Land Use Mapping and Zoning Using Machine Learning and Satellite Images

Surya Vardhan Karella,[1, a)] Reddy Neeraj.R,[1, b)] and Aruna M[2, c)]

[1] *Final Year B.Tech student, Department of Computing Technologies, College of Engineering and Technologies, SRM Institute of Science and Technology, Kattankulathur, Chennai, 603203, Tamilnadu, India*
[2] *Associate Professor, Department of Computing Technologies, Faculty of Engineering and Technology, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Chennai, 603203, Tamilnadu, India*

[a)] Corresponding author: *suryavardhankarella@gmail.com*
[b)] *r.reddyneeraj@gmail.com*
[c)] *arunam@srimst.edu.in*

Abstract.
Territorial management has emerged as a critical component of sustainable development, particularly in rapidly expanding metropolitan regions. Traditional terrain categorization frameworks, which often depend on human assessment and fundamental picture analysis techniques, struggle to handle the complex and voluminous datasets generated by modern remote sensing technologies. This study explores innovative approaches to environmental zoning by integrating satellite imagery with advanced machine learning algorithms, specifically neural networks, to automate the identification of surface elements. The goal is to optimize geographic resource allocation for crop cultivation, city development and industrialization. In this research, a specialized ground feature identification framework was developed, calibrated, and validated using high-resolution Sentinel-2 satellite data alongside location-specific satellite imagery. The model focuses on identifying key land cover categories such as barren land, vegetation, and water bodies. Using high-resolution imagery, the approach enables a precise classification, which is then visually represented by overlaying the results on the corresponding satellite images. Each land type is annotated and analyzed to enhance clarity and facilitate further investigation. The findings demonstrate that the integration of remote sensing data with machine learning techniques offers an efficient, reliable, and scalable solution to land zoning challenges. This methodology not only streamlines the decision-making process but also supports environmentally and economically sustainable land use practices. By automating the classification process, the study highlights the potential of these technologies to address pressing land management issues in fast-developing urban and rural landscapes.

Keywords: Urban planning, environmental sustainability, remote sensing, sustainable development, neural networks, TensorFlow, machine learning, land use zoning, satellite imagery analysis, image patch processing.

BACKGROUND

Land zoning plays a pivotal role in urban planning, serving as a cornerstone for environmental management and the sustainable development of urban regions. It involves the allocation of specific land areas for distinct purposes, such as agriculture, industry, residential use, and conservation. Over time, the methods for land use analysis have evolved significantly. Traditional approaches, reliant on manual mapping, historical records, and local expertise, were time-consuming, prone to errors, and struggled to keep pace with the rapid urbanization and dynamic land use changes. These outdated practices no longer meet the demands of modern land management and policy-making.

Recent advancements in high-resolution satellite imagery and state-of-the-art machine learning techniques offer transformative potential for land zoning processes. Satellite images have long been instrumental in capturing extensive land data, providing critical insights into vegetation, water bodies, infrastructure, soil types, and other key attributes. By integrating these images with advanced neural networks, it is now possible to automate land classification and zoning with unprecedented accuracy and efficiency.

This paper focuses on leveraging high-resolution satellite imagery, such as Sentinel-2 datasets, alongside neural networks to classify land cover types, including vegetation, barren land, and water bodies. The proposed methodology combines remote sensing, decision rules, and image interpretation to streamline the zoning process. Unlike traditional methods, this approach is scalable, environmentally friendly, and significantly reduces the complexity and effort associated with conventional land zoning practices. It equips decision-makers and metropolitan designers with a powerful instrument for environmentally responsible terrain allocation. The following chapters of this study explore relevant prior research, the neural network-based framework, experimental results, and the broader implications of this approach for sustainable urban development and environmental conservation.

LITERATURE SURVEY

Kerins [1] further expanded this understanding by analyzing urban expansion in the Asian subcontinent and Latin American nation, highlighting the relationship between socioeconomic variables and satellite-based monitoring methods.

Kussul [2] rdemonstrated the impressive precision capabilities of deep learning image processing frameworks when applied to farming evaluations through deep learning techniques.

Breiman's foundational investigations regarding Categorization and Prediction Structures (CPS) [3] formed an essential foundation for comprehending choice-branching computational processes [3]. Subsequently, his research on bagging predictors [4] introduced techniques to enhance algorithmic stability and predictive accuracy.

Ouchra [5] demonstrated the efficiency of Google Earth Engine in processing large datasets for land cover classification.

Batunacun [6] examined environmental shifts on the Mongolian Plateau, revealing critical interactions between human activities and land degradation from 1975 to 2015.

Feng [7] research on metropolitan district planning employed advanced probabilistic spatial models, incorporating multiple-perspective visual information to reveal the complexity of urban land use dynamics.

Tung [8] exemplified practical application by examining metropolitan growth in the Vietnamese capital using artificial intelligence algorithms and chronological orbital photography.

Bayas [9] explored terrain categorization utilizing European Space Agency's multispectral data, implementing diverse computational learning to achieve high-accuracy mapping. Complementing this, Weng [10] explored deep convolutional neural networks with transfer learning , showcasing improved classification performance in scenarios with limited training data.

Belgiu and Drăguț [11] conducted a comprehensive review in remote sensing regarding random forest applications, highlighting the algorithm's versatility in handling high-dimensional land cover classification tasks. Their research underscored the significance of machine learning techniques in environmental monitoring.

Mitra and Basu [12] have done a comprehensive survey of deep learning techniques machine learning for land cover classification, highlighting methodological evolution and persistent challenges.

Chen [13] advanced land cover mapping by combining object-based and pixel-based classification methodologies from remotely sensed imagery. Their approach demonstrated the potential of methodological convergence in improving classification outcomes.

Ojwang [14] proposed an innovative hierarchical classification methodology was employed to map land use within intricate social-ecological systems. The study highlighted the importance of adopting multi-tiered analytical frameworks that integrate a wide range of ecological and socio-economic variables for comprehensive analysis.

RESEARCH DESIGN AND METHODOLOGY

This study employs a neural network (NN) approach for discriminating between barren land, water, and vegetation using Sentinel satellite imagery. The research methodology encompasses a comprehensive strategy of image preprocessing, neural network design, and advanced classification techniques.

The dataset preparation involved creating three distinct class folders representing barren land, water, and vegetation. Image preprocessing was critical, standardizing all images to 256×256 pixels to ensure uniform input dimensions and facilitate consistent feature extraction. This preprocessing step is fundamental to maintaining data quality and model reliability.

The neural network architecture follows a carefully designed sequential model. The initial Flatten layer transforms the 256×256×3 image into a comprehensive array, enabling efficient data representation. A subsequent Dense layer with units of 128 count and activation function ReLU captures intricate image features, while the final Dense layer utilizes softmax activation to perform multi-class classification, precisely categorizing land cover types.

Training optimization incorporated sophisticated techniques to enhance model performance and generalization. The ReduceLROnPlateau callback dynamically adjusts the learning rate when validation loss plateaus, preventing model stagnation. Simultaneously, the EarlyStopping callback terminates training and restores optimal model weights based on validation loss trends. Data augmentation methods, including random rotations, shifts, and horizontal flips, were implemented to improve model robustness and prevent overfitting. The experimental setup employed the optimizer Adam with loss function as categorical cross-entropy, using accuracy as the primary performance metric. The training configuration consisted of 10 epochs with an 80-20 train-validation data split, providing a robust

approach to model development and validation.
Model evaluation involved comprehensive performance analysis using a dedicated test set, calculating test loss and test accuracy. Innovative visual overlay techniques were developed to demonstrate classification outcomes for each land cover type, offering intuitive interpretation of the model's performance. These visualization methods provide critical insights for urban planning, resource management, and environmental monitoring applications.

DATA ACQUISITION

The data collection methodology leveraged satellite imagery from publicly accessible platforms, including Sentinel-2, Earth Explorer, and curated datasets from Kaggle [15] [16]. These platforms provide comprehensive satellite imagery with spectral bands critical for distinguishing diverse land cover types.

The selected satellite images encompass three primary land cover categories: barren land, vegetation, and water bodies. By utilizing patch-based analysis, the research captures nuanced terrain variations across different geographical regions. The spectral indices and multispectral data incorporated into the neural network modeling facilitate robust training and validation processes.

Kaggle's diverse image repositories enabled the compilation of a representative dataset, ensuring the classification model's adaptability to varied terrain characteristics. The extensive image collection strategically represents global land cover diversity, enhancing the model's generalization capabilities and potential for comprehensive environmental classification. The methodological approach prioritizes data heterogeneity, selecting images that capture the intricate spatial and spectral variations across different ecological landscapes. This comprehensive data collection strategy forms the foundational framework for developing a sophisticated land cover classification model with broad applicability.

FIGURE 1. Sample Sentinel-2 Imagery: Water Region

DATA TRANSFORMATION

Data preprocessing represents a critical phase in developing robust neural network-based image classification models. The methodology implemented a systematic approach to prepare raw image data for effective computational analysis.The initial preprocessing stage involved loading a structured dataset with images organized in specific directory hierarchies. A standardized resizing procedure was applied, transforming all images to a uniform 256x256 pixel dimension. This dimensionality standardization ensures consistent input representation, enabling neural networks to process images uniformly across computational batches.

Pixel value normalization followed resizing, scaling pixel intensities from the original [0, 255] range to [0, 1] through division by 255.0. This normalization technique significantly enhances optimization algorithm convergence rates and facilitates more stable neural network training. The transformed pixel values provide a mathematically consistent input space, mitigating potential computational variations introduced by diverse image characteristics.

Image reshaping was subsequently performed to align the data format with the neural network's expected input architecture. This transformation enables efficient batch processing during both training and evaluation phases, creating a computational framework that supports advanced feature extraction and learning mechanisms.The comprehensive preprocessing strategy—encompassing loading, resizing, normalization, and reshaping—establishes an optimal foundation for neural network model training. By preparing the dataset through these refined techniques, the research ensures enhanced feature extraction capabilities, improved classification accuracy, and superior model generalization potential.

MODEL IMPLEMENTATION

The research implemented a comprehensive neural network approach for land cover classification, emphasizing robust dataset preparation and model architecture. Initial data preprocessing involved strategically organizing images into subdirectories representing distinct land cover categories, ensuring structured and systematically labeled datasets.

Dataset partitioning followed a carefully defined allocation strategy: 70% for training, 10% for validation, and 20% for testing. This methodological approach ensures comprehensive model evaluation while maintaining statistically significant data representation across different computational phases.The Keras ImageDataGenerator class was employed to facilitate advanced preprocessing techniques. This implementation enabled efficient pixel value rescaling and batch generation directly at the directory level, streamlining data preparation for subsequent neural network training.

Model architecture incorporated activation function called softmax in the output layer, facilitating probabilistic estimation of land cover class distributions. The Adam optimizer, coupled with categorical cross-entropy loss function, provided a robust computational framework for model training. Regularization techniques were strategically implemented to enhance model stability and prevent overfitting. Callback features, including ReduceLROnPlateau and EarlyStopping were integrated to dynamically manage learning rates and terminate training when performance plateaus. The model underwent comprehensive training across 10 iterations, with continuous validation data monitoring to assess performance and generalization capabilities. This methodological approach ensures a rigorous and systematic approach to land cover classification using neural network techniques.

```
In [8]: cnn_model.summary()
```

Model: "sequential"

Layer (type)	Output Shape	Param #
flatten (Flatten)	(None, 196608)	0
dense (Dense)	(None, 128)	25,165,952
dense_1 (Dense)	(None, 3)	387

Total params: 75,499,019 (288.01 MB)

Trainable params: 25,166,339 (96.00 MB)

Non-trainable params: 0 (0.00 B)

Optimizer params: 50,332,680 (192.00 MB)

FIGURE 2. Model Structural Summary

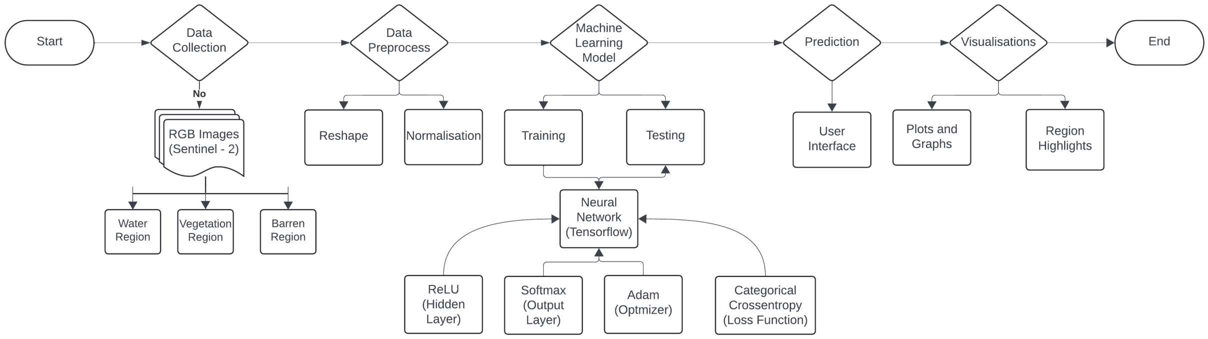
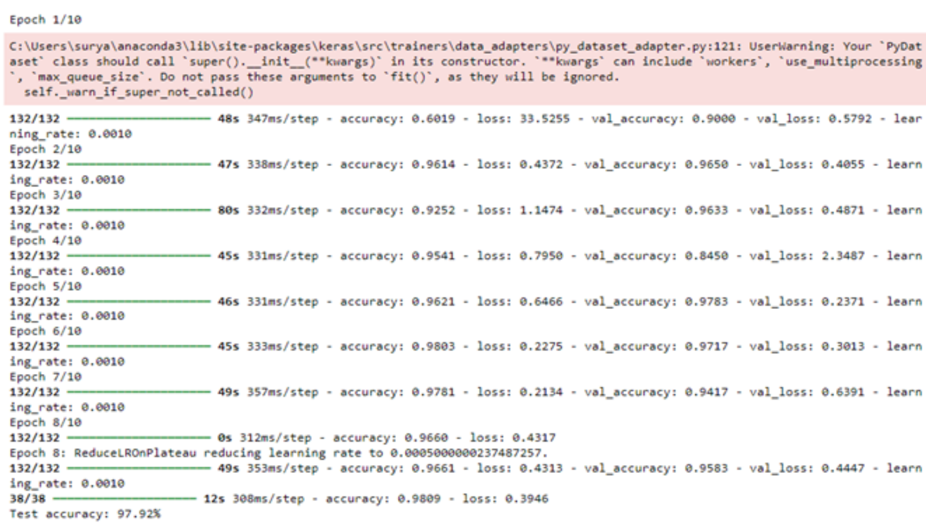


FIGURE 3. Architectural Framework

PERFORMANCE EVALUATION

The primary research objective was to develop a sophisticated neural network model capable of accurately classifying satellite imagery into Barren, Vegetation, and Water land cover categories. The proposed model underwent rigorous training and validation processes, demonstrating exceptional performance with a 97% classification accuracy.

This high-precision classification methodology demonstrates the significant potential of neural network techniques in environmental monitoring and land use forecasting. The research contributes valuable insights into advanced machine learning approaches for comprehensive ecological assessment and strategic land resource management.The model's remarkable accuracy underscores the effectiveness of deep learning algorithms in extracting complex spatial and spectral features from satellite imagery, providing a robust framework for precise land cover classification and environmental analysis.



```
Epoch 1/10
C:\Users\surya\anaconda3\lib\site-packages\keras\src\trainers\data_adapters\py_dataset_adapter.py:121: UserWarning: Your `PyDat
aset` class should call `super().__init__(**kwargs)` in its constructor. `**kwargs` can include `workers`, `use_multiprocessing
`, `max_queue_size`. Do not pass these arguments to `fit()`, as they will be ignored.
  self._warn_if_super_not_called()
132/132 ━━━━━━━━━━━━━━━━━━━━ 48s 347ms/step - accuracy: 0.6019 - loss: 33.5255 - val_accuracy: 0.9000 - val_loss: 0.5792 - lear
ning_rate: 0.0010
Epoch 2/10
132/132 ━━━━━━━━━━━━━━━━━━━━ 47s 338ms/step - accuracy: 0.9614 - loss: 0.4372 - val_accuracy: 0.9650 - val_loss: 0.4055 - learn
ing_rate: 0.0010
Epoch 3/10
132/132 ━━━━━━━━━━━━━━━━━━━━ 80s 332ms/step - accuracy: 0.9252 - loss: 1.1474 - val_accuracy: 0.9633 - val_loss: 0.4871 - learn
ing_rate: 0.0010
Epoch 4/10
132/132 ━━━━━━━━━━━━━━━━━━━━ 45s 331ms/step - accuracy: 0.9541 - loss: 0.7950 - val_accuracy: 0.8450 - val_loss: 2.3487 - learn
ing_rate: 0.0010
Epoch 5/10
132/132 ━━━━━━━━━━━━━━━━━━━━ 46s 331ms/step - accuracy: 0.9621 - loss: 0.6466 - val_accuracy: 0.9783 - val_loss: 0.2371 - learn
ing_rate: 0.0010
Epoch 6/10
132/132 ━━━━━━━━━━━━━━━━━━━━ 45s 333ms/step - accuracy: 0.9803 - loss: 0.2275 - val_accuracy: 0.9717 - val_loss: 0.3013 - learn
ing_rate: 0.0010
Epoch 7/10
132/132 ━━━━━━━━━━━━━━━━━━━━ 49s 357ms/step - accuracy: 0.9781 - loss: 0.2134 - val_accuracy: 0.9417 - val_loss: 0.6391 - learn
ing_rate: 0.0010
Epoch 8/10
132/132 ━━━━━━━━━━━━━━━━━━━━ 0s 312ms/step - accuracy: 0.9660 - loss: 0.4317
Epoch 8: ReduceLROnPlateau reducing learning rate to 0.0005000000237487257.
132/132 ━━━━━━━━━━━━━━━━━━━━ 49s 353ms/step - accuracy: 0.9661 - loss: 0.4313 - val_accuracy: 0.9583 - val_loss: 0.4447 - learn
ing_rate: 0.0010
38/38 ━━━━━━━━━━━━━━━━━━━━ 12s 308ms/step - accuracy: 0.9809 - loss: 0.3946
Test accuracy: 97.92%
```

FIGURE 4. Training Progress Across Epochs

Comprehensive model assessment involved computing area coverage percentages through patch-based classification of a test image. The quantitative analysis revealed nuanced land cover distribution: Barren land constituted 2.6%, Vegetation encompassed 46.75%, and Water dominated with 50.65% coverage. These results highlight the region's predominantly aquatic and vegetative landscape characteristics.Visualization techniques employed color-coded overlay masks to represent classification outcomes. The color mapping strategy assigned distinct colors to each land cover category: red for Barren land, yellow for Vegetation, and blue for Water. This approach enhances result interpretability, enabling stakeholders to rapidly comprehend spatial land cover distributions.

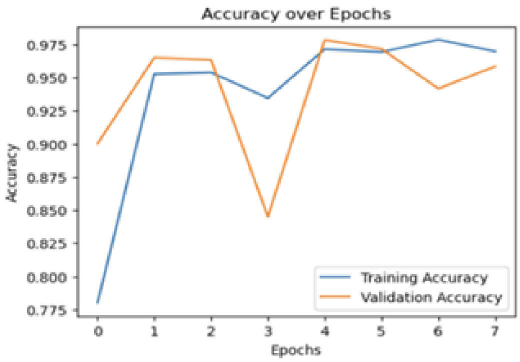


FIGURE 5. Epoch-wise Accuracy Analysis

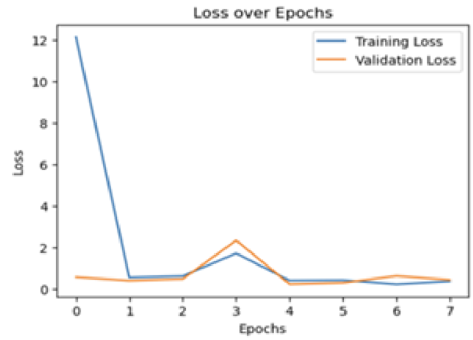


FIGURE 6. Epoch-wise Loss Analysis

The visual representations demonstrated the model's sophisticated classification capabilities, revealing intricate land cover boundaries and proportions. However, the analysis acknowledged inherent limitations of patch-based prediction methodologies, recognizing potential variability in classification precision due to complex imagery characteristics. The research has the potential of advanced neural network techniques in providing detailed, visually intuitive environmental monitoring insights, offering a robust framework for land cover assessment and ecological analysis.

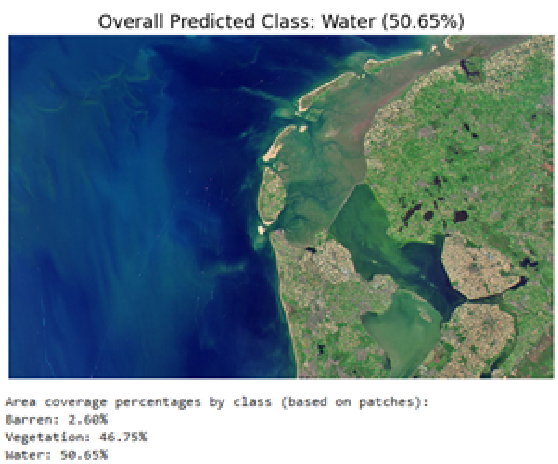


FIGURE 7. Model Assessment Metrics

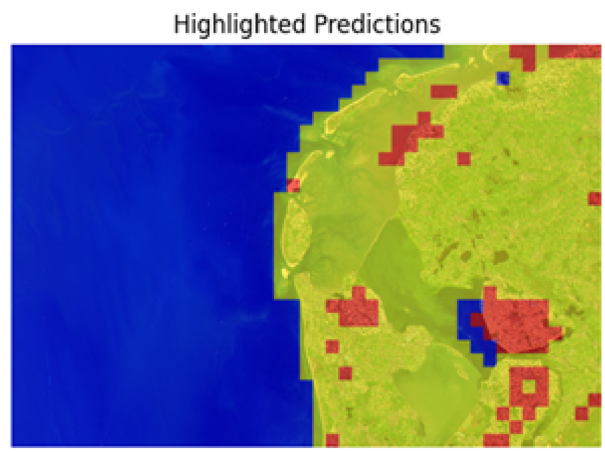


FIGURE 8. Visualizing Key Predictions

CONCLUSION

Neural network models for satellite image land cover classification represent a pivotal approach to territorial resource management. The research demonstrates the potential of high-resolution satellite imagery in precisely delineating Barren, Vegetation, and Water land cover classes, with inherent capabilities for continuous model refinement through iterative learning.The quantitative analysis of area coverage percentages provides critical insights for stakeholders, facilitating informed decision-making in land use planning and natural resource conservation. By developing a comprehensive classification framework, the research offers a sophisticated tool for sustainable land management strategies.

The integration of neural network techniques with remote sensing technologies illuminates a transformative pathway for environmental monitoring. Visualization techniques that translate complex classification data into intuitive color-coded representations bridge the gap between sophisticated computational analysis and practical ecological understanding. This methodological approach provides a robust framework for supporting sustainable development initiatives across diverse geographical contexts.

REFERENCES

1. E. M. T. R. P. Kerins, B. Guzder-Williams and E. Pietraszkiewicz, “Mapping urban land use in india and mexico using remote sensing and machine learning,” (2023).
2. S. S. N. Kussul, M. Lavreniuk and A. Shelestov, “Deep learning classification of land cover and crop types using remote sensing data,” arXiv (2015).
3. R. O. L. Breiman, J. Friedman and C. Stone, *Classification and Regression Trees* (Routledge, 1984).
4. L. Breiman, “Bagging predictors,” Mach. Learn. **24**, 123–140 (1996).
5. A. B. H. Ouchra and A. Erraissi, “Machine learning algorithms for satellite image classification using google earth engine and landsat satellite data: Morocco case study,” IEEE Access (2023).
6. Y. H. B. Batunacun, C. Nendel and T. Lakes, “Land-use change and land degradation on the mongolian plateau from 1975 to 2015—a case study from xilingol, china,” Land Degrad. Dev. (2019).
7. D. N. J. K. L.-F. Y. A. B. T. Feng, Q.-T. Truong and S.-K. Yeung, “Urban zoning using higher-order markov random fields on multi-view imagery data,” in *Proc. 2023 IEEE/CVF Conf. Comput. Vis. Pattern Recognit.* (2023).
8. H. T. T. N. D. H. D.T. Tung, N.T. Tung and P. Linh, “Analyzing urban expansion in hanoi using machine learning and multi-temporal satellite imagery,” Remote Sens. (2023).
9. I. D. P. K. S. Bayas, S. Sawant and A. Joshi, “Land use land cover classification using different ml algorithms on sentinel-2 imagery,” Int. J. Res. Appl. Sci. Eng. Technol. (2023).
10. C.-Y. Weng, “Land-use classification via transfer learning with a deep convolutional neural network,” Front. Environ. Sci. (2024).
11. M. Belgiu and L. Drăguţ, “Random forest in remote sensing: A review of applications and future directions,” ISPRS J. Photogramm. Remote Sens. **114**, 24–31 (2016).
12. S. Mitra and S. Basu, “Remote sensing-based land cover classification using machine learning and deep learning: A comprehensive survey,” Front. Remote Sens. **14** (2023).
13. Y. G. R. A. Y. Chen, Y. Zhou and Y. Chen, “Enhancing land cover mapping through integration of pixel-based and object-based classifications from remotely sensed imagery,” Remote Sens. **10** (2018).
14. M. S. M. O. S. K.-F. V. B. G. R. B. G.O. Ojwang, J.O. Ogutu and H. Olff, “An integrated hierarchical classification and machine learning approach for mapping land use and land cover in complex social-ecological systems,” Front. Remote Sens. (2023).
15. Kaggle, “Satellite images of regions- sentinel-2,” (2023).
16. Kaggle, “Satellite images of water bodies – sentinel-2,” (2023).