Comparative Analysis of Machine Learning Models for Sentinel-2 based Classification of the Bornean Heath Forest

Dwi Ahmad Dzulhijjah
Magister of Informatics, Universitas AMIKOM Yogyakarta, Indonesia | Department of Informatics and Computer Engineering, Politeknik Elektronika Negeri Surabaya, Indonesia
dwiahmadz@students.amikom.ac.id

Kusrini
Informatics Postgraduate Program, Universitas AMIKOM Yogyakarta, Indonesia | Fundacion para la Investigacion y Desarrollo Tecnologico de la Sociedad del Conocimiento, Murcia, Spain
kusrini@amikom.ac.id (corresponding author)

Rodrigo Martinez-Bejar
Informatics, Faculty of Computer Science, University of Murcia, Murcia, Spain | Fundacion para la Investigacion y Desarrollo Tecnologico de la Sociedad del Conocimiento, Murcia, Spain
rodrigo@um.es

Received: 11 January 2025 | Revised: 13 February 2025 | Accepted: 24 February 2025

 DOI: https://doi.org/10.48084/etasr.10173

ABSTRACT

Bornean heath forests, known as hutan kerangas, are fragile ecosystems that face significant anthropogenic threats. This study integrates Sentinel-2 satellite imagery with Machine Learning (ML) models to accurately classify these forests and assess their current spatial distribution. The Random Forest (RF) and Gradient Tree Boost (GTB) models achieved the highest classification performance, with overall accuracy scores of 96.66% and 96.69%, respectively, and Kappa coefficients of 0.945. These metrics were obtained using a test dataset with an 80:20 train-test split and validated through a 5-fold cross-validation process, ensuring the robustness of the models. Compared to previous studies employing unsupervised classification with Landsat-9 data, this approach demonstrates improved classification reliability and spatial accuracy. The findings highlight the substantial potential of combining remote sensing technologies with advanced ML techniques for large-scale ecosystem monitoring. This approach provides valuable insights for conservation planning and sustainable management of Bornean heath forests, addressing the growing environmental pressures that threaten their integrity.

Keywords-heath forest classification; Sentinel-2 imagery; machine learning models; remote sensing; climate

I. INTRODUCTION

Bornean heath forests, or hutan kerangas [1, 2], are unique ecosystems that thrive in nutrient-poor, acidic, and quartz-sand soils [3]. Despite their ecological roles in carbon sequestration, biodiversity conservation, and hydrological regulation, they face severe threats from deforestation, uncontrolled fires, sand mining, illegal logging, and land-use conversion [4]. Between 2009 and 2024, approximately 32,895,682.96 m^2 of heath forests have been lost, exacerbating climate challenges. These forests support endemic plant species such as Nepenthes gracilis and Rhodomyrtus tomentosa [5, 6], regulate hydrological cycles, and provide economic resources for local communities. However, their fragmented distribution and vulnerability to degradation present challenges for effective monitoring and conservation efforts.

Traditional monitoring methods, such as manual surveys and unsupervised classification with remote sensing, often fail to capture dynamic changes [7]. Integration of remote sensing platforms such as Landsat and Sentinel-2 with Geographic Information Systems (GIS) has improved large-scale forest monitoring [8, 9]. Previous studies in East Kalimantan successfully mapped heath forests using Landsat-9 and unsupervised classification [1], but the approach was limited in spatial resolution, spectral detail, and classification accuracy. Accurate Land Use and Land Cover (LULC) classification remains a research priority [10], particularly for heterogeneous ecosystems such as heath forests [11].

Machine learning (ML) models, such as Random Forest (RF), Support Vector Machines (SVM), Gradient Tree Boost (GTB), and K-Nearest Neighbors (KNN), have shown superior performance in handling high-dimensional remote sensing data [12, 13]. Applied to Sentinel-2 imagery, these models capture vegetation structure, canopy density, and soil characteristics effectively. Sentinel-2, with its finer spatial (10-20 m) and spectral resolution, offers notable advantages over Landsat for detailed ecosystem classification [14]. Multi-temporal remote sensing has further improved LULC classification by capturing seasonal and phenological variations, enhancing classification accuracy beyond single-date imagery approaches [15-17].

Research in Indonesia has focused on heath forest distribution, especially around the new capital city, Nusantara. A study employing Landsat-9 and GIS-based methods mapped these forests using parameters such as elevation, soil pH, and NDVI [1], aligning with previous findings on the ecological uniqueness of these forests in Southeast Asia [18]. Sentinel-2's higher spatial and spectral resolution enables the detection of subtle vegetation and soil variations essential for distinguishing heath forest types [19], while its improved temporal resolution enhances ecosystem monitoring.

This study refines Bornean heath forest classification by integrating Sentinel-2 data with ML techniques. The objectives are:

- Evaluate the classification performance of ML models, focusing on their ability to capture the spatial and spectral complexity of heath forest ecosystems using Sentinel-2 data. These models are compared with a baseline derived from a previous study [1] employing unsupervised classification with Landsat-9 data, demonstrating improvements in classification accuracy and spatial detail.
- Demonstrate the effectiveness of Sentinel-2 imagery in capturing finer spatial and spectral details of heath forest ecosystems compared to conventional approaches.
- Provide a robust and replicable method to monitor changes in heath forests using ML and high-resolution remote sensing data.

This research enhances forest classification precision, achieving more than 94% accuracy across models. Comparative evaluation offers insights into the best-performing models for classifying heterogeneous ecosystems, supporting scalable monitoring systems, conservation strategies, and sustainable forest management.

II. METHODOLOGY

A. Research Framework

This study adopts a structured framework for the classification of Bornean heath forests using Sentinel-2 satellite imagery and ML techniques. The framework consists of five stages: on-site observation, data collection and annotation, data preprocessing, model training and validation, and performance evaluation. This iterative process integrates remote sensing data with ML models for classifying heath forest regions. The workflow ensures accurate data representation and model performance validation. Figure 1 illustrates the overall method.

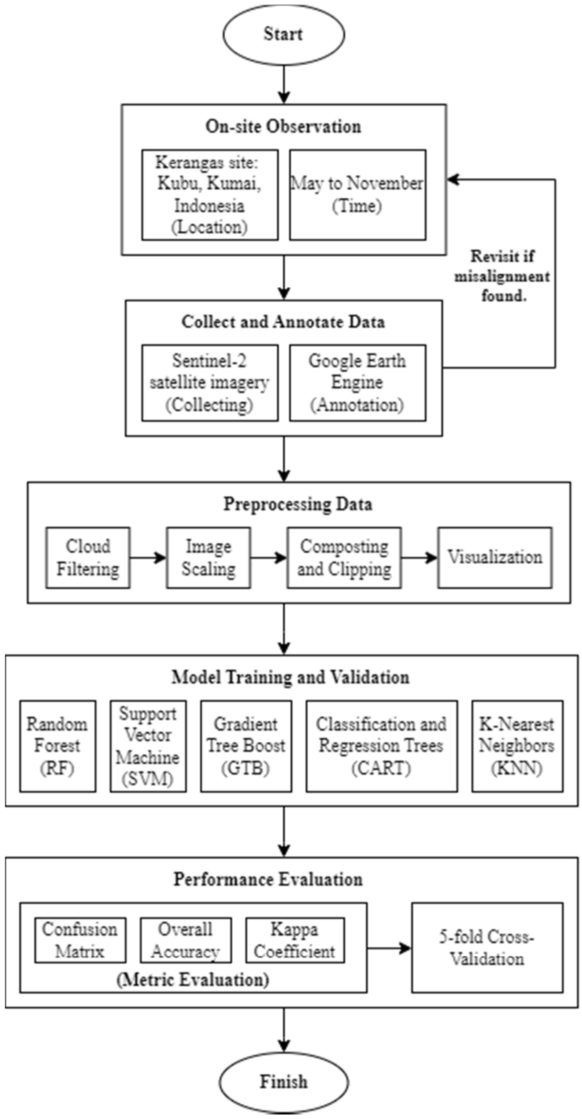

Fig. 1. Workflow of the research framework for heath forest classification.

B. On-site Observation

On-site observations were conducted to validate the alignment between Sentinel-2 imagery and actual ground conditions across 30 locations in Kubu, Kumai, and Kotawaringin Barat Regency. These observations, carried out over 12 sessions between May and November, served as essential references for ensuring accurate classification and improving training data quality. The field observations highlighted the distinctive ecological traits of heath forests (kerangas) across three main classes:

1. Kerangas Kolam: Waterlogged heath forest areas with sparse vegetation and lower canopy density.
2. Kerangas Hutan: Densely vegetated heath forest with well-defined tree canopies and high biodiversity.
3. Kerangas Rawa: Wetland-dominated heath forest with periodic flooding and adapted vegetation.

Fig. 2. Field observation documentation for heath forest classification.

Figure 2 provides field documentation of kerangas characteristics, emphasizing their unique traits. For instance, kerangas ecosystems are often associated with white sandy soils and sparse vegetation, indicating nutrient-poor conditions. Additionally, black water commonly observed in wetlands reflects high acidity (low pH), a hallmark of heath forest environments. These observations substantiate the spectral and ecological features captured in Sentinel-2 imagery, enhancing the robustness of the training dataset.

C. Data Annotation and Preprocessing

Ground truth annotations were established based on the collected field samples, supplemented by high-resolution Google Earth imagery and Sentinel-2 verification to ensure consistency in labeling. The dataset includes Sentinel-2 imagery covering the period from January 2024 to November 2024 as primary input data. Table I summarizes the preprocessing steps applied to the imagery before classification.

TABLE I. DATA PREPROCESSING STEPS FOR SENTINEL-2 IMAGERY

Step	Table Column Head
Cloud filtering	Exclude images with CLOUDY_PIXEL_PERCENTAGE > 10%.
Image scaling	Scale pixel values by a factor of 0.0001 to ensure compatibility with the reflectance range.
Compositing and clipping	A median composite of filtered images is generated to reduce noise, and the resulting image is clipped to the Region of Interest (ROI).
Visualization	The preprocessed image is visualized using the True Color Composite bands (B4, B3, B2) to assess its quality before analysis.

After the data preprocessing steps, Regions of Interest (ROIs) were delineated for each class of heath forests: Kerangas Kolam, Kerangas Rawa, and Kerangas Hutan. These ROIs represent diverse ecological characteristics and serve as the foundation for annotated training data. Figure 3 showcases a visual representation of the preprocessed Sentinel-2 imagery, highlighting the ROIs for each class. The True Color Composite bands (B4, B3, B2) were used to visualize the distinct features of these regions, ensuring the clarity and precision of the annotated samples.

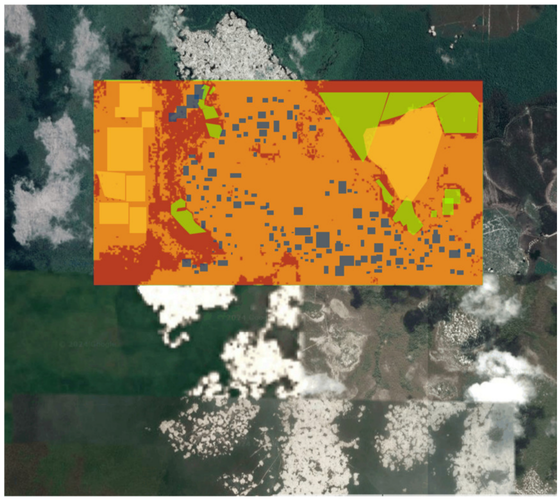

Fig. 3. Regions of Interest (ROI) for heath forest classes.

D. Feature Selection and Data Preparation

The spectral bands used for classification include Sentinel-2 bands: B2, B3, B4, B5, B6, B7, B8, B9, B11, and B12. These bands serve as input features for the ML models. Table II provides detailed information on the selected spectral bands, including their wavelengths and spatial resolutions.

TABLE II. SELECTED SENTINEL-2 SPECTRAL BANDS FOR FEATURE INPUT

Band	Wavelength (nm)	Spatial Resolution (m)
B2 (Blue)	490	10
B3 (Green)	560	10
B4 (Red)	665	10
B5 (Red Edge 1)	705	20
B6 (Red Edge 2)	740	20
B7 (Red Edge 3)	783	20
B8 (NIR)	842	10
B9 (Water Vapor)	945	60
B11 (SWIR 1)	1610	20
B12 (SWIR 2)	2190	20

The labeled training data were generated by sampling the preprocessed image over annotated regions with the following properties:

- Class labels: ML models were trained to classify heath forests into three distinct classes, based on their ecological characteristics and spectral signatures in Sentinel-2 imagery.
- Spatial resolution: 10-meter scale to match Sentinel-2 spatial resolution.

E. Machine Learning Model Training and Classification

The core of this study involves the implementation and comparison of multiple ML models for heath forest classification. Table III summarizes the ML models used, their key parameters, and their respective advantages.

TABLE III. ML MODELS USED FOR CLASSIFICATION

Model	Key parameters and advantages
RF	100 decision trees; Handles high-dimensional data effectively.
SVM	Linear kernel; Suitable for small datasets.
GBT	50 iterations; Robust against overfitting.
CART	Depth: 5; Trees: 100; Produces interpretable results.
KNN	5 neighbors; Euclidean distance; Simple and non- parametric.

For each model, the following steps were executed:

- Training: Models were trained using 80% of the data and the selected spectral bands as input features.
- Classification: The trained models classify the Sentinel-2 image into heath forest classes.
- Validation: The remaining 20% of the data were used to validate the classification results, generating confusion matrices and overall accuracy metrics.

F. Performance Metrics

The performance of each model was quantified using the following metrics:

- Confusion matrix: Provides a detailed breakdown of true and predicted classifications.
- Overall accuracy: Measures the proportion of correctly classified samples.
- Kappa coefficient: Evaluates the agreement between predicted and actual classifications while accounting for chance.

The results are compared across models to determine the most effective technique for heath forest classification.

G. K-Fold Cross Validation

A 5-fold Cross-Validation (K-fold CV) was performed to further evaluate the robustness of each model. The process involved the following steps:

- The labeled dataset was randomly partitioned into 5 equal folds.
- For each fold, one subset served as the testing data while the remaining four subsets were used for training.
- Models were trained and evaluated iteratively over all five folds.
- The accuracy and Kappa statistics were averaged across the folds to assess overall performance.

The use of K-fold CV ensures that the evaluation is unbiased and accounts for variability in the data distribution.

H. Experimental Workflow Summary

The experimental workflow can be summarized as follows:

- Sentinel-2 data collection, preprocessing, and annotation of training regions.
- Feature extraction using Sentinel-2 bands and data splitting into training and testing sets.
- Implementation of ML models.
- Validation using confusion matrices, overall accuracy, and K-fold CV.
- Comparative analysis of model performance and identification of the best-performing classifier.

III. RESULTS AND DISCUSSION

A. Model Selection and Rationale

The selection of ML models for the classification of Bornean heath forests was driven by their capacity to process high-dimensional spectral data from Sentinel-2 imagery effectively. Ensemble learning methods, particularly RF and GBT, leverage multiple decision trees to optimize classification performance. These models are well-suited for handling the spectral heterogeneity inherent in Sentinel-2 imagery, making them superior to linear classifiers such as SVM, which may struggle with non-linearly separable features in high-dimensional remote sensing data.

B. Comparison with Baseline Approach

To further evaluate the improvements achieved in this study, the supervised classification approach using Sentinel-2 imagery was compared with the unsupervised classification method used in [1], which relied on Landsat 9 imagery and GIS-based classification techniques. Table IV presents the key differences between these approaches.

TABLE IV. COMPARISON WITH BASELINE APPROACH

Aspects	Baseline [1]	This study
Satellite data	Landsat 9 (30m)	Sentinel-2 (10–20m)
Classification method	Unsupervised (ISO Cluster)	Supervised (RF, GTB, SVM, KNN, CART)
Input features	NDVI, land cover, elevation, soil pH, texture	Sentinel-2 Bands (B2–B12)
Accuracy	Not explicitly stated	Best performance on RF (96.66%) and GTB (96.69%)

This comparison highlights the advantages of integrating high-resolution Sentinel-2 imagery with ML models, as it significantly improves classification accuracy. The baseline approach relied on Landsat 9 data (30 m resolution) and unsupervised classification [1], which lacked the precision offered by supervised models. By leveraging Sentinel-2's finer spatial and spectral details, this method enables a more granular and reliable classification of heath forest ecosystems.

C. Classification Accuracy

Table V summarizes the performance of each classifier, demonstrating the high classification accuracy achieved across all models, with RF and GBT being the most effective.

TABLE V. PERFORMANCE FOR HEATH FOREST CLASSIFICATION

Model	Overall accuracy (%)	Kappa coefficient
RF	96.72	0.945
SVM	94.09	0.9
GBT	96.63	0.945
CART	95.92	0.932
KNN	96,13	0.938

These results indicate that RF and GBT deliver the most reliable classification, achieving an overall accuracy exceeding 96% with a Kappa coefficient above 0.94. The Kappa coefficient values suggest a near-perfect agreement between predicted and actual classifications, confirming the robustness of these models.

D. Confusion Matrix Analysis

The confusion matrix presented in Table VI summarizes the classification performance of different machine learning models across the three heath forest classes: Kerangas Kolam (Class 1), Kerangas Rawa (Class 2), and Kerangas Hutan (Class 3). Each cell in the table follows the format (True Positives, False Negatives, False Positives) for each class, separated by commas.

TABLE VI. CONFUSION MATRIX FOR MACHINE LEARNING MODELS

Model	Class 1 (Kerangas Kolam)	Class 2 (Kerangas Rawa)	Class 3 (Kerangas Hutan)
RF	1053, 15, 34	15, 704, 6	36, 0, 1370
SVM	1023, 12, 67	40, 664, 21	50, 1, 1355
GBT	1044, 20, 38	11, 708, 6	34, 0, 1372
CART	1040, 23, 39	16, 700, 9	45, 0, 1361
KNN	1043, 15, 44	21, 696, 8	36, 1, 1369

The key observations from the confusion matrix are:

- RF and GBT achieved the highest accuracy, with minimal misclassifications. RF correctly classified 1053 Kerangas Kolam samples, misclassifying 15 as Kerangas Rawa and 34 as Kerangas Hutan.
- SVM struggled with Class 2 (Kerangas Rawa), misclassifying 40 samples as Kerangas Kolam and 21 as Kerangas Hutan.
- CART and KNN exhibited moderate performance. CART showed higher misclassification for Kerangas Hutan (45 samples misclassified as Kerangas Kolam).
- Overall, ensemble models (RF and GTB) outperform traditional methods (SVM, CART, KNN), confirming their effectiveness in handling Sentinel-2 spectral complexity.

E. K-Fold Cross Validation

A 5-Fold CV was performed to further validate the stability and generalizability of the models. The results, presented in Table VII, confirm that RF and GBT maintain consistent performance across multiple training-test partitions. The cross-validation results demonstrate that the top-performing models maintain their effectiveness across different training splits, further validating their suitability for heath forest classification.

TABLE VII. K-FOLD CV RESULTS

Model	K-fold mean accuracy (%)	Kappa coefficient
RF	1053, 15, 34	15, 704, 6
SVM	1023, 12, 67	40, 664, 21
GBT	1044, 20, 38	11, 708, 6
CART	1040, 23, 39	16, 700, 9
KNN	1043, 15, 44	21, 696, 8

F. Discussion

The integration of Sentinel-2 imagery with ML models has proven highly effective in classifying Bornean heath forests. The superior performance of RF and GBT can be attributed to their ensemble learning capabilities, which mitigate overfitting while capturing intricate spectral patterns. KNN, despite being a relatively simple algorithm, also exhibited commendable performance, suggesting that Sentinel-2 spectral features are highly discriminative. However, SVM demonstrated lower accuracy, which may be due to its reliance on linear decision boundaries that are not optimally suited for the complex spectral distribution of heath forest ecosystems.

G. Implications for Conservation and Land Use Policy

The high classification accuracies achieved in this study highlight the potential for operationalizing ML-based forest monitoring systems. These systems could be employed to:

- Identify regions at risk of degradation through continuous remote sensing analysis.
- Generate high-resolution ecological maps to support conservation efforts and sustainable land-use planning.
- Inform policymakers and environmental agencies on data-driven strategies for mitigating climate change effects on heath forest ecosystems.

H. Limitations and Future Work

Despite the promising results, this study has several limitations that should be addressed in future research:

- Seasonal and atmospheric variability: Sentinel-2 imagery used in this study was limited to the period of January-November 2024. Multi-temporal data should be incorporated to account for seasonal variations and improve model robustness.
- The study focused on three primary heath forest classes. Additional subclasses and ecosystem variations should be explored for more comprehensive classification models.
- Potential for deep learning integration: Future studies should investigate the applicability of Convolutional Neural Networks (CNNs) for classifying complex forest structures, leveraging Sentinel-2's spatial and spectral richness to identify intricate ecological patterns.

To further enhance classification reliability, future work should explore data fusion techniques integrating multiple remote sensing platforms (e.g., Sentinel-1 radar data) to provide complementary structural and spectral insights into heath forest ecosystems.

IV. REPLICATION PACKAGE

To ensure reproducibility, all scripts, datasets, and model configurations used in this study are provided in the replication package available at [20]. This package includes Sentinel-2 imagery preprocessing steps, annotated training regions, model training scripts, and evaluation metrics. Users can follow the guidelines provided to replicate the experiments or adapt the method to other land cover classification tasks.

V. THREATS TO VALIDITY

Although the findings are promising, several threats to validity exist. Reliance on Sentinel-2 imagery can introduce biases due to seasonal variations or cloud cover, which were mitigated through strict preprocessing steps but could still influence the results. Additionally, the models were trained on specific heath forest types, limiting their generalizability to other ecosystems or regions. Another potential threat is the variability in training data quality. Although K-fold CV was employed to ensure robustness, the small sample size for certain classes could affect model reliability. Future work should address these limitations by incorporating larger and more diverse datasets, as well as evaluating the models under different environmental conditions.

VI. CONCLUSION

This study demonstrates the effectiveness of combining Sentinel-2 imagery with ML models for the classification of Bornean heath forest ecosystems. Unlike previous studies that primarily relied on Landsat-9 imagery with unsupervised classification methods, this study leverages the higher spatial and spectral resolution of Sentinel-2 and applies supervised ML models, resulting in significantly improved classification accuracy. The RF and GBT classifiers achieved overall accuracies above 96%, with strong Kappa coefficients, indicating high reliability. The application of 5-fold CV further confirmed the robustness of these models across varied data partitions. This novel approach refines heath forest classification by integrating high-dimensional spectral data with ensemble learning techniques, providing a more granular and precise monitoring method. The results underscore the potential of using high-resolution remote sensing data to detect subtle ecological variations, enabling scalable and precise monitoring of fragmented ecosystems. These findings are instrumental in conservation planning, land-use management, and policy-making aimed at preserving heath forests amidst escalating anthropogenic pressures. Future work should explore temporal dynamics to assess seasonal changes in vegetation, incorporate additional forest classes to improve generalization, and investigate deep learning architectures, such as CNNs, to further enhance classification accuracy and automation. Additionally, integrating multi-source remote sensing data could further improve classification robustness and applicability for large-scale ecosystem monitoring.

ACKNOWLEDGMENT

This research was funded by the IEEE Student Branch of Universitas AMIKOM Yogyakarta under the IEEE ASB Research Grant program for the 2024 fiscal year.

REFERENCES

[1] R. M. Nisaa, U. K. Sari, and Y. B. Sulistioadi, "Heath forest identification using remote sensing in the surrounding area of the new capital city of Indonesia," *IOP Conference Series: Earth and Environmental Science*, vol. 1266, no. 1, Sep. 2023, Art. no. 012068, https://doi.org/10.1088/1755-1315/1266/1/012068.

[2] M. A. Soendjoto, M. K. Riefani, and S. Diana, "Floristic diversity and composition of heath forest in Kuala Tambangan Village, Tanah Laut District, Indonesia," *Biodiversitas Journal of Biological Diversity*, vol. 24, no. 10, Nov. 2023, https://doi.org/10.13057/biodiv/d241024.

[3] I. S. Mohd Ikbal, S. Md Jaafar, N. Ahmad, and R. S. Sukri, "Contrasting environmental drivers of tree community variation within heath forests in Brunei Darussalam, Borneo," *Biodiversity Data Journal*, vol. 12, Dec. 2024, Art. no. e127919, https://doi.org/10.3897/BDJ.12.e127919.

[4] D. Oktavia, S. D. Pratiwi, S. Munawaroh, A. Hikmat, and I. Hilwan, "The potential of medicinal plants from heath forest: Local knowledge from Kelubi Village, Belitung Island, Indonesia," *Biodiversitas Journal of Biological Diversity*, vol. 23, no. 7, Jul. 2022, https://doi.org/10.13057/biodiv/d230731.

[5] M. A. Brovelli, Y. Sun, and V. Yordanov, "Monitoring Forest Change in the Amazon Using Multi-Temporal Remote Sensing Data and Machine Learning Classification on Google Earth Engine," *ISPRS International Journal of Geo-Information*, vol. 9, no. 10, Oct. 2020, Art. no. 580, https://doi.org/10.3390/ijgi9100580.

[6] D. Hoekman *et al.*, "Wide-Area Near-Real-Time Monitoring of Tropical Forest Degradation and Deforestation Using Sentinel-1," *Remote Sensing*, vol. 12, no. 19, Jan. 2020, Art. no. 3263, https://doi.org/10.3390/rs12193263.

[7] Y. Gao, M. Skutsch, J. Paneque-Gálvez, and A. Ghilardi, "Remote sensing of forest degradation: a review," *Environmental Research Letters*, vol. 15, no. 10, Jun. 2020, Art. no. 103001, https://doi.org/10.1088/1748-9326/abaad7.

[8] D. Phiri, M. Simwanda, S. Salekin, V. R. Nyirenda, Y. Murayama, and M. Ranagalage, "Sentinel-2 Data for Land Cover/Use Mapping: A Review," *Remote Sensing*, vol. 12, no. 14, Jan. 2020, Art. no. 2291, https://doi.org/10.3390/rs12142291.

[9] K. O. Abduraufovich, "Use of GIS Technologies and Satellite Data for Geological Research," *Western European Journal of Modern Experiments and Scientific Methods*, vol. 2, no. 11, pp. 24–30, Nov. 2024.

[10] J. Wang, M. Bretz, M. A. A. Dewan, and M. A. Delavar, "Machine learning in modelling land-use and land cover-change (LULCC): Current status, challenges and prospects," *Science of The Total Environment*, vol. 822, May 2022, Art. no. 153559, https://doi.org/10.1016/j.scitotenv.2022.153559.

[11] M. S. -Dastjerdi and A. Lak, "Towards resilient place emphasizing urban form: An assessment framework in urban design," *Sustainable Cities and Society*, vol. 96, Sep. 2023, Art. no. 104646, https://doi.org/10.1016/j.scs.2023.104646.

[12] C. Sothe *et al.*, "A comparison of machine and deep-learning algorithms applied to multisource data for a subtropical forest area classification," *International Journal of Remote Sensing*, vol. 41, no. 5, pp. 1943–1969, Mar. 2020, https://doi.org/10.1080/01431161.2019.1681600.

[13] O. Okwuashi and C. E. Ndehedehe, "Deep support vector machine for hyperspectral image classification," *Pattern Recognition*, vol. 103, Jul. 2020, Art. no. 107298, https://doi.org/10.1016/j.patcog.2020.107298.

[14] R. H. Topaloğlu, E. Sertel, and N. Musaoğlu, "Assessment of Classification Accuracies of Sentinel-2 and Landsat-8 data for Land Cover/Use Mapping," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLI-B8, pp. 1055–1059, Jun. 2016, https://doi.org/10.5194/isprsarchives-XLI-B8-1055-2016.

[15] R. N. Masolele, "Assessing land use following tropical deforestation : Combining remote sensing and deep learning," Ph.D. dissertation, Wageningen University, 2023.

[16] A. Arfa and M. Minaei, "Utilizing multitemporal indices and spectral bands of Sentinel-2 to enhance land use and land cover classification with random forest and support vector machine," *Advances in Space Research*, vol. 74, no. 11, pp. 5580–5590, Dec. 2024, https://doi.org/10.1016/j.asr.2024.08.062.

[17] A. Tariq, J. Yan, A. S. Gagnon, M. Riaz Khan, and F. Mumtaz, "Mapping of cropland, cropping patterns and crop types by combining optical remote sensing images with decision tree classifier and random forest," *Geo-spatial Information Science*, vol. 26, no. 3, pp. 302–320, Jul. 2023, https://doi.org/10.1080/10095020.2022.2100287.

[18] C. Brown, S. Sjögersten, M. J. Ledger, F. Parish, and D. Boyd, "Remote Sensing for Restoration Change Monitoring in Tropical Peat Swamp Forests in Malaysia," *Remote Sensing*, vol. 16, no. 15, Jan. 2024, Art. no. 2690, https://doi.org/10.3390/rs16152690.

[19] P. S. Nagendram, P. Satyanarayana, and P. R. Teja, "Mapping Paddy Cropland in Guntur District using Machine Learning and Google Earth Engine utilizing Images from Sentinel-1 and Sentinel-2," *Engineering, Technology & Applied Science Research*, vol. 13, no. 6, pp. 12427–12432, Dec. 2023, https://doi.org/10.48084/etasr.6460.

[20] "Replication package for land cover classification study," *Google Earth Engine Code Repository*. https://code.earthengine.google.com/8fae384fe296584613c96049f3cd15d3.

AUTHORS PROFILE

Dwi Ahmad Dzulhijjah is currently a Ph.D. student in Cyber-Physical Systems at Politeknik Elektronika Negeri Surabaya (PENS), Indonesia, focusing on real-time traffic management and anomaly detection. He holds two master's degrees in Artificial Intelligence from Universitas Gadjah Mada (UGM) and Informatics focusing on Business Intelligence from Universitas AMIKOM Yogyakarta, as well as bachelor's degrees in Biology and Informatics Engineering. Additionally, he has earned a professional engineering degree in Informatics. He is also an AI Engineer at PT. Piranti Kecerdasan Buatan (PIKEBU), specializing in computer vision and edge computing solutions. His research interests include multi-agent systems, distributed AI, and the integration of AI in smart city applications.

Kusrini is a professor at Universitas AMIKOM Yogyakarta Indonesia. She fulfilled her Ph.D. in the Computer Science Study Program at Universitas Gadjah Mada Yogyakarta Indonesia in 2010. She is interested in exploring many things related to machine learning and other AI fields. She also loves to research decision support systems and databases. She is a member of the IEEE and IEEE Systems, Man, and Cybernetics Society. Her website is http://kusrini.com/.

Rodrigo Martínez-Béjar is a Professor of Computer Science in the Department of Information and Communication Engineering at the University of Murcia in Spain. He received his Master's degree in Computer Science from the University of Malaga in 1992 and his Ph.D. from the University of Murcia in Spain in 1997. In addition, he holds a Master's degree in Applied Sociology from the University of Murcia. He has held visiting positions at several European, Australian, and Latin-American Universities, including the Australian Universities of Adelaide and New South Wales and the French University of Southern Paris. He has served on several Editorial Boards and coauthored more than 100 scientific publications, most of them in peer-reviewed journals and conferences. He has led more than 30 national or international competitive grants and research contracts with industry. In 2006, he became the Head of an institutional, inter-disciplinary University research group.