Proceedings of ISERME 2025
19-20 September 2025
https://doi.org/10.31705/ISERME.2025.9

Evaluation of Image-Based Land Use Classification with Multi-Feature Fusion Approach

Wijethunga MGSP, Senaratne UAADT, Sivasubramaniam S, Gunathilaka JPDAR, Thiruchittampalam S and *Chaminda SP

Department of Earth Resources Engineering, University of Moratuwa, Sri Lanka

*Corresponding author – chaminda@uom.lk

Abstract

Traditional land use and land cover (LULC) classification approaches using spectral bands display persistent limitations in distinguishing similar landcover types, particularly within heterogeneous urban–rural transition zones. This research examines the effectiveness of multi-feature fusion for improving classification accuracy by systematically evaluating combinations of spectral indices with machine learning classifiers. Sentinel-2 imagery of Kaduwela, Sri Lanka, was utilized, focusing on six LULC categories: water, forest, vegetation, roads, buildings, and rock exposure. The methodology involved an initial evaluation of baseline RGB performance using Support Vector Machine (SVM), Random Forest (RF), and Decision Tree (DT) classifiers. Subsequently, 27 spectral indices were generated and integrated with RGB to form a 30-dimensional feature set, followed by Principal Component Analysis (PCA) to reduce dimensionality, retaining five components that explained over 95% of the variance for subsequent classification. The RGB feature set yielded accuracies of 39.22% (SVM), 44.19% (DT), and 50.45% (RF); integration of RGB with 27 indices improved accuracies to 53.22% (SVM), 52.01% (DT), and 52.94% (RF); while the PCA-reduced feature set provided 54.70% (SVM), 46.18% (DT), and 49.34% (RF). The findings highlight that in heterogeneous and urban–rural interfaces, PCA-based reduction of spectral indices has improved SVM classification and reduced computational load. For future studies, further hyperparameter tuning could enhance accuracy.

Keywords: Land use classification, Machine learning, Multi-feature fusion, Remote sensing, Spectral Indices

1 Introduction

Land Use Land Cover (LULC) classification is essential for generating accurate land use maps, which are vital for environmental monitoring, urban planning, and resource management. Satellite remote sensing provides timely and consistent LULC data, including multispectral and spatial information. Despite technological advances, key challenges persist. Notably, the "mixed pixel" issue and "spectral data confusion" remain unresolved [12]. Additionally, distinguishing land use from land cover continues to be difficult, as spectrally distinct land functions often exhibit similar reflectance across wavelengths [3]. This issue is especially pronounced in complex environments such as the urban rural interface, where rapid transitions between residential zones, farmlands, forests, waterbodies, and mixed developments create spectral ambiguity [1].

To tackle this complexity, high-resolution satellite data across standard bands (Red, Green, Blue, SWIR, and Red Edge) are essential. However, raw spectral bands alone may fail to capture subtle heterogeneity. Derived features such as spectral indices (e.g., NDVI, NDBI, MNDWI) are often needed to improve class separability [2], [6]. One study in Banda Aceh City [6] compared Maximum Likelihood classification using only spectral bands versus including spectral indices. The latter improved overall accuracy from 83.56% to 92.22%, underscoring the importance of index-enhanced features for classification.

While spectral indices aid in separating classes, further accuracy improvements often require advanced analytical approaches. Machine Learning (ML) has demonstrated strong capabilities in handling complex spectral and textural patterns. For instance, [14] applied ML to relate spectral and textural features in geotechnical contexts, and [2] showed accuracy gains through textural feature integration. Similarly, [10] assessed four ML algorithms Random Forest (RF), Support Vector Machine (SVM), Maximum Likelihood (ML), and K Nearest Neighbors (KNN) for LULC classification using Sentinel 2 imagery. Incorporating NDVI, NDBI, and NDWI led to improved outcomes, with RF achieving the highest accuracy (93.1%).

Building upon these results, [4] employed a post classification multi feature fusion approach, combining outputs from RF, SVM, and Decision Trees to create an ensemble map. This strategy outperformed individual classifiers, further enhancing classification accuracy.

Beyond algorithm choice, the relevance of input features is crucial. Feature extraction (e.g., spectral bands, vegetation indices, texture) and feature selection significantly affect model performance. A study on crop mapping using Sentinel 2 time series data [8] found that Random Forest based feature selection combined with Extreme Gradient Boosting yielded over 90% accuracy, outperforming RF and SVM. This highlights the importance of selecting informative variables.

Dimensionality reduction also plays a key role. In [1], Principal Component Analysis (PCA) was applied to Landsat 8 data for LULC mapping in Egypt's North Delta. PCA reduced seven spectral bands to three principal components retaining over 98% of the variance. Among tested classifiers (SVM, KNN, RF), SVM applied to the PCA reduced dataset achieved the highest accuracy (96.97%), demonstrating the benefit of optimized spectral inputs.

The findings of [15] also underscore the importance of structured training and validation. Their supervised approach used ground truth samples from field data and high-resolution maps, employing a supervised train validation split. This reinforces the value of integrating diverse, optimized features with robust classification strategies to enhance LULC mapping outcomes.

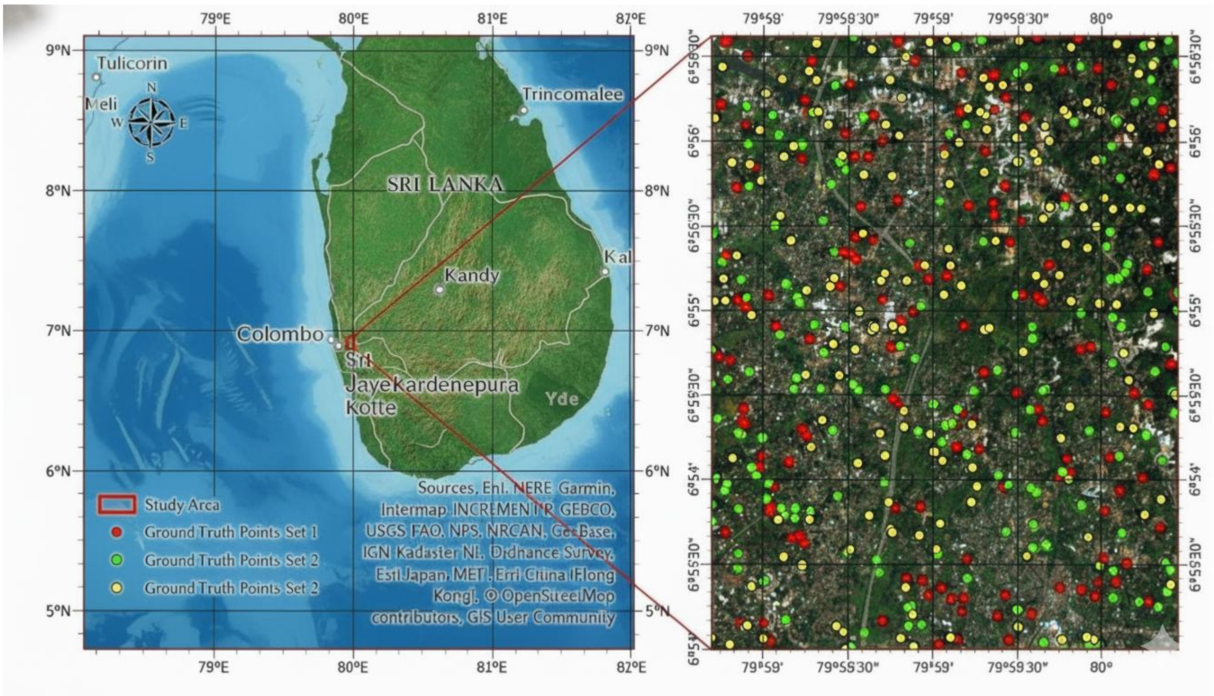

Figure 1: Study area and Ground truth point distribution covering the classes (Water, Forest, Vegetation, Roads, Buildings, Rock Exposure) - Esri Topographic map – Sentinel-2 (R10m).

2 Methodology

2.1 Study Area

The study area selected for this research encompasses 63.28 km^2 in Kaduwela, Sri Lanka, representing a complex urban-rural boundary region that provides presence of quarry site contribute to a high degree of class variability, making it an ideal environment for evaluating land use classification performance (Figure 1). This extensive area exhibits significant landscape heterogeneity, containing diverse land cover types including built-up areas, rock outcrops with exposed surfaces, water bodies, road networks, vegetation patches, and agricultural features. The selection of this region was made to assess classification algorithms across spectrally similar and heterogeneous classes, which commonly pose significant challenges in remote sensing applications [2].

This condition creates a complex mosaic of land cover types that exhibit high spectral variability and spatial heterogeneity, making it particularly suitable for testing multi-feature fusion approaches in land use classification [3], [4].

2.2 Material and Methods

2.2.1 Methodological Overview

The methodology proceeds with the feature extraction (generation of composite images, including basic RGB, NDVI, NDWI, and multiple other spectral indices), which the extracted features are then combined in various configurations. These composite datasets serve as inputs for training machine learning models with different algorithms.

Feature selection techniques are applied to identify the most effective spectral index combinations, which are then used for final classification and accuracy assessment. The workflow culminates in a comparative analysis of classification accuracies, leading to the identification of the optimal set of spectral features for land use classification in complex urban-rural environment (Figure 2).

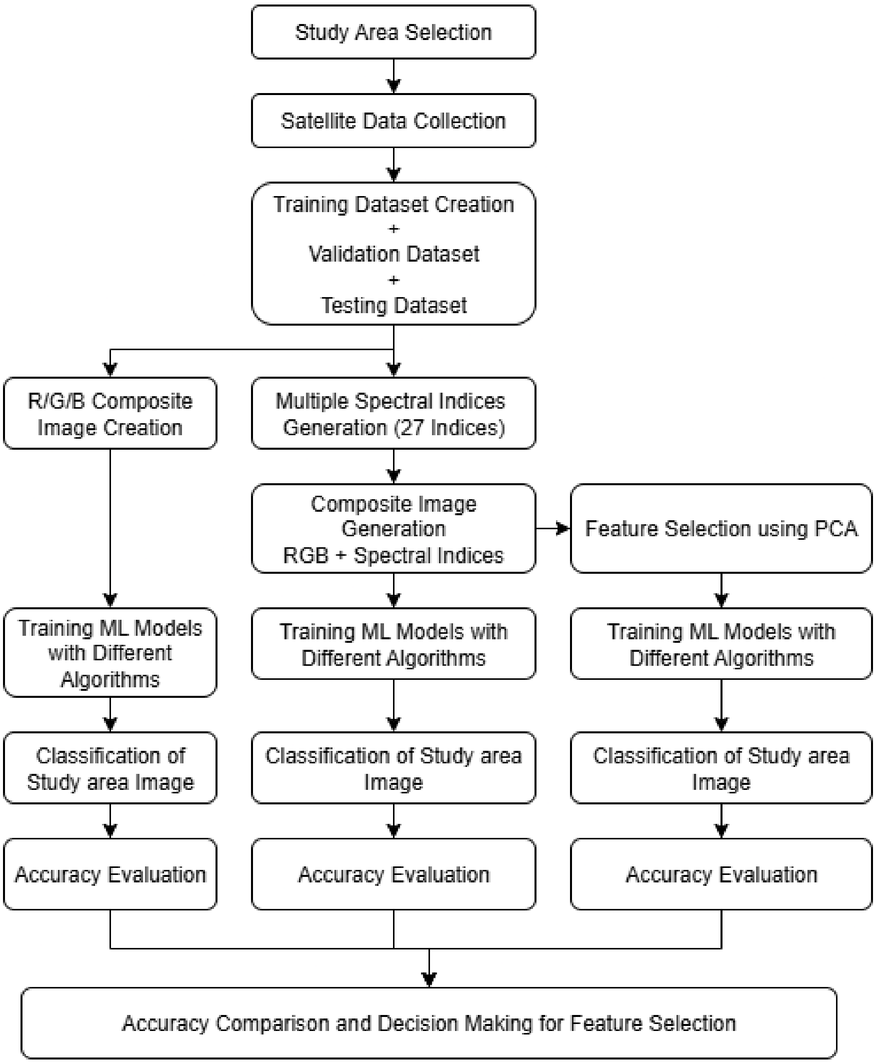

Figure 2: Workflow of Multi-Feature Fusion and PCA-Based Land Use Classification Using Machine learning.

2.2.2 Satellite data

The study utilized Sentinel-2C level-2A imagery acquired on February 8, 2025, from the Copernicus Open Access Hub. The data provides atmospherically corrected bottom-of-atmosphere reflectance values, essential for accurate land cover classification [12]. The Multi-Spectral Instrument (MSI) aboard Sentinel-2C offers 13 spectral bands with spatial resolutions of 10m, 20m, and 60m, where RGB and NIR bands at 10-meter resolution and SWIR bands at 20-meter resolution were specifically selected for this analysis. The Level-2A processing eliminates the need for additional atmospheric correction as the data has been radiometrically corrected to surface reflectance values with a quantification factor of 10,000 [5]. Training and validation dataset.

Two independent, randomly selected ground truth data sets for training (90% of ground truth polygons) and validation (10% of ground truth polygons) were acquired through visual interpretation and manual digitization of high-resolution imagery available through Google Earth. Google Earth provides access to very high-resolution satellite imagery with spatial resolutions typically ranging from 0.5 to 1 meter for most urban and semi-urban areas, which is adequate for accurate identification and delineation of land cover features in heterogeneous landscapes. The platform's integration of multi-temporal imagery from various commercial satellite providers. Ensures access to cloud-free, high-quality imagery suitable for ground truth data collection in remote sensing applications [5].

Six distinct land cover classes were identified and manually digitized as polygon features with corresponding attribute data: Water, Forest, Vegetation, Roads, Buildings, and Rock Exposure. This classification scheme was designed to capture the major land cover types present within the urban-rural boundary environment, representing both natural and anthropogenic features that exhibit distinct spectral characteristics in satellite imagery [4], [6]. The digitization process involved careful visual interpretation of landscape features, with particular attention to maintaining consistent class definitions and avoiding mixed-pixel effects at class boundaries.

Cumulative areal calculations were made to balance the areal coverage of each land cover class to minimize potential bias in classifier training, though complete balance was constrained by the natural distribution of land cover types within the study area [2], [7] (Table 1). An independent validation dataset was generated using stratified random sampling to ensure proportional representation of each land cover class. Model performance was evaluated using a confusion matrix, from which overall Accuracy, Producer's Accuracy, and the Kappa Coefficient were derived (Table 2).

Table 1: Count and area of polygons in training dataset

Class	Count	Area [m^2]
Water	158	201388.165
Forest	276	2400321.551
Vegetation	47	236256.883
Roads	98	25538.335
Buildings	192	166664.106
Rock Exposure	164	172697.943

Table 2: Count and area of polygons in validation dataset

Class	Count	Area [m^2]
Water	20	66699.362
Forest	20	35170.606
Vegetation	20	46659.963
Roads	12	1718.836
Buildings	19	15442.768
Rock Exposure	20	17777.897

2.2.3 Testing dataset

Three independent datasets were generated through stratified random sampling across the study area to assess the classification accuracy. The first dataset contained 300 randomly distributed points, while the second and third datasets each comprised 400 points, total 1100 testing points (Figure 1). Each point was manually verified against high-resolution Google Earth imagery (0.5–1 m spatial resolution) to confirm its land cover class, following the same ground truthing protocol. used for polygon-based training and validation datasets. The multi-dataset approach enables evaluation of classification consistency across different sampling densities and spatial distributions

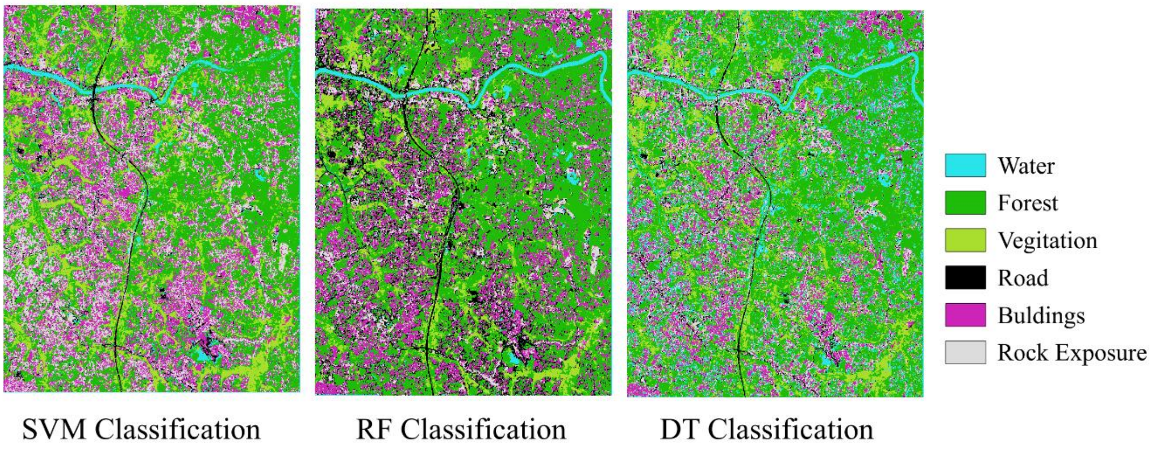

Figure 3: Comparison of SVM, RF and DT classifiers for Land Cover Classification using 11 Spectral Indices.

2.3 Feature extraction

To improve classification performance, a multi layered feature extraction approach was adopted. Spectral indices NDVI (vegetation), NDBI (Buitl-up areas) and MNDWI (water bodies) were computed to enhance class separability. These features were combined with the original spectral bands to generate excessive indices. In that manner 27 indices were generated. Generated indices were again combined with RGB and a single composite was created.

2.4 Feature selection

In Feature selection process, principal component analysis (PCA) was applied to reduce dimensionality and redundancy within the high-dimensional Sentinel-2 dataset. PCA was selected over other methods due to its ability to transform correlated spectral features into a compact set of uncorrelated components, thereby preserving most information content while minimizing noise and computational complexity. This made it particularly suitable for enhancing classification performance in the subsequent modeling process. The methodology involved transforming the original feature space, which contained multiple spectral indices and derived features, into a reduced dimensional representation while preserving derived features, into a reduced dimensional representation while preserving the most significant variance in the data.

The PCA implementation involved standardizing the input feature matrix, computing the covariance matrix, and performing eigenvalue decomposition to extract principal components ordered by explained variance criterion, retaining components until a predetermined threshold of total variance was achieved.

The dataset for Feature Selection was A composite image was created containing 27 standard spectral indices and RGB. The order **[RGB, AWEI, VARI, NDLI, NDVI705, OSAVI, MSAVI, NDRE, MCARI, TVI, NDBI, NDWI, NDVI, RGB, TNDVI, PSSR, NIRI, GCI, BAI, EVI2, ARVI, NDSI, NCSI, SIPI, WI, VSDI, NDTI, UI, MSI].**

To reduce the high dimensionality of the 27-index and RGB composite image, Principal Component Analysis (PCA) was applied. PCA transformed the correlated spectral indices into a new set of uncorrelated components that capture the maximum variance in the data.

The resulting eigenvalues from the PCA indicated that the first few components accounted for the majority of variance in the dataset. A scree plot and cumulative variance analysis were used to determine the optimal number of components to retain. Based on this, the top 10 principal components capturing over 95% of the total variance and that was selected for the classification. The selected components were then used as input features for the SVM classifier.

While PCA improved processing efficiency and reduced overfitting risk, its unsupervised nature means it does not explicitly prioritize class-discriminative features. This limitation may have contributed to suboptimal classification accuracy, indicating that integrating PCA with supervised feature selection methods could improve the performance in future studies.

2.5 Image Classification

The selection of DT, SVM and RF classifiers was informed by prior studies that demonstrated their effectiveness in LULC classification tasks [2,4,10]. DT offers a simple, interpretable baseline, SVM handles complex, non-linear boundaries effectively and RF as an ensembled method, improves accuracy and reduces overfitting. This combination allows for a balanced comparison across a spectrum of commonly used machine learning approaches. Separate models were trained using each algorithm for every spectral composite. The classification process involved training each model on the labeled training dataset, followed by applying the trained model to the test data to produce a classified study area map. This approach allowed for a comparative Evaluation of how different classifiers performed with varying feature sets. Accuracy metrices, including overall accuracy and confusion matrices, were computed to assess the effectiveness of each model and to identify the most suitable algorithm-feature combination for the study.

3 Results and Discussion

This study evaluated how different feature combinations influence land use classification accuracy using Decision Tree (DT), Support Vector Machine (SVM), and Random Forest (RF) classifiers. Performance was assessed using Overall Accuracy and Kappa Index, averaged across two independent validation datasets to ensure robustness and reduce sampling bias.

Using only RGB bands resulted in the lowest performance. RF achieved an average accuracy of **(50.45%)**, followed by DT **(44.19%)** and SVM **(41.8%)**. The corresponding Kappa values were (0.3580), (0.2800), and (0.2196), respectively, underscoring the limitations of basic spectral input in complex and heterogeneous landscapes. (Table 1, Figure 4, Figure 5).

Table 3: RGB classification accuracy

Model	**Accuracy (%)**	**Kappa Index**
SVM	39.22	0.2197
Decision Tree	44.19	0.2801
Random Forest	50.45	0.3580

The next approach, adding 27 indices and RGB composite enhanced the classification accuracy in all 3 classifiers DT, RF, SVM (Table 4. Figure 4, Figure 5). Kappa indices have also been enhanced, claiming number of features that are added elevate the accuracy of the classification.

Table 4: RGB + 27 indices accuracy (%) classification accuracy

Model	**Accuracy (%)**	**Kappa Index**
SVM	53.22	0.3598
Decision Tree	52.01	0.3347
Random Forest	52.94	0.3487

To manage feature dimensionality and redundancy, Principal Component Analysis (PCA) was applied, preserving the most informative variance components. This step elaborated acceptable performance with SVM classifier demarcating a **54.79%** accuracy level. In the context of imagery (sentinel - 2), This value is acceptable while this can be improved by moving into a high resolution (<10m) imagery. Additionally performing a classifier hyper parameter optimization can improve classification accuracy further.

Table 5: Accuracy using PCA (RGB + 27 Indices)

Model	**Accuracy (%)**	**Kappa Index**
SVM	54.79	0.3702
Decision Tree	46.18	0.3598
Random Forest	49.34	0.3107

Adding 27 indices to RGB improves accuracy for all models. SVM gains the most with PCA, reaching the highest accuracy. RF and DT perform better with indices, but PCA slightly reduces their accuracy.

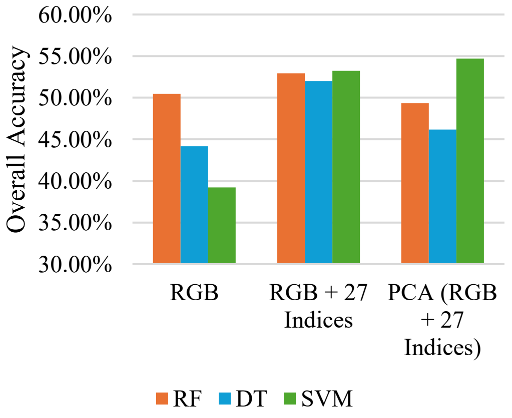

Figure 4: Overall accuracy comparison of RGB, RGB + 27 Indices, and PCA of RGB + 27 Indices against RF, DT, and SVM

Kappa values improve with more features. SVM shows the largest jump, especially with PCA. RF remains steady across all setups. DT shows the lowest performance throughout.

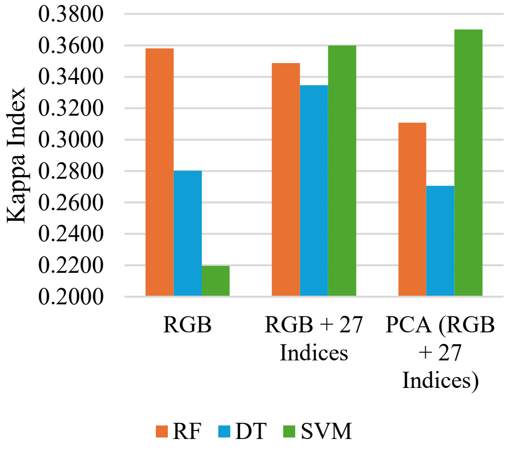

Figure 5: Kappa index comparison of RGB, RGB + 27 Indices, and PCA of RGB + 27 Indices against RF, DT, and SVM

Among all scenarios, Random Forest consistently yielded the highest overall accuracy, reflecting its strength in handling high-dimensional and nonlinear data. SVM showed the greatest benefit from PCA, indicating its sensitivity to feature scaling and structure.

Further accuracy gains could be achieved through temporal satellite data; hence the various environmental conditions directly impact on the Accuracy.

In conclusion, the combined use of spectral and textural features with dimensionality reduction significantly improves classification accuracy. When paired with robust classifiers like Random Forest, this integrated approach is well suited for accurately mapping heterogeneous land cover environments.

4 Conclusions

This research showed that multi-feature fusion can significantly improve land use classification accuracy compared to using RGB imagery alone. Integrating indices like NDVI, NDBI and NDWI contributed to better class separability, while adding many indices without proper selection resulted in mixed performance. Applying PCA and selecting only the first 10 principal components based on variance explained helped reduce dimensionality and improved accuracy, especially for SVM. Among the classifiers, Random Forest consistently achieved higher accuracy with multiple features, confirming its effectiveness for handling diverse land cover data. Overall, this study demonstrates that careful selection of features and dimensionality reduction are key to improving classification accuracy for complex landscapes, supporting better land use mapping and planning.

References

[1] J. Zhang, X. Yuan, X. Tan, and X. Zhang, “Delineation of the Urban-Rural Boundary through Data Fusion: Applications to Improve Urban and Rural Environments and Promote Intensive and Healthy Urban Development,” *Int. J. Environ. Res. Public. Health*, vol. 18, no. 13, p. 7180, Jul. 2021, doi: 10.3390/ijerph18137180.

[2] 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 Sci. Inform.*, vol. 12, no. 1, pp. 71–86, Mar. 2019, doi: 10.1007/s12145-018-0369-z.

[3] C. Zhang *et al.*, “An object-based convolutional neural network (OCNN) for urban land use classification,” *Remote Sens. Environ.*, vol. 216, pp. 57–70, Oct. 2018, doi: 10.1016/j.rse.2018.06.034.

[4] Y. O. Ouma, A. Keitsile, B. Nkwae, P. Odirile, D. Moalafhi, and J. Qi, “Urban land-use classification using machine learning classifiers: comparative evaluation and post-classification multi-feature fusion approach,” *Eur. J. Remote Sens.*, vol. 56, no. 1, p. 2173659, Dec. 2023, doi: 10.1080/22797254.2023.2173659.

[5] C. Zhang *et al.*, “Joint Deep Learning for land cover and land use classification,” *Remote Sens. Environ.*, vol. 221, pp. 173–187, Feb. 2019, doi: 10.1016/j.rse.2018.11.014.

[6] Mulkal, P. Rizkya, N. Aflah, and Muchlis, “Exploring the role of Spectral Indices on improving Land Cover Classification accuracy based on Sentinel-2 satellite imagery In Banda Aceh City, Indonesia,” *Indones. J. Environ. Sustain.*, vol. 2, no. 2, pp. 96–107, Dec. 2024, doi: 10.22373/ijes.v2i2.5957.

[7] E. A. L. Salas and S. S. Kumaran, “Hyperspectral Bare Soil Index (HBSI): Mapping Soil Using an Ensemble of Spectral Indices in Machine Learning Environment,” *Land*, vol. 12, no. 7, p. 1375, Jul. 2023, doi: 10.3390/land12071375.

[8] R. Tufail, P. Tassinari, and D. Torreggiani, “Assessing feature extraction, selection, and classification combinations for crop mapping using Sentinel-2 time series: A case study in northern Italy,” *Remote Sens. Appl. Soc. Environ.*, vol. 38, p. 101525, Apr. 2025, doi: 10.1016/j.rsase.2025.101525.

[9] A. Humeau-Heurtier, “Texture Feature Extraction Methods: A Survey,” *IEEE Access*, vol. 7, pp. 8975–9000, 2019, doi: 10.1109/ACCESS.2018.2890743.

[10] L. K. Jasim, R. H. Hasan, and O. A. Ibrahim, “Evaluation of the Accuracy of Machine Learning Classifiers and Spectral Indices in Land Cover Classification,” *Eng. Technol. Appl. Sci. Res.*, vol. 15, no. 3, pp. 22548–22552, Jun. 2025, doi: 10.48084/etasr.10406.

[11] X. Wei, W. Zhang, Z. Zhang, H. Huang, and L. Meng, “Urban land use land cover classification based on GF-6 satellite imagery and multi-feature optimization,” *Geocarto Int.*, vol. 38, no. 1, p. 2236579, Dec. 2023, doi: 10.1080/10106049.2023.2236579.

[12] D. Phiri and J. Morgenroth, “Developments in Landsat Land Cover Classification Methods: A Review,” *Remote Sens.*, vol. 9, no. 9, p. 967, Sep. 2017, doi: 10.3390/rs9090967.

[13] Z. Wang, J. Bai, and R. Feng, “A Multi-Feature Fusion Method for Urban Functional Regions Identification: A Case Study of Xi’an, China,” *ISPRS Int. J. Geo-Inf.*, vol. 13, no. 5, p. 156, May 2024, doi: 10.3390/ijgi13050156.

[14] S. Thiruchittampalam, B. P. Banerjee, N. F. Glenn, and S. Raval, “Geotechnical characterisation of coal spoil piles using high-resolution optical and multispectral data: A machine learning approach,” *Eng. Geol.*, vol. 329, p. 107406, Feb. 2024, doi: 10.1016/j.enggeo.2024.107406.

[15] S. Talukdar *et al.*, “Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review,” *Remote Sens.*, vol. 12, no. 7, p. 1135, Apr. 2020, doi: 10.3390/rs1207113