stand-alone data classification, high costs associated with data fusion, and the need for frequent data collection. To address these issues, this paper proposes an Iterative Pixel-Based Dimensional Voting model (Pixel-DMV), which enhances classification accuracy by iteratively analyzing pixel similarities within 3 x 3 neighborhood pixels. The model assigns a label to unknown pixels based on computed similarity measures and ranks to predict the class for the unknown pixel. The performance is then measured using statistical indices such as Overall Accuracy (OA) and the Kappa Index of Agreement (Kappa). Pixel-DMV outperformed conventional methods, including Support Vector Machine (SVM), Random Forest (RF), and Maximum Likelihood (ML) classifications. The findings indicate that Pixel-DMV simplifies the classification process by relying on a single dataset, reducing data preparation costs, and is suitable for frequent data collection tasks. Given its high accuracy, the proposed model is well-suited for applications in agricultural management, urban planning, and disaster response.

Keywords: Machine learning algorithms; Pleiades satellite; Remote sensing; Satellite images; Supervised classification; SPOT 7 satellite.

1. INTRODUCTION

The field of remote sensing, particularly in the domain of land use and land cover (LULC), involves the utilization of advanced technologies to collect and analyze data related to the Earth's surface. Among the applications of remote sensing classification are in urban planning (assessing urban expansion, monitoring infrastructure development, and managing urban growth), agriculture (monitoring crop health and estimating yields), and environmental monitoring (change monitoring in natural ecosystems, tracking deforestation, and assessing the health of water bodies). The classification process is carried out within two distinct types of areas of interest (AOI): 1) mixed-land zones, as discussed by [1], and 2) urban areas, which have been examined by [2-4] Notably, distinguishing between mixed-land zones and urban areas relies on several key differences: composition, building types, vegetation, density, and water features.

A mixed-land zone is a diverse range of land uses coexisting within the same space, including industrial areas, port facilities, roads, and natural landscapes [5]. Meanwhile, urban areas are predominantly characterized by densely built-up structures and infrastructure, focusing on residential, commercial, and industrial developments. In mixed-land zones, building structures are typically low to medium-rise, reflecting various functions such as factories, warehouses, and open spaces. Meanwhile, the structures in urban areas are typically high-rise buildings, reflecting the concentrated and vertical nature of urban development. For vegetation presence, mixed-land zones can include significant green spaces or vegetation, such as parks, forests, or agricultural areas. In contrast, urban areas have smaller and more scattered green spaces, such as parks or tree-lined streets, compared to the larger vegetation areas in mixed-land zones. On the other hand, building density in mixed-land zones is often more moderate, with a mix of developed and undeveloped spaces contributing to the overall composition compared to high building density. This is a defining feature, with structures closely spaced and occupying a significant portion of the landscape in urban areas. In regard to water features, mixed-land zones may include rivers, lakes, or coastal areas,

contributing to the diversity of the landscape. While urban areas may have rivers or lakes, they generally have fewer water features compared to mixed-land zones.

Despite the heterogeneity of these features, the classification task is further complicated by the inclusion of seawater in the mixed-land zone and coastal area. The dynamic nature of water characteristics, influenced by human activities, the daily operation of boats by fishermen of various sizes, and water pollution introduce continuous changes over time. In addition, these factors collectively contribute to a challenging classification environment, complicating the use of different data types. Due to rapid changes in vast segments in mixed-land zones, frequent classification works are required to be served in diverse applications.

The motivation for this study arises from the limitations of traditional classification methods, which often struggle with ineffective stand-alone data classification, high costs associated with data fusion, and the need for frequent data collection. Despite the remarkable enhancement in the spatial resolution of various sensors over the preceding years and the wide range of capabilities offered by diverse remote sensing data types, the classification task remains challenging. For instance, Li *et al*. [6] proposed a low-to-high network (L2HNet) that automatically generates high-resolution land-cover maps from high-resolution images using only low-resolution land-cover products as training labels, eliminating the need for finely labelled samples during large-scale map updates. While such advancements present opportunities for accurate classification of land use and land cover, the prominent application of learning algorithms is still necessary to address the complexities of the task.

Recently, conventional classification methodologies like Support Vector Machine (SVM), Random Forest (RF), and diverse ensemble methods have given way to more advanced Neural Network (NN) architectures. This transition enables the incorporation of various facets of multi-modal data, as highlighted by [7]. Adriano *et al.* [8] proposed rapid building damage mapping during the 2018 Sulawesi earthquake and tsunami in Palu, Indonesia, using Sentinel-2 (S-2), PlanetScope, Sentinel-1 (S-1), and ALOS-2 PALSAR-2 data. Several ensembles of Decision Tree (DT) were evaluated, such as RF, Rotation Forests (ROF), and Canonical Correlation Forests (CCF). These classifiers performed differently in every scenario. CCF outperformed the RF and ROF in scenarios 1, 2, 3, 4, 5, 6, and 7, while for scenario 8, the best result was held by ROF. Nevertheless, the Overall Accuracy (OA) differences between classifiers for every scenario are relatively low, ranging from 0.20-1.10%. They are 0.20-0.52% (scenario 1), 0.29-0.50% (scenario 2), 0.70-0.81% (scenario 3), 0.39-0.48% (scenario 4), 0.51-1.01% (scenario 5), 0.41-0.85% (scenario 6), 0.72-1.10% (scenario 7) and 0.27-0.56% (scenario 8). Ghorbanzadeh *et al.* [9] achieved high accuracy with low complexity using RF compared to Artificial Neural Network (ANN).

Nguyen *et al.* [10] discovered that unsupervised learning can achieve lower complexity compared to RF. Methods like *k*-Means are particularly effective when within-group variances are homogeneous, as distance-based group assignments are more likely to be accurate. However, when variances are heterogeneous, this approach becomes prone to misclassification [11]. To address this issue, they proposed an improved version of *k*-Means classification by introducing probabilistic *k*-Means, utilizing datasets from Pavia University, Okavango, Southern France, and eastern Thailand. This approach has demonstrated enhanced precision, the ability to determine the optimal number of clusters, and effective management and analysis of large datasets.

Lv *et al.* [12] also proposed an enhancement of the *k*-Means algorithm by developing the *k*-Means_AMV (adaptive majority voting) approach for several bitemporal remote sensing images: 1) a landslide event in Lantau Island, Hong Kong, 2) land use change in Ji Nan City, Shan Dong Province, China, and 3) Sardinia Island, Italy. This approach starts with the generation of an adaptive region around the central pixel and then labeling each pixel using the *k*-Means algorithm. The *k*-Means AMV approach demonstrated better detection accuracy than several other methods. Despite that, the *k*-Means algorithm presents specific limitations, such as issues related to the random initialization of centroids, resulting in unforeseen convergence challenges. Moreover, this clustering algorithm necessitates the pre-definition of the number of clusters, contributing to diverse cluster shapes and susceptibility to outlier effects. A fundamental drawback of the *k*-Means algorithm lies in its inability to effectively manage various data types [13].

On the other hand, a comparative analysis of SVM and Machine Learning (ML) classification was conducted by [14] in Selangor, Malaysia, using Landsat-8 TM data. The study suggested that SVM performed slightly better than ML with 0.04 Kappa difference. Swetanisha *et al.* [15] studied change detection using RF, Minimum Distance (MD), and SVM using Landsat satellite-8 ETM+ data. Located in the Puri District of Odisha, India, the results demonstrated that SVM yielded the best Kappa accuracy (0.8742), followed by RF (0.7872) and MD (0.413). SVM also produced the best results for [16] compared to RF, k-Nearest Neighbors (KNN), and NN during their study for change detection analysis in a particular region surrounding Nirma University, Ahmedabad, India, using S-2 data.

As a result, there is no common suitable classifier suggested among researchers, and each classifier often produces inconsistent results. Recently, Das *et al.* [17] attempted LULC classification by implementing several machine learning models, namely RF, KNN, SVM, DT, and Gradient booster, using S-2 data. Next, the performance of an ensemble of the classifiers has been investigated. Among individual classifiers, the proposed ensemble classifiers produced the best classification results compared to individual classifiers. However, a relatively small difference in OA between the ensemble method and SVM's (ranked second) has been observed. Another attempt at the ensemble classifier method was made by [18] using hyperspectral imagery and LiDAR-derived Digital Surface Model (DSM) at the University of Houston, USA.

Generally, the proposed ensemble method produced the highest OA in the study, with a 9% improvement over the single classifier algorithm. Nevertheless, for certain classes like Tree and Parking Lot, an individual classifier method outperformed the outcomes achieved by the ensemble. The classifier ensemble method is associated with increased computational steps, time, and complexity. Increased time consumption and limitations in classification accuracy are challenges faced by a machine learning technique [19]. Ebrahimy and Zhang [20] proposed Per-pixel Accuracy-based Ensemble of Extreme Learning Machine (PAELM). However, due to its complexity, the method was documented to have the longest processing time compared to other methods, including SVM, in all six experimental sites. Amine Douad *et al.* [21] conducted an experiment on multicriteria ensemble learning for multispectral image classification in Sundarbans, Bangladesh, using S-2 data. In their

study, KNN, RF, SVM, and their proposed ensemble method, namely the Analytical Hierarchy Process (AHP), were investigated. Nonetheless, the AHP shared the same accuracy with KNN and SVM at 93% but took the longest computational time. Hence, with the relatively small or no difference of improvement observed, it is advisable to reconsider before employing the classifier ensemble method.

Adrian *et al.* [22] employed a Deep Learning (DL) method by using S-1 and S-2 data over an agricultural site at Bradford Research Center in Columbia, USA. For S-2, the proposed 3D-Net was outperformed by RF. Meanwhile, for S-1 and S-2 fusion, the 3D-Net managed to top the competition with a 4.9% (OA) difference with U-Net. [23] investigated the performance of individual classifiers with the proposed U-Net using stand-alone S-1 and S-2, as well as their fusion. In this study, RF, KNN, KD-tree Nearest Neighbor (KD Tree), and ML were also considered. For S-1 imagery, U-Net outperformed the others with 54.19 Kappa accuracy, while for S-2 imagery, the U-Net lost to ML. For S-1 and S-2 fusion, the U-Net outperformed the others by 78.89 Kappa accuracy. Nonetheless, 54.19 and 78.89 Kappa accuracy are in the "good agreement" class [24]. They explained that a "very good agreement" class is in the range of 80-100.

Mishra *et al.* [25] utilized SVM in their study by manually selecting the optimal combination of textural features. Costache *et al.* [26] reported that Multilayer Perceptron (MLP) produced an acceptable performance. As multiple research studies have highlighted both the advantages and disadvantages of specific methods, it becomes crucial to address the insufficient understanding stemming from incomparable results regarding the potential and limitations of these methods. incomparable results regarding the potential and limitations of these methods.

The novelty of this study was to introduce a new approach to LULC classification by proposing the Iterative Pixel-Based Dimensional Voting (Pixel-DMV) model, which offered unique advantages over traditional methods. Unlike conventional classifiers, Pixel-DMV employed an iterative pixel-based dimensional voting mechanism, where labels were assigned and ranked based on an iterative analysis of pixel similarities within a localized 3 x 3 neighbourhood. This approach enhanced classification accuracy by considering contextual relationships between adjacent pixels, a feature that many existing methods overlooked. Another distinguishing feature of Pixel-DMV was its independence from multi-source data fusion. Traditional classification methods often required integrating multiple remote sensing datasets, which increased costs and complexity. Pixel-DMV eliminated this requirement by achieving high classification accuracy using a single dataset. Moreover, it did not rely on extensive training data or complex preprocessing, making it a streamlined and cost-effective alternative to deep learning and ensemble-based classification models. The adaptability of Pixel-DMV to dynamic environments further enhanced its applicability. Given its ability to operate with frequent and rapid classification, it was well-suited for monitoring fast-changing landscapes such as mixed-land areas and coastal environments impacted by human activities and natural changes. Furthermore, it improved accuracy in challenging classification environments by addressing common issues such as cloud cover, shadows, and variations in land cover types.

1.1 Study Area

Miri district of Sarawak state serves as a representative example of a coastal mixed-land zone. It is situated at latitude 04° 21' N and longitude 113° 55' E. Sarawak is renowned for its agriculture, logging, fishery sectors, and tourism. Miri, often referred to as the "City of Oil", is located approximately 797 km from the Kuching district, the capital of Sarawak. Given its distinctive characteristics and coastal proximity to the South China Sea, Miri is categorized as a mixed-land zone. Figure 1 presents an overview of the Miri district. Situated immediately north of the Equator, Sarawak spans between latitude 0° 50' and 5° N and longitude 109° 36' and 115° 40' E. Extending approximately 800 kilometers along the northwest coast of Borneo, the state covers a vast area of 124,449.51 square kilometers.

The South China Sea acts as a natural boundary, separating 600 kilometers between Sarawak and Peninsular Malaysia. To the northeast, Sarawak shares a direct connection with the State of Sabah, with the Sultanate of Brunei forming a double enclave. Inland, Sarawak shares its borders with Kalimantan, Indonesia. Sarawak is geographically divided into three distinct regions: the coastal lowlands characterized by peat swamp, narrow deltaic, and alluvial plains; a vast expanse of undulating hills with elevations reaching about 300 meters; and the mountain highlands stretching to the Kalimantan border. The region is traversed by the Batang Rajang, Malaysia's longest river, originating from the Iran Mountain range and flowing southwest to Kapit. From Kapit, it takes a westward course, covering a total distance of 563 kilometers before reaching the South China Sea. Notably, the Batang Rajang is navigable for 130 kilometers to Sibu by ocean-going vessels and extends an additional 160 kilometers into the otherwise inaccessible interior through shallow-draft vessels and small boats such as sampans.

1.2 Dataset

In this study, the remote sensing datasets consisted of imagery from Pleiades and SPOT 7 with spatial resolutions of 0.5 m and 1.5 m, respectively. The Pleiades dataset was obtained in July 2022, while the SPOT 7 dataset was obtained in May 2022. The selected projection was RSO East Malaysia (Borneo) with Timbalai 1948 datum to be geometrically co-locatable. Each imagery dataset was taken over 12 sq/km total of enclosed area across the Miri district in the state named Sarawak.

These datasets were obtained from the Malaysian Space Agency (MYSA: https://www.mysa.gov.my/) and consisted of panchromatic (PAN) and multispectral sensors. These multispectral bands were red (R; 0.63-0.69 μm), green (G; 0.52-0.59 μm), blue (B; 0.45-0.52 μm), and near-infrared (NIR; 0.77-0.89 μm). In addition to the PAN, RGB, and NIR bands, the SPOT 7 dataset also has a short-wave infrared (SWIR) band, which is often used for vegetation analysis. Both datasets offered advantages over current land-cover datasets due to their extensive coverage, broad distribution, and high spatial resolution.

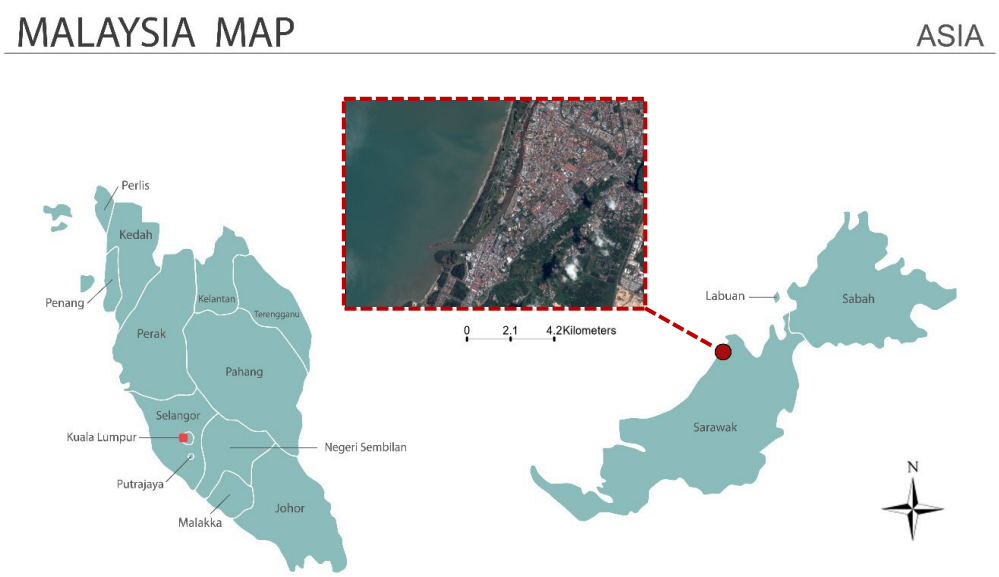


Figure 1. Miri district, Sarawak [27].

2. METHODOLOGY

Traditional pixel-based LULC classification involves analyzing remotely sensed imagery, where each pixel is classified independently based on its spectral signature. This method typically uses algorithms such as ML classification, DT, or SVM, which evaluate the reflectance values in various spectral bands to assign each pixel to a specific LULC category, like forest, urban, water, or agricultural land. The classification process relies on training data, where known samples of each class are used to train the classifier, enabling it to categorize the entire image based on similarities in spectral properties. While effective, this approach can sometimes misclassify pixels due to mixed land cover types or spectral similarities between different classes. The increased variation in the spectral characteristics of a specific type of land cover can result in misidentified pixels within categories, often causing inconsistent pixel classification [28].

This study introduced a voting model designed to enhance the traditional pixel-based image classification, with the goal of improving accuracy and reducing misclassifications. Figure 2 illustrates the proposed Pixel-DMV model, which began with an image filtering process, followed by feature extraction. This extraction involved analyzing intensity histograms for each class within the segment of interest (SOI) in the training set images. The classification process started by reading the image pixel-by-pixel and comparing each pixel with the information from the training set. When an unknown pixel was encountered, the algorithm initiated an iterative labeling process within a 3 x 3-pixel area surrounding the unknown pixel to determine potential pixel labels. Using the possible labels identified in this neighborhood, similarity measures were computed. Based on these measures, the pixel labels were ranked to predict the most likely class for the unknown pixel. Algorithm 1 shows a more detailed explanation of the Pixel-DMV model.

2.1 Image Pre-Processing

In optical images, such as Pleiades and SPOT 7, various types of noise, such as Gaussian noise, salt-and-pepper noise, and atmospheric noise, significantly degraded data quality and hindered accurate interpretation. This underscores the importance of selecting the most effective image filtering methods, as they are essential for enhancing image quality, eliminating noise, and extracting critical features necessary for further analysis.

In [29], Mean, Median, and Wiener filters were compared for image denoising, with the Median filter proving to be the most effective for denoising the majority of their noisy images. Similarly, Abdullah *et al.* [30] conducted a qualitative and quantitative investigation of Gaussian, Median, and Wiener filters, further exploring their performance in denoising applications. Among the image filtering results, the Median filter proved to be the most effective at preserving image edges. On the other hand, the Wiener filter stood out for its ability to deblur the image background and enhance grayscale contrast. Although the Wiener filter successfully maintained black edge colors and sharpness, its performance diminished as noise density increased. In high-noise scenarios, while the filter managed to deblur the noise, it struggled to retain the original characteristics of the image. Similarly, even in low noise-density conditions, the Gaussian filter faced challenges in preserving edges, ultimately compromising the overall integrity of the image.

Recently, Ahamed *et al.* [31] compared the performance of Gaussian, Median, and Wiener filters. The Median filter demonstrated effectiveness in restoring noisy images affected by noise and motion blur. While the Wiener filter performed well in deblurring motion blur, its reliability in restoring images degraded by additive noise remained limited. The Gaussian filter offered certain benefits; however, it had limitations, particularly in its inability to fully eliminate white and black pixel impulse noise. In a brief analysis conducted by [32], Gaussian, Median, Mean, Wiener, and Average filters were evaluated. Among these, the Median filter demonstrated the best performance. Consequently, the Median filter was chosen for this study.

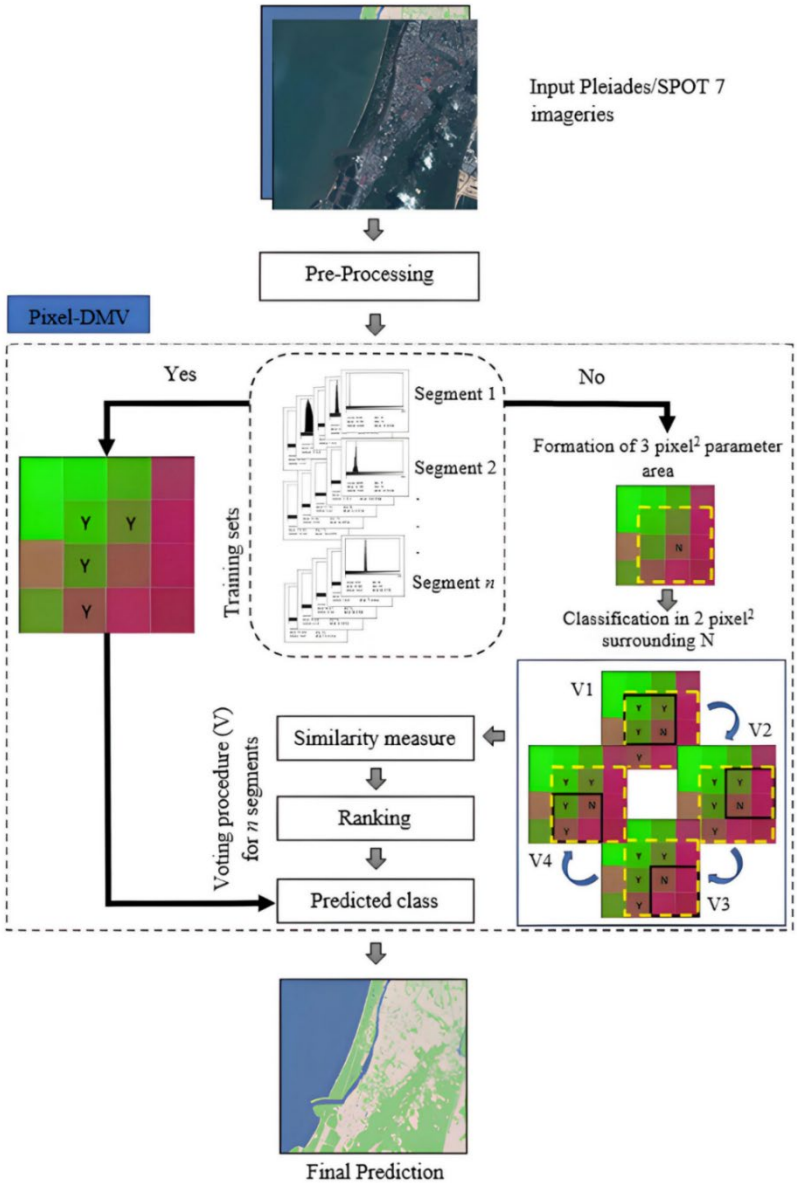


Figure 2. The proposed Pixel-DMV model.

2.2 Feature Extraction

Unlike [33], which combined similar land cover types into 6 broader classes (such as grouping rivers, lakes, and reservoirs into a single Water class), this study categorized vegetation, urban terrain, water, and natural ground segments into 12 distinct land cover classes. While this approach increased computational complexity, the risk of spectral similarity issues, and potential classification errors, it tested the efficiency of the proposed method in handling these challenges and ensured a more precise and reliable classification. The dataset consisted of high-resolution images acquired from the Pleiades and SPOT 7 satellites, with resolution grids of 12,288 × 9,178 and 5,853 × 4,590, respectively. However, small regions were excluded due to the sensor's malfunctions.

To assess the efficacy of classification algorithms in generating high-quality outcomes with limited training data, the datasets were divided into training and testing samples. Ouchra *et al.* [34] utilized between 2.96% to 18.84% of the total area as training samples while the rest were testing samples. In this study, 10% of the total area was randomly selected for training, while the remaining 90% was reserved for testing. Table 1 and Figure 3 show the distribution of train and test samples for the Pleiades dataset, while Table 2 and Figure 4 show the distribution of train and test samples for the SPOT 7 dataset. Training samples were labelled data points used to develop machine learning models, enabling them to recognize patterns in different land cover classes. Test samples, on the other hand, served as independent data points to evaluate the model's performance and ensure its ability to classify new, unseen data accurately. The classification process involved analyzing the reflectance for each pixel and determining the most similar signature.

Algorithm 1: Pixel-DMV

input
Rsd: Pleiades and SPOT 7.
Ct: Pixel-DMV.
R, G and *B*: Red, Green and Blue band, respectively.
label: a function that returns a pixel label for every pixel.
output
N: Target pixel.
n: Number of dimensional voting surrounding *N*, $n = 4$.
V_i, $i = 1,\ldots n$: Order of each dimensional voting surrounding *N*.
R_i, $i = 1,\ldots 12$: Rank of matched class surrounding *N*.
begin
for (preprocessing)
    Apply Median filter.
end for
Perform signature analysis of various land cover class in AOI as training sets.
Conduct testing sets.
for every pixel *N*
    Find *RGB* value of *N* in training sets.
    if found, **do**
        label pixel *N* with matched class.
    else
        Form 3 x 3-pixel parameter area surrounding *N*.
        for ($V_i <= V_n$)
            Form 2 x 2-pixel parameter area surrounding *N*.
            Find *RGB* value of pixels surrounding *N* in training sets.
            if found, **do**
                label the surrounding pixels with matched class.
            else
                Repeat step 21 to 35.
            end if
        end for
    end if
end for
for land cover class determination of pixel *N*, **do**
    Rank matched class surrounding pixel *N* with R_1, R_2, R_3 and so on.
    label pixel *N* with R_1.
end for
end

To assess the efficacy of classification algorithms in generating high-quality outcomes with limited training data, the datasets were divided into training and testing samples. Ouchra *et al.* [34] utilized between 2.96% to 18.84% of the total area as training samples while the rest were testing samples. In this study, 10% of the total area was randomly selected for training, while the remaining 90% was reserved for testing. Table 1 and Figure 3 show the distribution of train and test samples for the Pleiades dataset, while Table 2 and Figure 4 show the distribution of train and test samples for the SPOT 7 dataset. Training samples were labelled data points used to develop machine learning models, enabling them to recognize patterns in different land cover classes. Test samples, on the other hand, served as independent data points to evaluate the model's performance and ensure its ability to classify new, unseen data accurately. The classification process involved analyzing the reflectance for each pixel and determining the most similar signature.

Table 1. Pleiades dataset with number of train and test samples.

	Class	Train samples	Test samples
1	Road	315,255	2,837,293
2	Artificial ground	3,529	31,762
3	Buildings	2,951,624	26,564,612
4	Trees	1,658,647	14,927,823
5	Healthy grass	678,916	6,110,243
6	Stressed grass	415,129	3,736,164
7	Ocean	4,639,628	41,756,648
8	Lake	14,256	128,302
9	Swimming pool	601	5,409
10	Swamp water	156,137	1,405,233
11	Bare earth	206,506	1,858,558
12	Sand	237,699	2,139,292
	Total	11,277,926	101,501,338
	Percentage (%)	10	90

Table 2. SPOT 7 dataset with number of train and test samples.

	Class	Train samples	Test samples
1	Road	36,952	332,572
2	Artificial ground	1,358	12,226
3	Buildings	707,047	6,363,424
4	Trees	744,152	6,697,365
5	Healthy grass	185,299	1,667,688
6	Stressed grass	63,048	567,428
7	Ocean	847,162	7,624,454
8	Lake	11,813	106,314
9	Swimming pool	58	518
10	Swamp water	36,697	330,276
11	Bare earth	39,627	356,641
12	Sand	12,842	115,574
	Total	11,277,926	101,501,338
	Percentage (%)	10	90

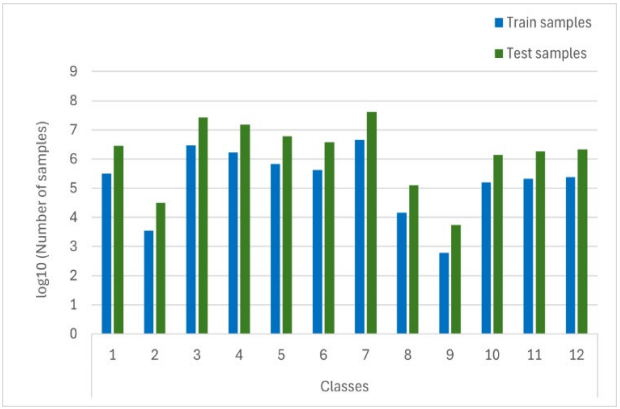


Figure 3. Logarithmic bar chart representing train and test samples for Pleiades dataset.

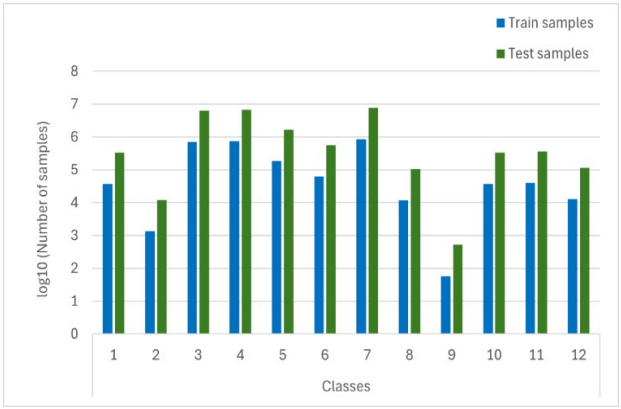


Figure 4. Logarithmic bar chart representing train and test samples for the SPOT 7 dataset.

2.3 Supervised Classification Methods

In this study, we employed several well-known classifiers to validate our method. These classifiers, namely SVM, RF, and ML, were selected based on their recent outstanding performances. For SVM, the Radial Basis Function (RBF) kernel is employed since it works well in most cases [35].

2.4 Accuracy Assessment

Accuracy assessment involves evaluating the performance and reliability of the classification results compared to ground truth data or reference data. The classified map was compared against the reference data in confusion matrices. Then, True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) are computed. The performance was evaluated using the User's Accuracy (UA), Producer's Accuracy (PA), Average Accuracy (AA), OA, and Kappa. A good classification result is indicated by high UA, PA, AA, OA, and Kappa values.

The UA measures the probability that a pixel classified on the map represents the land cover type indicated by that classification, as shown in Equation (1).

$$UA = \frac{TP}{TP + FP} \tag{1}$$

The PA measures the probability that a pixel on the ground of a particular land cover type is classified as such on the map, as shown in Equation (2).

$$PA = \frac{TP}{TP + FN} \tag{2}$$

The AA measures the performance across the entire landscape and provides a comprehensive evaluation of its effectiveness in accurately mapping land cover types, as shown in Equation (3).

$$AA = \frac{\sum Individual\ Class\ Accuracies}{Number\ of\ Class} \tag{3}$$

The OA determines the degree of closeness between the classified image and the reference data, as shown in Equation (4).

$$OA = \frac{TP + TN}{TP + FP + FN + TN} \tag{4}$$

Kappa, as shown in Equation (5), assesses the agreement between the classification results and the expected classification results by chance alone while accounting for the possibility of agreement occurring by random chance. It considers both the OA of the classification and the proportion of agreement that could be expected by chance. Kappa values range from -1 to 1, where 1 indicates perfect agreement, 0 indicates agreement equivalent to chance, and negative values suggest agreement worse than chance.

$$Kappa = \frac{Pr(a) - Pr(e)}{1 - Pr(e)} \tag{5}$$

where $Pr(a)$ is the relative observed agreement among raters and $Pr(e)$ is the hypothetical probability of chance agreement whereby the observed data is used to calculate the probabilities of each observer randomly saying each category.

3. EXPERIMENTAL RESULTS AND DISCUSSION

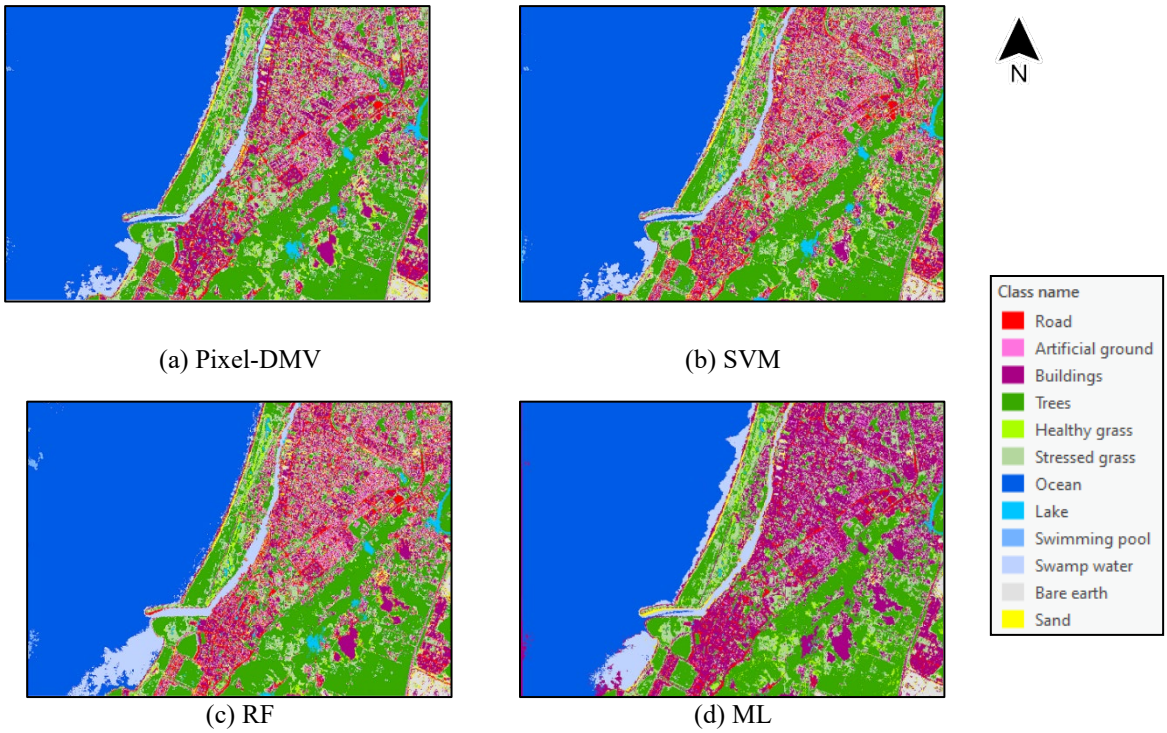
3.1 Pleiades Results

Figure 5 presents an overview of classification results obtained from different algorithms for Pixel-DMV, SVM, RF, and ML, providing valuable insights into their performance across various land cover classes. Table 3 shows the accuracy analysis with distinct land cover types that were identified in this study. Each percentage value represented the precision of the classification algorithms in distinguishing specific land cover categories. For Pixel-DMV, the analysis began with the identification of the Road class and achieved an accuracy of 91.81%, indicating a robust capability in discerning road networks within the studied area. Similarly, the classification achieved a commendable accuracy of 80.96% for the Artificial Ground class, signifying the successful identification of developed or urbanized regions. The buildings class was identified with 93.65% accuracy, reflecting the algorithm's effectiveness in isolating built structures.

The algorithm also demonstrated exceptional accuracy in identifying vegetative features. The Trees, Healthy Grass, and Stressed Grass classes achieved accuracies of 94.37%, 82.27%, and 95.32%, respectively. This highlighted the algorithm's capability to differentiate between healthy and stressed vegetation, which is crucial for environmental monitoring and habitat assessment. Additionally, the classification results indicated strong performance in identifying aquatic features. The Ocean class was classified with an exceptional accuracy of 99.77%, underscoring the algorithm's proficiency in delineating large water bodies. However, the accuracy for the Lake classification was relatively lower at 72.41%, suggesting potential challenges in accurately identifying smaller water bodies within the landscape.

The classification also excelled in identifying human-made features, such as the Swimming Pool class with 95.47% accuracy. Natural water bodies like Swamp Water were identified with notable precision at 99.26%. These results underscored the algorithm's effectiveness in distinguishing between different water types, which is critical for hydrological studies and ecosystem management. Additionally, the classification process successfully identified areas of minimal vegetation cover, with Bare Earth and Sand achieving accuracy of 76.75% and 96.80%, respectively. These results indicated the algorithm's ability to differentiate between barren land surfaces and sandy terrains, which is valuable for geological mapping and land use planning efforts. To summarize the overall performance, the proposed method achieved an OA of 95.35%, serving as a fundamental indicator of the classification methodology's effectiveness in accurately identifying and categorizing different land cover types within the study area. Moreover, the AA of 89.90% and a Kappa of 0.94 further reinforced the classification's consistency and reliability across individual land cover classes.

A more detailed analysis, as presented in Figure 6, revealed that Pixel-DMV achieved exceptional accuracy for most classes based on the obtained UA and PA. For example, the Roads class achieved a UA of 78.83%, indicating that approximately 78.83% of the pixels designated as Roads on the map corresponded to actual Road areas on the ground. Conversely, the PA for the Roads class stood at 91.81%, meaning that 91.81% of the ground pixels classified as Roads were accurately identified on the map. Similarly, the Buildings, Trees, and Stressed Grass classes achieved remarkable success, with UA values of 96.60%, 92.19%, and 85.82%, respectively. This reflected precise identification of building structures and tree-covered areas on the map. The corresponding PA values of 93.65% for Buildings, 94.37% for Trees, and 95.32% for Stressed Grass further underscored the fidelity of the classification process in capturing these features on the ground.



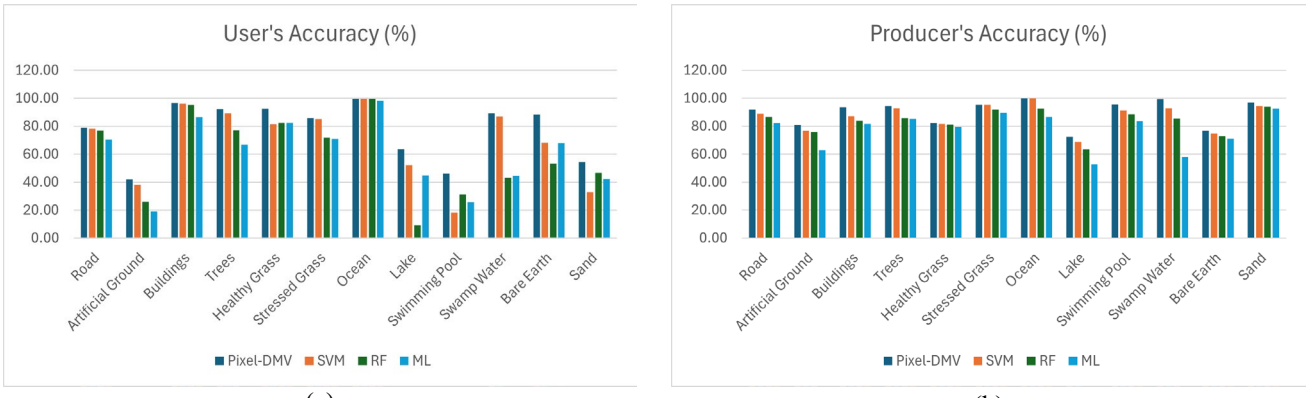
(a) Pixel-DMV (b) SVM

(c) RF (d) ML

Figure 5. Overview of classification results using the Pleiades dataset.

Table 3. Accuracy analysis of Pixel-DMV, SVM, RF, and ML (in %) for the Pleiades dataset.

Class name	Pixel-DMV	SVM [14][36][37][38][39]	RF [8][37][40][41]	ML [14]
Road	91.81	88.95	86.69	82.26
Artificial Ground	80.96	76.83	75.84	62.99
Buildings	93.65	87.25	83.92	81.67
Trees	94.37	92.57	85.88	85.17
Healthy Grass	82.27	81.77	81.06	79.34
Stressed Grass	95.32	95.18	91.82	89.66
Ocean	99.77	99.75	92.33	86.66
Lake	72.41	68.64	63.41	52.79
Swimming Pool	95.47	91.33	88.69	83.73
Swamp Water	99.26	92.82	85.66	57.90
Bare Earth	76.75	74.74	72.96	70.96
Sand	96.80	94.41	93.73	92.38
OA	**95.35**	**93.04**	**87.70**	**83.88**
AA	**89.90**	**87.02**	**83.50**	**77.12**
Kappa	**0.94**	**0.90**	**0.83**	**0.78**



(a) (b)

Figure 6. UA and PA of Pixel-DMV, SVM, RF and ML using Pleiades dataset. (a) UA. (b) PA.

The Healthy Grass class demonstrated a UA of 92.57%, indicating that approximately 92.57% of the pixels classified as Healthy Grass on the map accurately represented Healthy Grass areas on the ground. The corresponding PA for Healthy Grass stood at 82.27%, suggesting that 82.27% of the ground pixels classified as Healthy Grass were correctly identified on the map. Meanwhile, Swamp Water achieved a UA of 89.19% and a PA of 99.26%, highlighting accurate detection of swampy regions. The Ocean class also performed exceptionally well, with UA and PA values of 99.74% and 99.77%, respectively, attesting to the algorithm's proficiency in identifying oceanic areas. The Bare Earth class displayed a UA of 88.36%, indicating accurate identification of Bare Earth areas on the map. However, the corresponding PA for Bare Earth was 76.75%, suggesting potential challenges in accurately identifying Bare Earth areas on the ground. The Lake class obtained a UA of 63.48% and a PA of 72.41%, indicating room for improvement in depicting Lake areas on the ground.

Among the classes, Artificial Ground, Swimming Pool, and Sand displayed notable disparities between UA and PA. For the Artificial Ground class, the UA of 42.03% indicated that only 42.03% of the pixels classified as Artificial Ground on the map accurately represented actual Artificial Ground areas. In contrast, the PA for this class was significantly higher at 80.96%, meaning that 80.96% of the actual Artificial Ground areas on the ground were correctly identified on the map. Similarly, the Swimming Pool class showed a UA of 46.08%, suggesting that less than half of the pixels classified as Swimming Pool on the map accurately corresponded to actual Swimming Pool areas. However, the PA for this class was much higher at 95.47%, indicating that 95.47% of the actual Swimming Pool areas on the ground were correctly mapped. Meanwhile, the Sand class exhibited 54.50% of UA, meaning that just over half of the pixels classified as Sand on the map accurately represented Sand areas on the ground. On the other hand, the PA for Sand was remarkably high at 96.80%, demonstrating a strong ability to correctly identify Sand areas on the ground. These discrepancies highlighted varying levels of precision and accuracy across the different classes.

Compared to other methods in this study, Pixel-DMV achieved the highest OA at 95.35%, followed by SVM at 93.04%, RF at 87.70%, and ML at 83.88%. These values highlighted Pixel-DMV's superior ability to classify pixels accurately across all land cover classes. Similarly, Pixel-DMV attained the highest AA at 89.90%, followed by SVM at 87.02%, RF at 83.50%, and ML at 77.12%. These results demonstrated Pixel-DMV's capability to classify different land cover types consistently and accurately, making it a reliable choice for remote sensing analysis tasks. The Pixel-DMV model also demonstrated exceptional agreement with a Kappa accuracy of 0.94, indicating highly accurate predictions across classes such as Roads, Trees, Ocean, and Stressed Grass. In contrast, the SVM, RF, and ML methods exhibited lower Kappa accuracies of 0.90, 0.83, and 0.78, respectively, suggesting less consistent agreement between predictions and actual classes, particularly for challenging classes like Lake and Swamp Water.

Pixel-DMV consistently outperformed SVM, RF, and ML algorithms across all categories. For instance, Pixel-DMV achieved an accuracy of 99.26% for the Swamp Water class, while SVM, RF, and ML achieved 92.82%, 85.66%, and 57.90%, respectively. In classifying the Stressed Grass class, Pixel-DMV achieved 95.32% accuracy, outperforming SVM (95.18%), RF (91.82%), and ML (89.66%). Pixel-DMV also demonstrated remarkable accuracy in distinguishing the Ocean and Swimming Pool classes, with accuracies of 99.77% and 95.47%, respectively, while SVM, RF, and ML lagged behind in these categories. For the Ocean class, SVM obtained 99.75%, followed by RF (92.33%) and ML (86.66%). For the Swimming Pool class, SVM achieved 91.33%, while RF (88.69%) and ML (83.73%) followed. For the Road class, Pixel-DMV exhibited the highest accuracy at 91.81%, followed by SVM (88.95%), RF (86.69%), and ML (82.26%). In the case of the Artificial Ground class, Pixel-DMV demonstrated the most accurate classification at 80.96%, outperforming SVM (76.83%), RF (75.84%), and ML (62.99%).

Similarly, in the classification of the buildings class, Pixel-DMV achieved an accuracy of 93.65%, followed by SVM (87.25%), RF (83.92%), and ML (81.67%). For the Trees class, Pixel-DMV achieved the highest accuracy at 94.37%, followed by SVM with 92.57%, RF with 85.88%, and ML with 85.17%. For the Healthy Grass class, Pixel-DMV attained an accuracy of 82.27%, while SVM, RF, and ML achieved 81.77%, 81.06%, and 79.34%, respectively. In the case of the Bare Earth class, Pixel-DMV exhibited an accuracy of 76.75%, surpassing SVM (74.74%), RF (72.96%), and ML (70.96%). Pixel-DMV also demonstrated superior accuracy in classifying the Sand class at 96.80%, followed by SVM (94.41%), RF (93.73%), and ML (92.38%). For the Lake class, Pixel-DMV (72.41%) outperformed the other methods by noticeable margins, with SVM, RF, and ML achieving 68.64%, 63.41%, and 52.79%, respectively.

Overall, the results underscored the efficacy of the Pixel-DMV algorithm in remote sensing classification, highlighting its ability to achieve higher accuracies and outperform other methods across diverse land cover classes. These findings emphasized the importance of selecting appropriate classification algorithms tailored to specific application needs to ensure accurate and reliable land cover mapping in remote sensing analysis.

In this study, several significant challenges were identified, including confusion factors, cloud cover, and shaded areas. As shown in Figure 7, examples alongside corresponding classified images revealed pertinent issues encountered during the classification process. In the first image, some regions from the Ocean class were misclassified as the Swamp Water class due to the close similarity in color information. Similarly, in the second image, some regions from the Sand class were misclassified as the buildings class due to comparable color characteristics, acting as interference for the classification algorithms. Moreover, the presence of shade and cloud cover was evident in the third and fourth images, respectively.

Despite these challenges, the Pixel-DMV voting mechanism successfully classified certain regions within shaded and cloudy areas. Notably, the method accurately identified some areas as Trees and Healthy Grass classes by considering neighboring pixels, thereby mitigating the misclassification of areas as Lake and Buildings classes in the third and fourth images, respectively. These findings underscored the complexities involved in remote sensing classification and highlighted the importance of developing robust methodologies capable of addressing various challenges, including confusion factors and environmental conditions, to enhance the accuracy and reliability of classification outcomes.

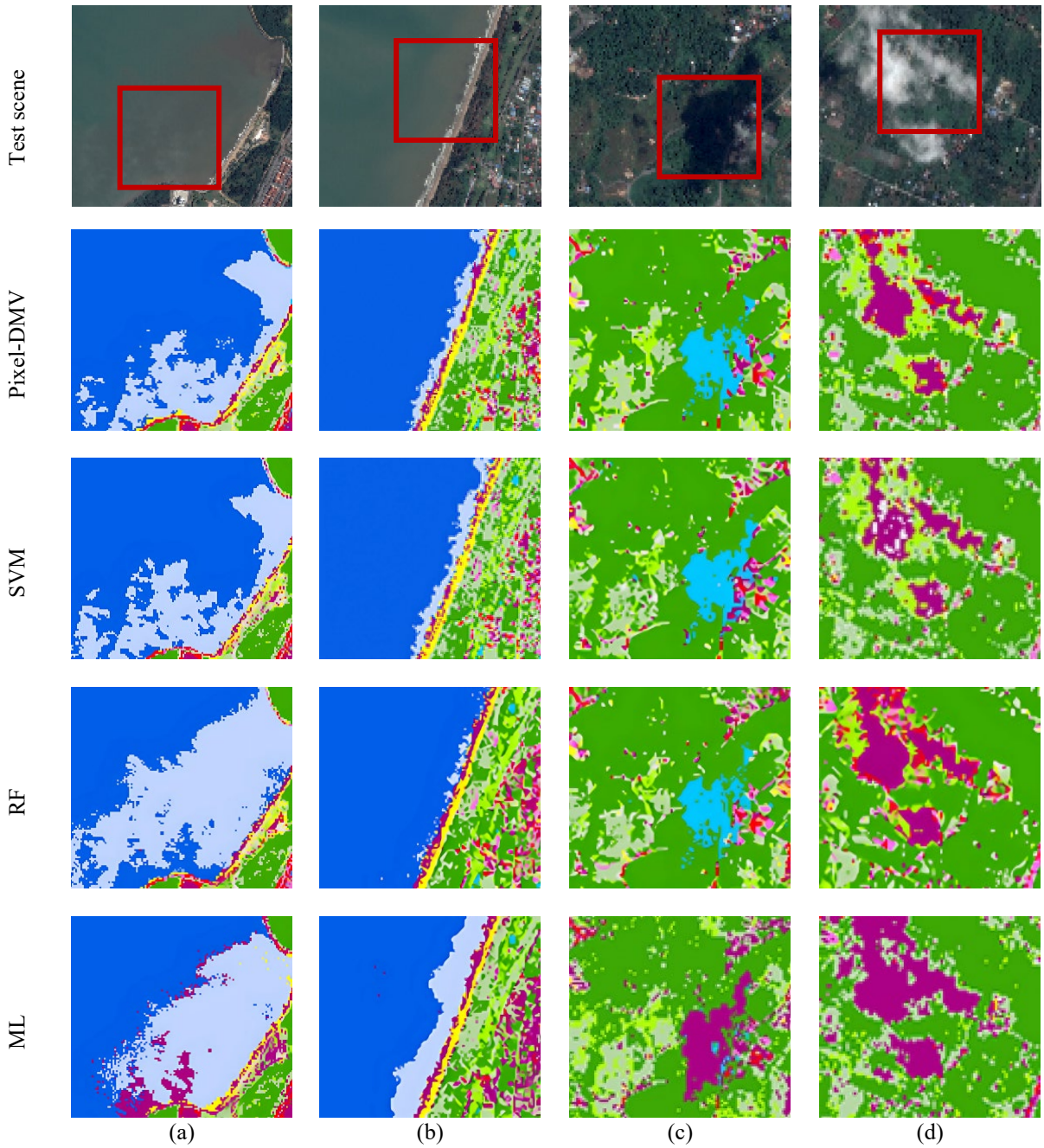


Figure 7. Classified images of Pixel-DMV, SVM, RF, and ML using the Pleiades dataset with the presence of confusion factors and shaded as well as cloudy areas (highlighted in red box)

3.2 SPOT 7 Results

Figure 8 provides a visual summary of the classification results, encompassing Pixel-DMV, SVM, RF, and ML, offering valuable insights into their efficacy across different land cover classes. Table 4 presents a comparative analysis of classification results obtained from these methods, offering valuable insights into the performance of the Pixel-DMV method when applied to the SPOT 7 dataset. The results highlighted the precision with distinct land cover types were identified through remote sensing techniques. For instance, the Road class was classified with an accuracy of 89.45%, demonstrating the method's robust capability in discerning road networks within the study area. Similarly, the Artificial Ground class achieved a commendable accuracy of 65.43%, indicating successful identification of developed or urbanized regions.

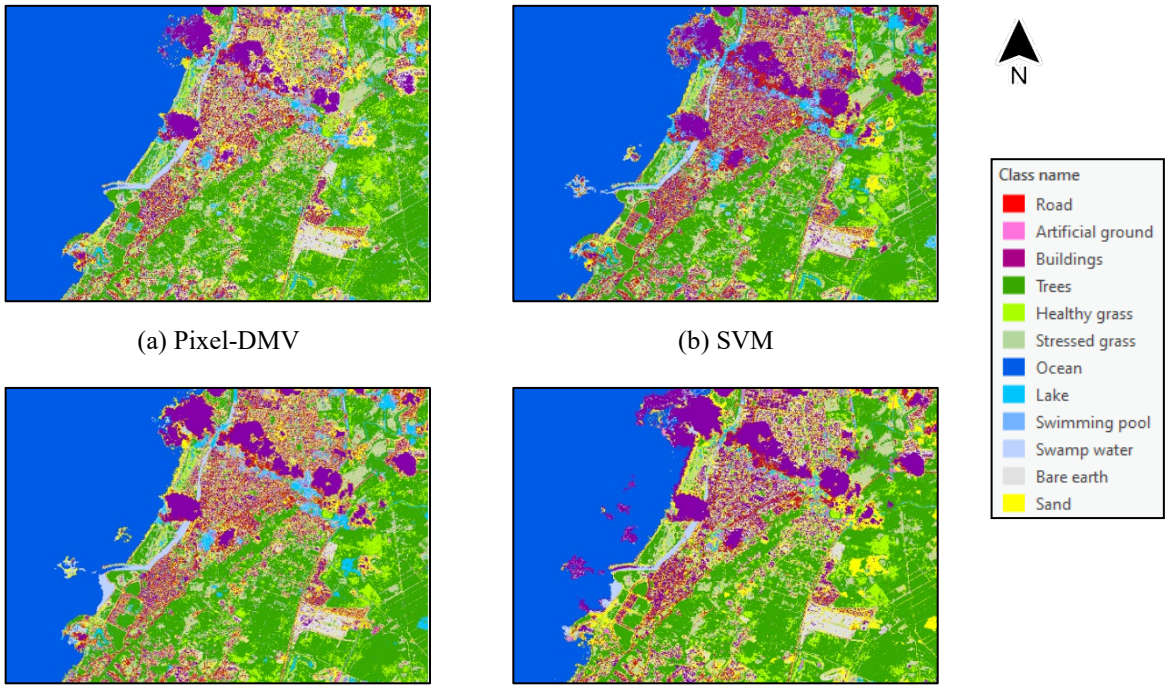
The method also excelled in identifying various other land cover types. Buildings were detected with high precision at 96.63%, reflecting the algorithm's effectiveness in isolating built structures. Vegetative features such as Trees, Healthy Grass, and Stressed Grass were classified with accuracies of 88.57%, 74.07%, and 89.13%, respectively, showcasing the method's proficiency in differentiating between healthy and stressed vegetation.

In addition to terrestrial features, the Pixel-DMV method performed exceptionally well in classifying aquatic features. The Ocean class was identified with a remarkable accuracy of 99.29%, while the Lake and Swamp Water classes achieved accuracies of 85.02% and 97.87%, respectively. These results underscored the method's ability to distinguish between different water types, which is crucial for hydrological studies and ecosystem management. Human-made features, such as Swimming

Pools, were also accurately detected, with an accuracy of 92.36%. Furthermore, areas with minimal vegetation cover, such as Bare Earth and Sand, were identified with accuracies of 83.94% and 94.51%, respectively, highlighting the method's capability to differentiate between diverse land cover categories.

Overall, the Pixel-DMV method achieved an OA of 93.16% across all classes, serving as a strong indicator of its effectiveness in accurately identifying and categorizing various land cover types within the SPOT 7 dataset. The AA of 88.02% further underscored the method's consistency and reliability across individual land cover classes, making it highly suitable for remote sensing applications and environmental monitoring. The inclusion of a Kappa accuracy of 0.91 provided additional insight into the agreement between observed and expected classifications, further validating the method's performance.

A closer examination of Figure 9 revealed the UA and PA percentages for various land cover categories classified by all methods using the SPOT 7 dataset. For the Road class, the UA was 69.60%, indicating that the algorithm correctly identified approximately 69.60% of the Road pixels. In contrast, the PA for the Road class was notably higher at 89.45%, reflecting the algorithm's effectiveness in accurately classifying Road pixels relative to the reference data. Similarly, for the Artificial Ground class, the UA stood at 48.47%, while the PA was 65.43%, suggesting a relatively higher level of accuracy in correctly classifying Artificial Ground surfaces.



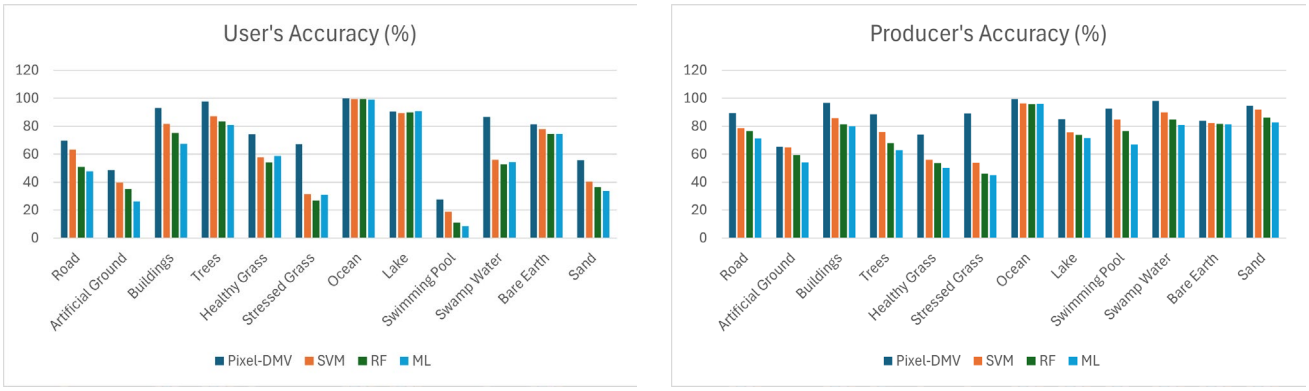
(a) Pixel-DMV (b) SVM

(c) RF (d) ML

Figure 8. Overview of classification results using the SPOT 7 dataset.

Table 4. Accuracy Analysis of Pixel-DMV, SVM, RF, and ML (in 100%) for the SPOT 7 dataset.

Class name	Pixel-DMV	SVM [15][17][20][21][42]	RF [15][17][21] [22][40]	ML [23]
Road	89.45	78.63	76.51	71.24
Artificial Ground	65.43	64.71	59.20	53.95
Buildings	96.63	85.82	81.50	80.05
Trees	88.57	75.92	67.95	62.83
Healthy Grass	74.07	55.94	53.53	50.30
Stressed Grass	89.13	53.91	46.16	44.99
Ocean	99.29	96.28	95.71	95.92
Lake	85.02	75.64	73.69	71.56
Swimming Pool	92.36	84.72	76.39	66.84
Swamp Water	97.87	89.84	84.83	80.75
Bare Earth	83.94	82.12	81.72	81.34
Sand	94.51	91.99	86.02	82.82
OA	**93.16**	**83.44**	**79.42**	**77.28**
AA	**88.02**	**77.96**	**73.60**	**70.22**
Kappa	**0.91**	**0.78**	**0.73**	**0.70**



(a) (b)

Figure 9. UA and PA of Pixel-DMV, SVM, RF, and ML using SPOT 7 dataset. (a) UA. (b) PA.

Meanwhile, for the Healthy Grass class, the UA was reported at 74.27%, indicating that approximately 74.27% of the pixels classified as Healthy Grass were correctly identified. Similarly, the PA for the Healthy Grass class was 74.07%, suggesting that the algorithm accurately classified 74.07% of the reference pixels representing Healthy Grass areas. The buildings class demonstrated a high UA of 93.03% and a PA of 96.63%, indicating the algorithm's strong capability in identifying built structures. For Trees, both the UA and PA were relatively high, at 97.59% and 88.57%, respectively, showcasing the method's effectiveness in accurately identifying tree-covered areas. The Bare Earth class also showed satisfactory accuracy, with a UA of 81.32% and a PA of 83.94%. Aquatic features were particularly well-classified, with the Ocean class achieving a UA of 99.79% and a PA of 99.29%. The Lake class had a UA of 90.50% and a PA of 85.02%, while the Swamp Water class exhibited a UA of 86.69% and a PA of 97.87%. These results highlighted the method's proficiency in accurately identifying different water bodies.However, some classes exhibited discrepancies between UA and PA. For example, the Swimming Pool class had a notably low UA of 27.68%, indicating challenges in correctly identifying Swimming Pool areas, while its PA was relatively high at 92.36%. Similarly, the Stressed Grass class had a UA of 67.04% and a PA of 89.13%, suggesting difficulties in identifying stressed vegetation. The Sand class also showed a lower UA of 55.78% compared to its PA of 94.51%, indicating challenges in correctly identifying sandy areas.

In terms of overall performance, Pixel-DMV ranked first among all methods, achieving an OA of 93.16% and an AA of 88.02%. This performance significantly surpassed that of SVM (OA: 83.44%, AA: 77.96%), RF (OA: 79.42%, AA: 73.60%), and ML (OA: 77.28%, AA: 70.22%). The Kappa accuracy further reinforced Pixel-DMV's superiority, with a score of 0.91 compared to SVM (0.78), RF (0.73), and ML (0.70). Pixel-DMV demonstrated exceptional accuracy across various land cover classes. For instance, it achieved the highest accuracy in Road classification at 89.45%, outperforming SVM (78.63%), RF (76.51%), and ML (71.24%). For the Artificial Ground class, Pixel-DMV achieved an accuracy of 65.43%, surpassing SVM (64.71%), RF (59.20%), and ML (53.95%). This indicated Pixel-DMV's superiority in distinguishing developed or urbanized regions, an essential aspect of land cover classification. The buildings class was identified with an accuracy of 96.63%, significantly higher than SVM (85.82%), RF (81.50%), and ML (80.05%).

Similarly, Pixel-DMV performed incredibly well in identifying tree-covered areas with an accuracy of 88.57%, outperforming SVM (75.92%), RF (67.95%), and ML (62.83%). This underscored Pixel-DMV's effectiveness in delineating vegetation-rich regions, which is crucial for environmental monitoring and habitat assessment. In classifying the Healthy Grass class, Pixel-DMV achieved an accuracy of 74.07%, demonstrating its effectiveness compared to SVM (55.94%), RF (53.53%), and ML (50.30%). For the Stressed Grass classification, Pixel-DMV again led with an accuracy of 89.13% compared to SVM (53.91%), RF (46.16%), and ML (44.99%), indicating its superior ability to differentiate between healthy and stressed vegetation conditions.

Moreover, Pixel-DMV excelled in classifying oceanic areas with an accuracy of 99.29%, followed by SVM (96.28%), RF (95.71%), and ML (95.92%). It also demonstrated notable accuracy in identifying other water bodies, such as Lake, Swimming Pool, and Swamp Water. For the Lake class, Pixel-DMV demonstrated an accuracy of 85.02%, surpassing SVM (75.64%), RF (73.69%), and ML (71.56%). For the Swimming Pool class, Pixel-DMV achieved an accuracy of 92.36%, outperforming SVM (84.72%), RF (76.39%), and ML (66.84%). Moreover, for the Swamp Water class, Pixel-DMV exhibited an accuracy of 97.87%, exceeding the performance of SVM (89.84%), RF (84.83%), and ML (80.75%).

In classifying Bare Earth areas, Pixel-DMV demonstrated superior performance at 83.94%, surpassing SVM (82.12%), RF (81.72%), and ML (81.34%). In the case of the Sand class, Pixel-DMV achieved an accuracy of 94.51%, outperforming SVM (91.99%), RF (86.02%), and ML (82.82%). The success in accurately classifying these two land cover types highlighted Pixel-DMV's expertise in distinguishing between barren land surfaces and sandy terrains.

Despite its overall success, the investigation revealed several challenges, including confusion factors, cloud cover, and shaded regions. Figure 10 illustrated instances where misclassifications occurred, such as the Ocean class being labelled as Swamp Water due to similar color characteristics, or Sand being misclassified as Lake. The influence of shade and cloud cover also posed challenges, as seen in the third and fourth images. However, the Pixel-DMV voting mechanism demonstrated resilience by accurately classifying certain regions within shaded and cloudy areas, such as identifying Trees and Healthy Grass by analyzing neighboring pixels. These observations underscored the complexity of remote sensing classification and emphasized the need for robust methodologies capable of addressing diverse challenges, including environmental conditions and confusion factors, to enhance the accuracy and reliability of classification results.

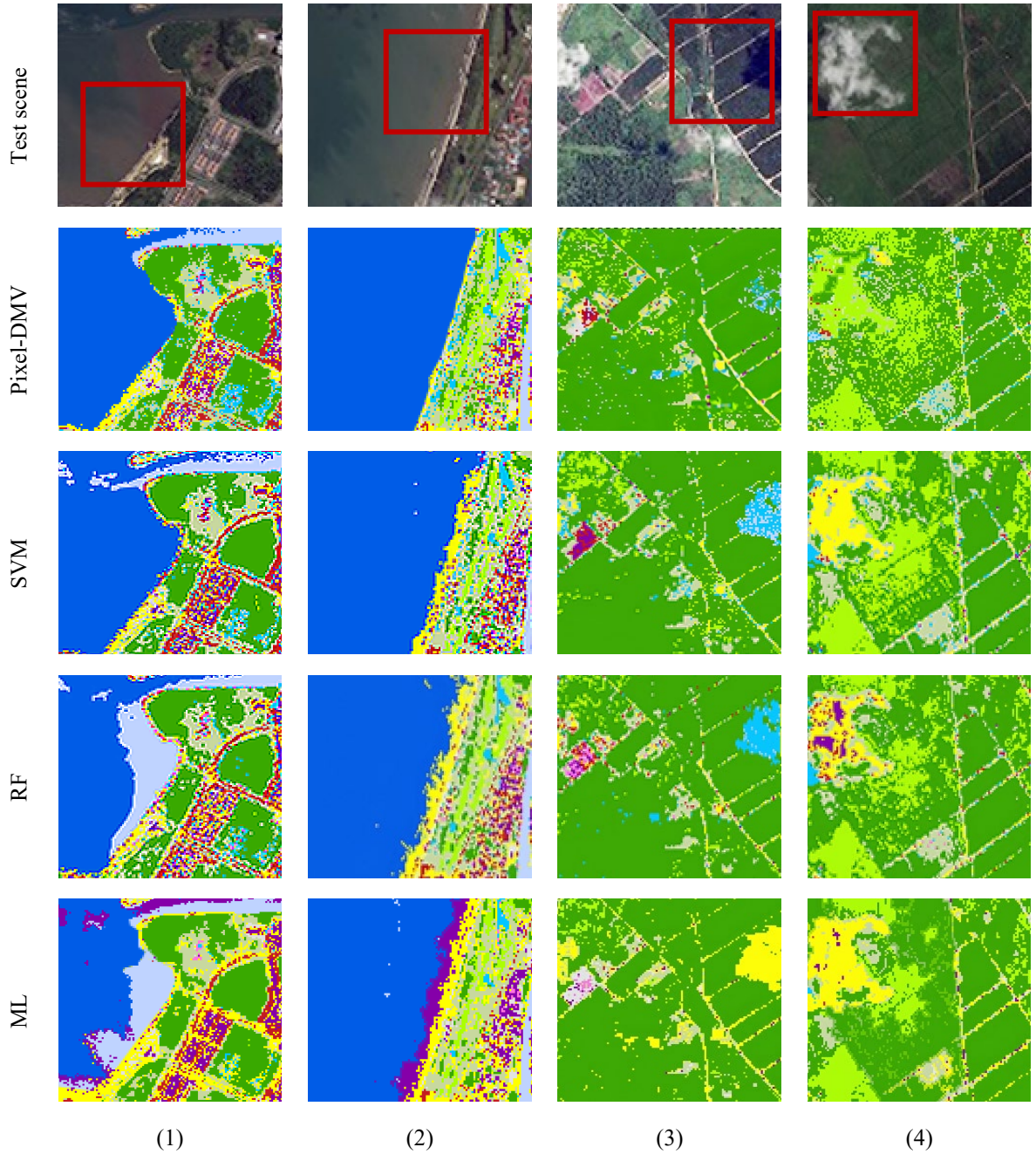


Figure 10. Classified images of Pixel-DMV, SVM, RF, and ML using the SPOT 7 dataset with the presence of confusion factors and shaded as well as cloudy areas (highlighted in red box).

3.3 Overall Results and Interpretation

The accuracy analysis of Pixel-DMV, SVM, RF, and ML for the Pleiades and SPOT 7 datasets revealed several key insights. Pixel-DMV consistently outperformed the other methods in terms of OA, AA, and Kappa across both datasets. For instance, in the Pleiades dataset, Pixel-DMV achieved OA of 95.35% and a Kappa of 0.94, compared to SVM (OA: 93.04%, Kappa: 0.90), RF (OA: 87.70%, Kappa: 0.83), and ML (OA: 83.88%, Kappa: 0.78). This indicated that Pixel-DMV had a higher agreement with the ground truth and fewer errors in classification.

Class-specific performance analysis showed that Pixel-DMV excelled in classifying classes with distinct spectral signatures, such as Ocean, Swamp Water, and Swimming Pool, achieving OA of 99.77%, 99.26%, and 95.47% respectively in the Pleiades dataset. This could be attributed to the algorithm's ability to effectively capture and utilize the unique spectral characteristics of these classes through its dimensional voting mechanism. Challenging classes like Artificial Ground and Bare Earth, which were often spectrally similar, also showed improved OA with Pixel-DMV. For example, Artificial Ground accuracy was 80.96% with Pixel-DMV compared to 76.83% with SVM in the Pleiades dataset. This improvement was likely due to Pixel-DMV's ability to leverage spatial context and neighboring pixel information, which helped in distinguishing between spectrally similar classes.

The dimensional voting mechanism of Pixel-DMV, combined with the Median filter preprocessing step, played a crucial role in its superior performance. The Median filter reduced noise and smoothed the data, enhancing the effectiveness of the

classification process. This was particularly beneficial for high-resolution datasets like Pleiades and SPOT 7, where noise could be more pronounced. The consistency of Pixel-DMV's performance across both datasets underscored its robustness and generalizability. The higher accuracy in the Pleiades dataset compared to SPOT 7 could be attributed to its superior spatial and spectral resolution, which allowed Pixel-DMV to better capture fine details and spectral variations.

For classes with an OA below 80%, such as Lake (72.41%) and Bare Earth (76.75%) for Pleiades, as well as Artificial Ground (65.43%) and Healthy Grass (74.07%) for SPOT 7, the lower performance could be attributed to the inherent complexity of these land cover types. These classes often exhibited spectral similarities with other classes, making it difficult for the algorithm to distinguish them accurately. If the surrounding pixels were misclassified or belonged to multiple classes, the voting process might have produced incorrect rankings. With further refinement, such as expanding the training sets to better represent heterogeneous or complex classes, Pixel-DMV could have achieved even greater accuracy for all land cover types.

4. CONCLUSION AND FUTURE WORK

The primary objectives of this study were to develop and evaluate the performance of Pixel-DMV for land use and land cover classification in mixed-land zones, in terms of OA, AA, and Kappa by using Pleiades and SPOT 7 datasets. Comparative analysis against traditional classification methods, including SVM, RF, and ML, highlighted the potential advantages of Pixel-DMV. The study affirmed the suitability of Pixel-DMV for applications requiring frequent and cost-effective classification, such as agricultural management, urban planning, and disaster response. The limitation of the study was that the accuracy of Pixel-DMV for all land cover types could have been further improved with additional refinements, such as expanding the training datasets to better represent heterogeneous or complex classes.

This work differed from those that existed in the literature in several key aspects: 1) unlike other studies that did not specifically focus on a particular type of area, this study considered a specific land zone (mixed-land zones); 2) while most studies relied on a single dataset, this study utilized multiple datasets for a more comprehensive analysis; and 3) instead of merely proposing a new method, this study selected the high-performing classifiers based on previous works, for a comparative study with the proposed method.

Future work based on this study includes incorporating additional datasets beyond Pleiades and SPOT 7 for a comparative analysis of their potential and limitations, as well as evaluating the performance of Pixel-DMV in urban areas, where different confusion factors may arise compared to its strong performance in coastal mixed-land zones. The classification technique needs to be further improved, especially for Lake and Bare Earth using Pleiades imagery as well as Artificial Ground and Healthy Grass using SPOT 7, with accuracies below 80%.

ACKNOWLEDGMENT AND FUNDING

This work was supported by the Universiti Teknologi Malaysia and GATES IT Solution Sdn Bhd (GATES) under a Contract Research DTD Grant (R.J130000.7628.4C652).

DECLARATION OF CONFLICTING INTERESTS

The authors declare no potential conflicts of interest with respect to the research and publication of this article.

REFERENCES

[1] J. He, X. Li, P. Liu and X. Wu, Accurate estimation of the proportion of mixed land use at the street-block level by integrating high spatial resolution images and geospatial big data, *IEEE Transactions on Geoscience and Remote Sensing*, 59(8), 2021, 6357-6370.

[2] R. Liu, H. Zhang and J. Ling, Hybrid transformer networks for urban land use classification from optical and SAR images, *International Geoscience and Remote Sensing Symposium (IGARSS 2022)*, Kuala Lumpur, Malaysia, 2022, 707-710.

[3] P. Pandey, Y. Dubey and S. V. Balamwar, Automatic feature abstraction from high resolution satellite data for urban land cover and land use, *8th International Conference on Advanced Computing and Communication Systems (ICACCS 2022),* Tamil Nadu, India, 2022, 912-917.

[4] R. M. Santiago, R. Gustilo, G. P. Arada, E. Magsino and E. Sybingco, Performance analysis of machine learning algorithms in generating urban land cover map of Quezon city, Philippines using sentinel-2 satellite imagery, *2021 IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM 2021)*, Manila, Philippines, 2021, 1-6.

[5] J. Garcia-Gutierrez, L. Gonalves-Seco and J. C. Riquelme-Santos, Automatic environmental quality assessment for mixed-land zones using lidar and intelligent techniques, *Expert Systems Applied*, 38(6), 2011, 6805-6813.

[6] Z. Li, H. Zhang, F. Lu, R. Xue, G. Yang and L. Zhang, Breaking the solution barrier: A low-to-high network for large-scale high-resolution land-cover mapping using low-resolution labels, *ISPRS Journal of Photogrammetry and Remote Sensing*, 192, 2022, 244-267.

[7] D. Ienco, R. Interdonato, R. Gaetano and D. Ho Tong Minh, Combining Sentinel-1 and Sentinel-2 satellite Image time series for land cover mapping via a multi-source deep learning architecture, *ISPRS Journal of Photogrammetry and Remote Sensing*, 158, 2019, 11-22.

[8] B. Adriano, J. Xia, G. Baier, N. Yokoya and S. Koshimura, Multi-source data fusion based on ensemble learning for rapid building damage mapping during the 2018 Sulawesi earthquake and Tsunami in Palu, Indonesia, *Remote Sensing*, 11(7), 2019, 886.

[9] O. Ghorbanzadeh, H. Shahabi, F. Mirchooli and K. V. Kamran, Gully erosion susceptibility mapping (GESM) using machine learning methods optimized by the multi-collinearity analysis and K-fold cross-validation, *Geomatics, Natural Hazards and Risk*, 11(1), 2020, 1653-1678.
[10] H. A. T. Nguyen, T. Sophea, S. H. Gheewala, R. Rattanakom, T. Areerob and K. Prueksakorn, Integrating remote sensing and machine learning into environmental monitoring and assessment of land use change, *Sustainable Production and Consumption*, 27, 2021,1239-1254.
[11] E. Paradis, Probabilistic unsupervised classification for large-scale analysis of spectral imaging data, *International Journal of Applied Earth Observation and Geoinformation*, 107, 2022, 102675,
[12] Z. Lv, T. Liu, C. Shi, J. A. Benediktsson and H. Du, Novel land cover change detection method based on k-Means clustering and adaptive majority voting using bitemporal remote sensing images, *IEEE Access*, 7, 2019, 34425-34437.
[13] M. Ahmed, R. Seraj and S. M. S. Islam, The k-means algorithm: A comprehensive survey and performance evaluation, *Electronics (Switzerland)*, 9(8), 2020, 1-12.
[14] M. F. Baig, M. R. U. Mustafa, H. B. Takaijudin and M. T. Zeshan, Comparative analysis of support vector machine and maximum likelihood classifications using satellite images of Selangor, Malaysia, *2021 3rd International Sustainability and Resilience Conference: Climate Change*, 2021, 405-409.
[15] S. Swetanisha, A. R. Panda and D. K. Behera, Change detection using machine learning models: A case study on the Puri District of Odisha, India, *2021 19th OITS International Conference on Information Technology (OCIT)*, 2022, 100-104.
[16] K. Patel, M. Jain, M. I. Patel and R. Gajjar, A novel approach for change detection analysis of land cover from multispectral FCC optical image using machine learning, *2nd International Conference on Range Technology, ICORT 2021*, 2021, 1-6.
[17] T. K. Das, D. K. Barik and K. V. G. R. Kumar, Land-use land-cover prediction from satellite images using machine learning techniques, *2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON 2022)*, 1, 2022, 338-343.
[18] J. Xia, N. Yokoya and A. Iwasaki, Fusion of hyperspectral and LiDAR data with a novel ensemble classifier, *IEEE Geoscience and Remote Sensing Letters*, 15(6), 2018, 957-961.
[19] N. D. Ponnaganti and R. Anitha, A Novel Ensemble Bagging classification method for breast cancer classification using machine learning techniques, *Traitement du Signal*, 39(1), 2022, 229-237.
[20] H. Ebrahimy and Z. Zhang, Per-pixel accuracy as a weighting criterion for combining ensemble of extreme learning machine classifiers for satellite image classification, *International Journal of Applied Earth Observation and Geoinformation*, 122, 2023, 103390.
[21] M. El Amine Douad, N. Aribi, S. Loudni and Y. Lebbah, Multicriteria ensemble learning for multispectral image classification, *2023 International Conference on Earth Observation and Geo-Spatial Information (ICEOGI)*, 2023, 1-6.
[22] J. Adrian, V. Sagan, and M. Maimaitijiang, Sentinel SAR-optical fusion for crop type mapping using deep learning and Google earth engine, *ISPRS Journal of Photogrammetry and Remote Sensing*, 175, 2021, 215-235.
[23] R. D. D. Altarez, A. Apan and T. Maraseni, Deep learning U-Net classification of Sentinel-1 and 2 fusions effectively demarcates tropical montane forest's deforestation, *Remote Sensing Applied*, 29, 2023, 100887.
[24] S. Gupta, K. Saluja, A. Goyal, A. Vajpayee and V. Tiwari, Comparing the performance of machine learning algorithms using estimated accuracy, *Measurement: Sensors*, 24, 2022, 100432.
[25] V. N. Mishra, R. Prasad, P. K. Rai, A. K. Vishwakarma, and A. Arora, Performance evaluation of textural features in improving land use/land cover classification accuracy of heterogeneous landscape using multi-sensor remote sensing data, *Earth Science Informatics*, 12(1), 2019, 71-86.
[26] R. Costache, Q. B. Pham, E. Corodescu-Rosca, C. Cimpianu, H. Hong, N.T. Thuy Linh, C.M. Fai, A. N. Ahmed, M. Vojtek, S.M. Pandhiani, G. Minea, N. Ciabotaru, M. C. Popa, D. C. Dianconu and B. T. Pham, Using GIS, remote sensing, and machine learning to highlight the correlation between the land-use/land-cover changes and flash-flood potential, *Remote Sensing* 12 (9), 2020, 1422.
[27] T. G. Whiteside, G. S. Boggs and S. W. Maier, Comparing object-based and pixel-based classification for mapping savannas, *International Journal of Applied Earth Observation and Geoinformation,* 13(6), 2011, 884-893.
[28] D. Gupta, R. Rathi and S. Gupta, An efficient approach of filtering for noises on images, *2018 4th International Conference on Computing Communication and Automation, ICCCA 2018*, 2018, 1-5.
[29] M. H. Ishak, N. N. Sofia, M. Marzuki, M. F. Abdullah, Z. H. Che Soh, I. S. Isa and S. N. Sulaiman, Image quality assessment for image segmentation algorithms: Qualitative and quantitative analyses, *Proceedings - 9th IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2019*), 2019, 66-71.
[30] C. S. K. Abdulah, M. N. K. H. Rohani, B. Ismail, M. A. M. Isa, A. S. Rosmi and W. A. Mustafa, Comparison of image restoration using median, Wiener, and Gaussian filtering techniques based on electrical Tree, *IEACon 2021 - 2021 IEEE Industrial Electronics and Applications Conference*, pp. 2021, 163-168.
[31] B. B. Ahamed, D. Yuvaraj and S. S. Priya, Image denoising with linear and non-linear filters, *Proceedings of 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE 2019*), 2019, 806-810.
[32] H. Houtsitjoki and M. Juhola, Comparing the one-vs-one and one-vs-all methods in benthic macroinvertebrate image classification, *Proceedings of the 7th International Conference on Machine Learning and Data Mining in Pattern Recognition,* 2011, 399-413.
[33] Sarawak Government Official Portal. http://www.ictu.tmp.sarawak.gov.my/seg.php?recordID=M0001&mainmenuID=M0001 (accessed Mar. 7, 2025).

[34] H. Ouchra, A. Belangour and A. Erraissi, Machine learning algorithms for satellite image classification using Google earth engine and Landsat satellite data: Morocco case study, *IEEE Access,* 11, 2023, 71127-71142.
[35] S. Pande and B. Banerjee, HyperLoopNet: Hyperspectral image classification using multiscale self-looping convolutional networks, *ISPRS Journal of Photogrammetry and Remote Sensing*, 183, 2022, 422-438.
[36] P. Ahmadi, S. B. Mansor, H. Ahmadzadeh Araji and B. Lu, Convolutional SVM networks for detection of *Ganoderma boninense* at early stage in oil palm using UAV and multispectral Pleiades images, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, X-4/W1-2022*, 2023, 25-30.
[37] M. Hamimeche, S. Niculescu, A. Billey and R. Moulai, Identification and mapping of Algerian island vegetation using high-resolution images (Pléiades and SPOT 6/7) and random forest modeling, *Environmental Monitoring and Assessment,* 193(9), 2021, 617.
[38] X. Li, L. Yu, Q. Zhou, D. Wu, L. Ren and Y. Luo, Detection of tree species in Beijing plain afforestation project using satellite sensors and machine learning algorithms. *Forests*, 14(9), 2023, 1889.
[39] Zylshal, S. Sulma, F. Yulianto, J. T. Nugroho and P. Sofan, A support vector machine object-based image analysis approach on urban green space extraction using Pleiades-1A imagery, *Modelling Earth Systems and Environment*, 2, 2016, 1-12.
[40] P. Widayani, A. Fadilah, I. Z. Irawan and K. Ghosh, Implementing support vector machine algorithm for early slum identification in Yogyakarta City, Indonesia using pleiades images, *Forum Geografi,* 37(1), 2023, 88-97.
[41] N. Rahimizadeh, S. Babaie Kafaky, M. R. Sahebi and A. Mataji, Forest structure parameter extraction using SPOT-7 satellite data by object- and pixel-based classification methods, *Environmental Monitoring and Assessment,* 192(1), 2020, 43.
[42] W.T. Ng, P. Rima, K. Einzmann, M. Immitzer, C. Atzberger and S. Eckert, Assessing the potential of Sentinel-2 and Pleiades data for the detection of *Prosopis* and *Vachellia* spp. in Kenya, *Remote Sensing,* 9(1), 2017, 74.