

Deep learning versus traditional artificial intelligence algorithms for remote sensing applications

Eslam Mostafa Elnahas^{1,*}, Ayman Rashad Elshehaby¹, Mahmoud Salah Gomaa¹

¹ Department of Geomatics Engineering, Faculty of Engineering at Shoubra, Benha university, Cairo, Egypt.

*Corresponding author

E-mail address: islam.elnahas@feng.bu.edu.eg, ayman.elshehaby@feng.bu.edu.eg, mahmoud.goma@feng.bu.edu.eg

Abstract: Remote sensing is an important mechanism for monitoring and oversight the environment, natural resources, and human activities. In this research, the focus was on utilizing deep learning (DL) algorithms, specifically a Convolutional Neural Network (CNNs) model with a U-Net architecture, for the classification of satellite image and because of there are multiple and extensive research contributions to the study of traditional artificial intelligence approaches in the field of remote sensing, this study was limited to a comprehensive review of some previous studies in this field. As for DL, a U-Net model was developed, utilized and it was trained on Sentinel-2 satellite image have a spatial resolution of 10 meters. The training stage was centered on data covering the Egypt especially the Delta region and the Nile Valley due to their agricultural and environmental importance. The model achieved a classification and segmentation accuracy of 94.14% and an Jaccard coefficient (IOU) value of 87.64%. These results confirm the strong potential of DL models, such as U-Net that resulted high accuracy in satellite image classification and segmentation tasks, especially when used specific geographical regions with distinctive characteristics.

Keywords: Remote Sensing, Deep Learning (DL), U-Net architecture, CNNs, Satellite Image, Image Segmentation, Google Earth Engine.

1. Introduction

The way we can see and understand what's happening on Earth from far away keeps getting better. There's been big improvements lately in using satellites and other remote Sensing techniques to keep track of nature, the environment, and what we humans are doing and the amount and types of satellite data have just exploded. So, we've really needed to come up with smarter ways to dig through it all to pull out the right info quickly. DL has been getting lots of attention because it's so good at finding patterns, even complicated ones in massive piles of data. These techniques totally blow away the old ways of analyzing satellite images. DL has kicked butt at picking out objects and labeling different parts of images. When it comes to making sense of remote sensing data, deep learning seems to really have a lot going for it. It's shown it can be extremely useful for types of tasks like identifying image classification, object detection, image Segmentation and Change detection. Khelifi and Mignotte [1]; Zhu et al. [2]; Dritsas and Trigka [3]

Land Use and Land Cover Change (LU/LC) is going through swift transformations on account of study of human population, superior to highly different countryside's, especially between urban and rural environments. This vital character demands studies across various dimensional and worldly scales. The study evaluates the accomplishment of machine learning algorithms in classifying multi-determination satellite images for LU/LC discovery. Three classification methods Random Forest (RF), Support Vector Machine (SVM), and their shapely combination were tested following harsh preprocessing. Results confirmed that SVM obtained accuracy above 0.96, stressing the importance of

selecting proper algorithms for monitoring land changes in quickly growing atmospheres. Rahman et al. [4]

In a study focused on classification in Beijing using Sentinel-2 satellite image established classifiers like SVM and default RF worked out categorization veracity levels of 47.1% and 79%, individually, when using RGB bands. Zhang and others. [5] Traditional AI algorithms like SVM, Decision Trees, and K-Means are productive for classifying Sentinel-2 satellite image, that offers extreme-resolution dossier for land cover tasks. These forms label features like houses, roads, and water bodies accompanying extreme veracity. Unsupervised methods like ISODATA to act well when refined to handle assorted pixels. Abburu and Golla. [6] The study presents a hybrid ensemble method for joining ANN, SVM, and Classification Trees, enhancing classification accuracy to 95.6% for constructions, roads, and plants. The approach outperforms individual classifiers and standard FMV while asserting computational effectiveness. Salah [7]. Artificial neural networks (ANN) are established in remote sensing for LU/LC categorization, upholding requests like urban preparation. While ANN offers strong pattern acknowledgment, added arrangements like SVM, fuzzy logic, and genetic algorithms each have singular substances and restraints inaccuracy, interpretability, and computational effectiveness. Abaidoo et al. [8]

DL algorithms have given better performance in satellite image classification, especially accompanying the chance of high- resolution imagery from platforms like Sentinel-2A and Sentinel-2B. These satellites supply 13 multispectral bands at variable spatial resolutions (10 m, 20 m, and 60 m), contribution rich spectral and geographical information

important for preparation deep neural networks. In particular, CNNs have demonstrated strong performance in learning hierarchical geographic and spectral features from satellite imagery without the need for manual feature extraction. In contrast to traditional methods, DL methods have the ability to independently acquire meaningful features directly from raw pixel data. Drusch et al. [9]; Kussul et al. [10]; Ienco et al. [11]; Zhang and Roy [12]; Ayush et al. [13]

Studies employing Sentinel-2 imagery have displayed that DL models outperform classical classifiers. Special scenarios, specifically when abundant, labeled datasets are convenient. For example, CNN based on U-Net located models applied to city land cover categorization have achieved overall accuracies surpassing 98% when prepared with Full-bandwidth Sentinel-2 data, significantly lowering misclassification in complex districts such as class confines. These models more exhibit greater strength to spectral contrast and are suitable for capture both worldwide framework and fine-grained analyses inside the imagery. Ma et al. [14]; Zhang et al. [15]

U-Net is a type of CNNs design developed specifically designed for semantic segmentation, and it was originally tailored for use in biomedical concept treatment applications. Ronneberger et al. [16]. U-Net has because gained extensive ratification in various fields, containing Remote sensing and segmentation of satellite images due to its ability to produce exact, pixel predictions Khan et al.[17] The construction trails a symmetric encoder-interpreter makeup: the contracting course captures circumstances through successive loop and pooling movements the broad pathway expedites exact localization by employing up sampling movements and merging high-determination dimensional details from the encoder through skip connections. Ronneberger et al.[16].

One of the key strengths of U-Net is its ability to do well with relatively limited training datasets, making it ideal for remote sensing uses where annotated data may be limited. In satellite image classification, U-Net is particularly direct at extracting detailed looks to a degree roads, buildings, and water bodies, even in complex city surroundings. Its end-to-end architecture enables it to learn both reduced-level and high-ranking features, happening in extreme classification veracity across diversified LC classes. As a result, U-Net has become a standard structure for DL-based segmentation of high-resolution remote sensing image. Zhang et al. [18]; Li et al. [19]

DL has emerged as a transformative solution, as demonstrated by the evaluation of seven architectures (SegNet, PSPNet, UNet, UNet++, DepLabV3+, SegFormer, and SegViT) on the ISPRS Potsdam dataset. The results show that SegViT achieves superior accuracy (85.2% of the unit of measure) and preserves edges but requires more computational resources highlighting the need to choose a task-specific model that balances accuracy and efficiency. Jiang.[19]

2. STUDY AREA AND DATA SOURCE

The study area includes sections of the Delta of the Nile and Valley regions in Egypt. A total of 8 Sentinel-2 tiles were downloaded using the Google Earth Engine Python

API. Each tile covers an area of $110 \times 110 \text{ km}^2$, its area is about $10,000 \text{ km}^2$ per tile, with imagery provided at a 10-meter spatial resolution. The total area covered by the tiles is $80,000 \text{ km}^2$ these tiles were selected to capture a representative variety of land cover types, including dense agricultural zones, expanding urban areas, and Water bodies. The dataset provides a rich foundation for analyzing spectral variability and detecting seasonal changes in this ecologically and economically vital region each tile corresponds to a $11,000 \times 11,000$ -pixel image, The dataset comprises approximately 800 million pixels distributed across eight image tiles. In the north, irrigation networks dominate, while the central areas include mixed croplands and the southern part shows increasing aridity. Hammam et al. [20]. The satellite images gathered from 2017 to 2022 were processed to correct for atmospheric interference and to standardize their radiometric values. Following this preprocessing, the data was split into three subsets: 70% for model training, 15% for validation purposes, and the remaining 15% for final testing, facilitating both classification and segmentation analyses. Goodfellow et al. [21]

Sentinel-2 is a dual-satellite Earth observation mission developed by the European Space Agency (ESA) under the Copernicus Programmer. It supplies multispectral data across 13 primary bands covering the visible, near-infrared, and shortwave infrared spectral regions. Each Sentinel-2 satellite has a swath width of 290 kilometers, allowing for broad-area coverage in a single pass. Moreover, the constellation's five-day revisit interval enhances its capacity to monitor dynamic environmental and agricultural changes over time. Drusch et al. [9].

TABLE 1 Attribute and its specification of specification of Study area

Attribute	Specification
Region	Nile Delta & Valley, Egypt
Tile Count	8
Single Tile Coverage	$110 \times 110 \text{ km}^2 \approx (10,000 \text{ km}^2)$
Total Area	$80,000 \text{ km}^2$
Image Dimensions	$11,000 \times 11,000$ pixels
Aggregate Pixels	800 million
Resolution	10m (MSI)
Temporal Coverage	2017–2022
Processing Level	L2A (Bottom-of-Atmosphere)

The satellite imagery used in this study was obtained from Sentinel-2 Level-2A (L2A) products, which provide Bottom-of-Atmosphere (BOA) reflectance values, ensuring consistency in radiometric and atmospheric conditions. Specifically, Bands 2 (Blue), 3 (Green), and 4 (Red) each with a 10-meter spatial resolution were individually downloaded using the Google Earth Engine Python API and later composited into true-color RGB images via Python scripting. These bands correspond to central wavelengths of approximately 496.6 nm (B2), 560 nm (B3), and 664.5 nm

(B4), offering high fidelity in capturing surface features and land cover patterns. All images were exported in GeoTIFF(.tif) format, which preserves geospatial metadata and ensures compatibility with remote sensing and GIS software for subsequent analysis. The Sentinel-2 Level-2A surface reflectance data, provided through Google Earth Engine, offers harmonized imagery suitable for land cover classification and change detection (Google Earth Engine, n.d.). [22]

3. METHODOLOGY

A general overview of the proposed methodology is illustrated in Figure 3. This framework summarizes the sequential stages starting from Sentinel-2 image acquisition, preprocessing steps, dataset annotation, model development and evaluation. Each of these phases is elaborated upon in the following subsections.

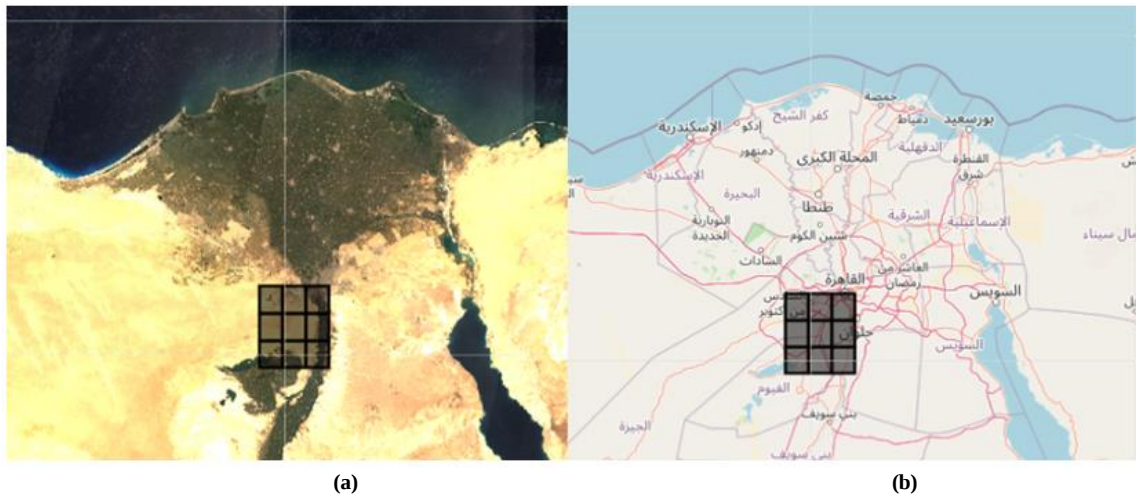


FIGURE 1 A general view of the study area from (a) Sentinel-2 and (b) OpenStreetMap, respectively. And it includes a tile divided into 9 images.

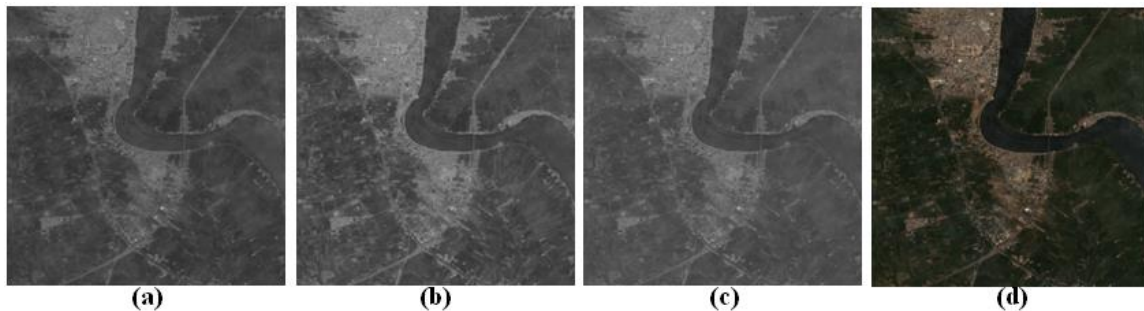


FIGURE 2 A sample of the uploaded images, bands 2, 3, and 4 in order, and lastly the color composite image.

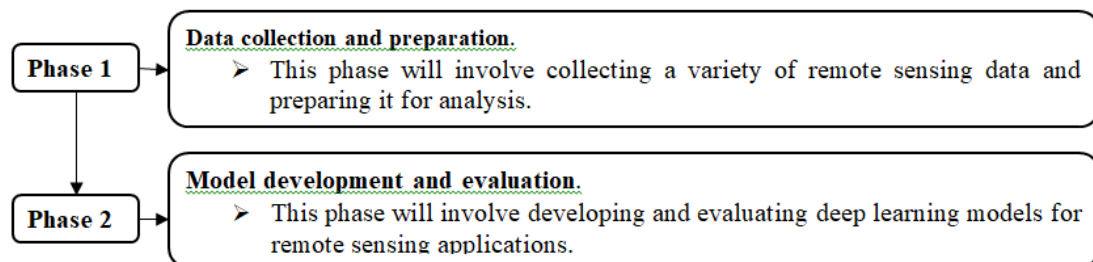


FIGURE 3 Methodology for Satellite Image Segmentation

3.1 Data collection and preparation:

The dataset was composed and pre-processed using Python scripts performed within a Jupyter Notebook environment. Specifically, Sentinel-2 Level-2A imagery was achieved via the Google Earth Engine Python API and were filtered using cloud masks (based on the QA60 band). In addition, atmospheric corrections were applied through the built-in functions where Bands 2 (Blue), 3 (Green), and 4 (Red) were separately downloaded each tile. These bands

were therefore composited into RGB (natural color) images utilizing Python, accompanying each Sentinel-2 tile processed individually to maintain spatial integrity. Successive the image preparation, these images were imported inside the image-annotator Tool up to manually annotate preparation masks, enabling the creation of marked datasets for supervised learning. That workflow ensured effective data treatment and correct annotation, facilitating the incident and preparation of the DL- CNN (U-Net) located model.

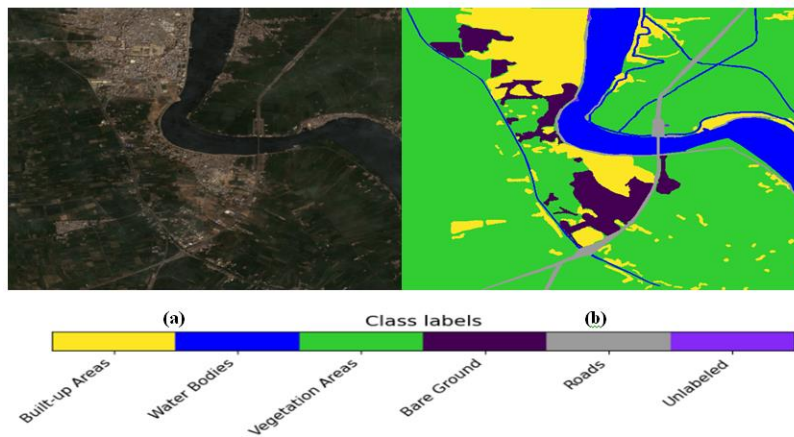


FIGURE 4 Sampling of the satellite images (a) for the study area accompanying Original masked image (b)

3.2 Model development and evaluation:

To do acceptable segmentation of Sentinel-2 satellite imagery, a CNNs based on the U-Net architecture was achieved. The preprocessed RGB images and their matching manually described masks were divide into training (70%), validation (15%), and test (15%) sets. To upgrade model generalization and strength, data augmentation methods containing flipping, turn, and shine adjustment were handled. This model was trained using a explicit cross-entropy loss function in addition to the Adam optimization algorithm. To improve generalization in addition avoid overfitting, early stopping in addition dynamic learning rate adjustments were included during the training phase. Goodfellow et al. [21]

3.2.1 CNNs and U-Net Architecture for Satellite Image Analysis

U-Net is a form of Convolutional Neural Network (CNN) that is optimized for image segmentation. Yele.et al. [24] Convolutional Neural Networks (CNNs) are effective class of deep learning models established in computer vision projects due to their skill to automatically extract spatial features from images. Upadhyay et al. [25]. In remote sensing, CNNs play an important role in defining satellite imagery, contribution high accuracy in classification, object

detection, and segmentation tasks. He et al. [26]. Among differing CNN-located models, the U-Net architecture is prominent all at once of the most active for semantic segmentation. Basically, developed for biomedical image, U-Net has been favorably used to remote sensing on account of its encoder-decoder construction and skip connections that maintain spatial analyses Ronneberger et al. [16].

3.2.2 Parameters and Hyperparameters in Deep Learning Models:

In DL models, **parameters** relate to the values automatically learned during training, for example convolutional layer weights and biases that are updated During backpropagation process. **Goodfellow et al. (2016)**. In contrast, hyperparameters are user-defined configurations set superior to training that rule the learning process itself, containing the learning rate, batch size, and dropout rate. In this study, hyperparameters were orderly optimized as itemized in Table 2, shared methods like fixed random seeding and input normalization (Min Max Scaler) to establish reproducibility. Notably, absolute hyperparameters (such as, the number of convolutional filters) are design-dependent, while possible choice (like, the loss function) generally management the optimization process

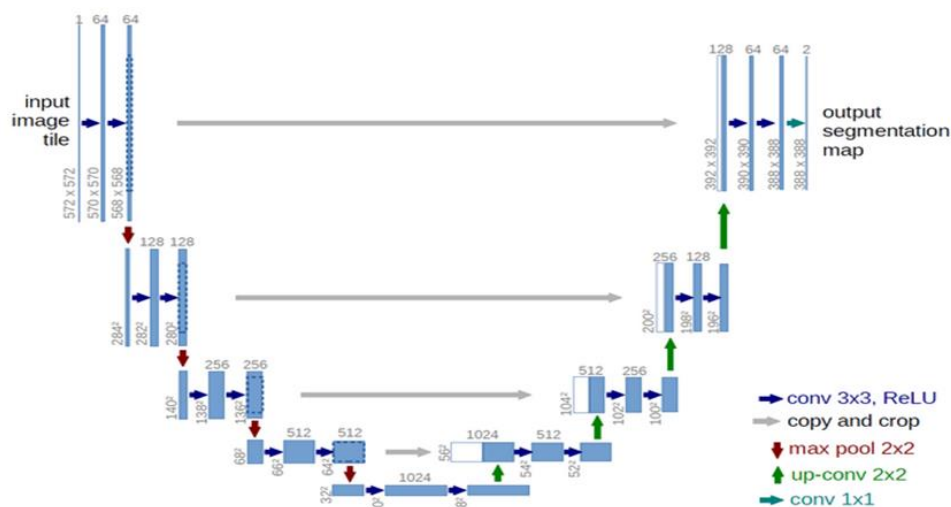


FIGURE 5 U-Net architecture adapted for semantic segmentation of Sentinel-2 satellite imagery, including encoder, decoder, skip connections, and output layer. Ronneberger et al. [16].

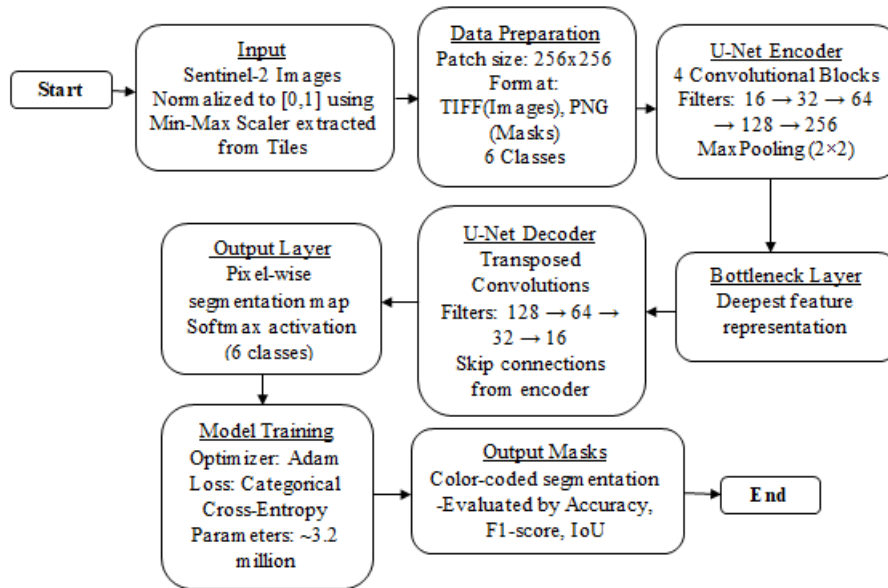


Figure 6. Workflow of the U-Net-Based Semantic Segmentation Process

TABLE2 Hyperparameter of the model

Hyperparameter	Value/Range
Image Patch Size	256×256
Image Channels	3 (RGB)
Train Size	70%
Validation Size	15%
Test Size	15%
Random State	100
Batch Size	16
Optimizer	Adam ($\beta_1=0.9$, $\beta_2=0.999$)
Learning Rate	0.001
Loss Function	Categorical Cross-Entropy
Number of Classes	6
Activation	Softmax

3.2.3 Comprehensive Evaluation of the Satellite Image Segmentation Model:

The performance of the DL model for satellite image segmentation was evaluated using standard metrics, including Precision, Recall, F1-score, Accuracy, and Jaccard coefficient that named in AI Intersection over Union (IoU). Minaee et al. [27]

The model was trained and tested using Google Colab Pro with a Tesla T4 GPU, achieving efficient resource utilization: 53.3% of system RAM, 60.0% of GPU RAM, and 14.6% of disk space. The overall Precision was 87.13%, reflecting the proportion of correct positive predictions among all predicted positives. The model achieved a recall score of 85.36%, indicating its effectiveness in detecting true positive instances within the dataset. The F1-score, which balances precision and recall, was 86.64%, and the IoU achieved 87.64%, indicating strong agreement between predicted and ground truth masks. The total accuracy was 94.14%. The metric evaluates the spatial agreement between the predicted segmentation output and the ground truth (GT) mask, reflecting the proportion of their intersection relative

to their union. It serves as an indicator of the model's effectiveness in accurately delineating target regions the following equations were used to calculate the metrics.

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + FP + FN + TN)} \quad \text{eq (1)}$$

$$\text{Precision} = \frac{TP}{(TP + FP)} \quad \text{eq (2)}$$

$$\text{Recall} = \frac{TP}{(TP + FN)} \quad \text{eq (3)}$$

$$\text{F1-score} = 2 \times \frac{(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \quad \text{eq (4)}$$

$$\text{IoU} = \frac{TP}{(TP + FP + FN)} \quad \text{eq (5)}$$

Here, TP indicates correctly identified positive cases, FP refers to incorrectly classified negatives, and FN represents missed positive instances. The evaluation results confirmed strong classification accuracy across most categories, with Water Boundaries (F1-score: 94.04 %) and Vegetation region (F1-score: 96.18%) achieving particularly high scores. In contrast, Road detection yielded a comparatively lower F1-score (60.00%), indicating potential room for improvement through model refinement for this specific class.

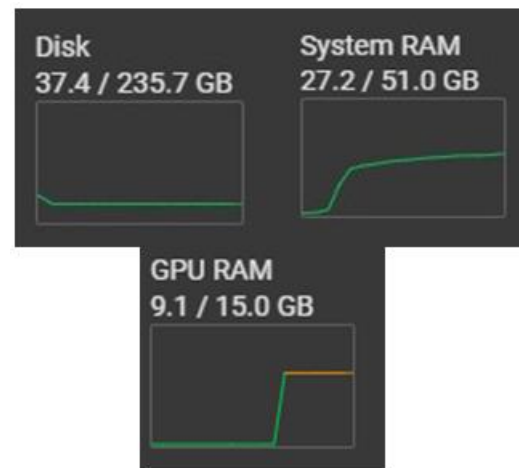


FIGURE 4 Computer Resources Used

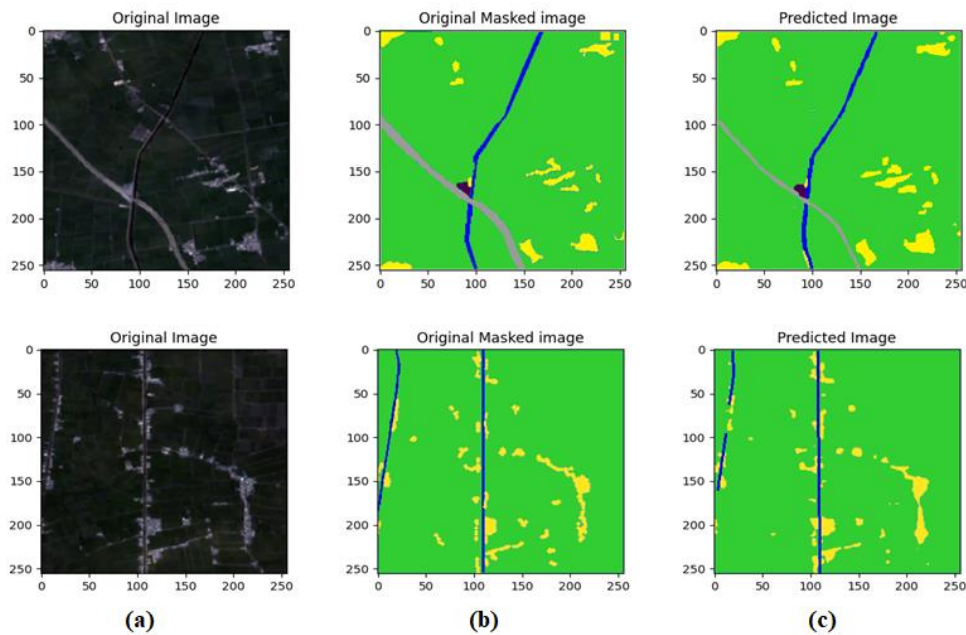


FIGURE 5 A sample of the results during the test phase

4. RESULTS ANALYSIS AND INTERPRETATION

4.1 CNN U-Net Model Performance:

CNNs established the U-Net architecture was grown and trained for the semantic segmentation (Classification) of satellite images. The model was evaluated utilizing several performance metrics containing overall accuracy, precision, recall, F1-score, and (IoU). The model reached an overall accuracy of 94.14%, signifying its powerful capability wrongly classifying the majority of pixels in the test dataset the class-wise appearance showed variation between the land cover types. The vegetation area class obtained the best classification accuracy of 97.18%, followed by water bodies accompanying 93.07%, and bare ground accompanying 95.01%. The built-up area class is well classified, with an accuracy of 81.41%. The model explained strong overall performance accompanying forceful test scores: 94.14% accuracy, 85.63% recall, an 86.64 F1-score, and 87.13% IoU. However, the road class demonstrated significantly lower accuracy at just 65.03%. This underperformance is probably attributable to two key challenges: the class's small-minded linear features and its spectral characteristics that approximately resemble adjacent sides in the imagery. These results express the model's powerful and balanced ability to discover and correctly segment the various land cover classes, making it suitable for high-accuracy remote sensing applications.

4.2 Class- wise Accuracy Analysis:

The class-wise accuracy study determines a detailed understanding of how efficiently the CNN U-Net model classified each land cover class. This evaluation helps identify that classes the model performed right on in addition that one's may require further cultivation. Generally, classes accompanying distinct spatial or spectral features likely to yield higher classification accuracies, even though more

complex or visually similar groups pose greater challenges. A comprehensive concise of the analysis accuracy each class is presented in the table 4, offering a more transparent view of the model's depiction across distinct land cover types. If we focus on the results, we will notice a difference between the overall accuracy and the rest of the criteria. this difference is a result of class imbalance such as bare ground, vegetation and water bodies dominate the dataset leading to higher accuracy while underrepresented classes like roads and unlabeled suffer from reduced precision and recall due to their limited spatial extent and spectral similarity with adjacent features and this will also be reflected in the overall criteria.

Confusion Matrix a table used to evaluate the performance of a classification model.to further evaluate the classification performance of the U-Net model, a confusion matrix was generated. This matrix provides a detailed breakdown of the correctly and incorrectly classified pixels across all land cover classes. It highlights class-wise prediction accuracy and helps identify where the model tends to misclassify certain classes, such as roads and unlabeled areas.

TABLE 3 Overall Evaluation

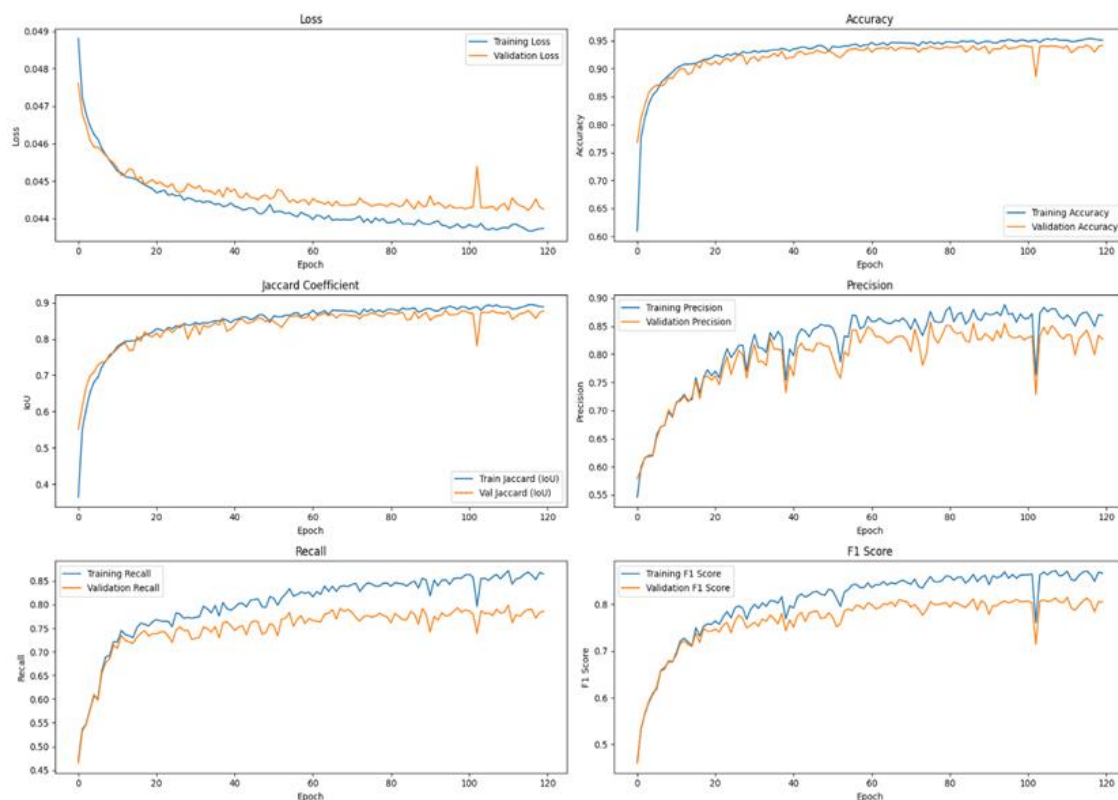
Hyperparameter	Value/Range
Metric	Value
Accuracy	94.14%
Precision	87.13%
Recall	85.63%
F1-Score	86.64%
IoU	87.13%

TABLE 4 Class-wise Evaluation

Class	Accuracy	precision	Recall	F1-Score
Built-up Areas	86.14 %	86.65 %	81.41 %	83.95 %
Water bodies	93.07 %	95.14 %	93.07 %	94.09 %
Vegetation Areas	97.18 %	95.24 %	97.18 %	96.20 %
Bare Ground	95.01 %	94.64 %	95.01 %	94.83 %
Roads	55.28 %	65.56 %	55.28 %	59.99 %
Unlabeled	49.65 %	59.10 %	49.65 %	53.97 %

Table 5. Confusion matrix showing the classification performance of the U-Net model across all land cover classes. Diagonal values represent correctly classified pixels, while off-diagonal values indicate misclassifications.

True \ Pred	Bare Ground	Water Bodies	Roads	Vegetation	Built-up	Unlabeled
Bare Ground	168,858,194	1,598,932	126,836	1,762,015	1,991,802	14,677
Water Bodies	3,909,055	74,933,249	59,621	648,118	157,265	119,632
Roads	205,139	57,815	977,065	322,507	250,490	4,116
Vegetation	815,714	791,568	223,668	65,271,102	729,050	24,544
Built-up	3,213,266	83,765	133,646	2,184,249	21,746,478	2,959
Unlabeled	81,398	381,500	18,377	116,420	29,003	1,068,125

**FIGURE 6.** Training and Validation metrics for semantic segmentation for satellite image

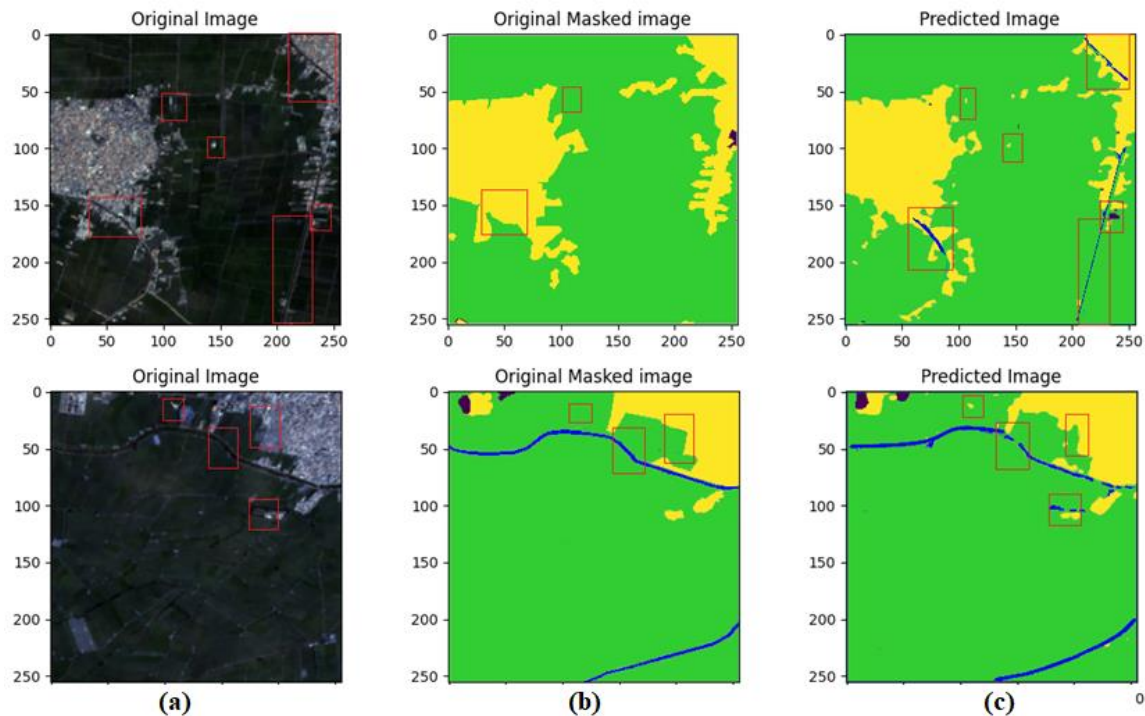


FIGURE 7. The predicted images representing 15% [Test the dataset]

The CNN U-Net model demonstrated remarkable performance, as represented in Figure (7), where it in spite outperformed the Ground Truth in certain areas, particularly in water bodies. This reflects its high accuracy in predicting water boundaries. The model also illustrated notable capability in identifying built-up regions, verifying its performance in this class. However, the model failed in some specific water boundaries, indicating challenges in accurately distinctive this class under certain conditions. Additionally, the model illustrated poor performance in detecting roads, which aligns with the quantitative results indicating low accuracy for this class likely Because of the narrow structure of roads and their spectral likeness to surrounding surfaces.

The CNN U-Net model was also used to satellite images inadequate Ground Truth data, so that evaluate its generalization capability. The model presented a high degree of accuracy in classifying Vegetation region and built-up areas, capturing the spatial patterns in addition to borders accompanying powerful clarity. Additionally, bare ground regions were well-identified, with minimum confusion with different classes in spite of these strengths, the model exhibited few disadvantages in classifying road segments, specifically due to their narrow and lengthened shapes, that often resemble neighboring classes in spectral characteristics. However, individual notable progress was the detection of a small-sized built-up area, that highlights the model's sensitivity in addition skill to capture even minor spatial features. That promises the model can be a good tool LC segmentation, even in the absence of classified data that highlights the model's sensitivity and capability to capture even small spatial features. This advises the model

maybe a reliable tool for LC segmentation, in spite of the absence of labeled data.

5. Discussion:

This study is based on previous research and illustrates the significant potential of deep learning algorithms particularly the U-Net architecture in remote sensing image classification and segmentation. This section interprets our results, discusses current limitations, and shows potential directions for future research.

5.1 Model Performance and Comparative Insights:

The U-Net model applied in this research gained a classification and segmentation accuracy of 94.14% in distinctive LC classes from satellite imagery. This result is consistent with existing research showing U-Net's success in pixel-level image classification and segmentation due to the encoder and decoder structure including skip connections of U-Net Keep spatial information through training, enabling more precise boundary mapping.

Versus to traditional artificial intelligence approaches such as SVM and RF and deep learning models especially the U-Nets provides an important benefit by automatically learning hierarchical and designed for a specific task feature from raw data, eliminating the need for manual feature engineering and enhancing model performance in complex remote sensing tasks and this is consistent with the results presented in the review phase this is consistent with the results presented in the review phase of the researches. Zhu et al. [2] , Rahman et al. [4] and Zhang et al. [5] and Abburu and Golla. [6] and Salah [7] Abaidoo et al. [8]. The results were consistent with previous results in U-Net. Ma et al. [14]; Zhang et al. [15]

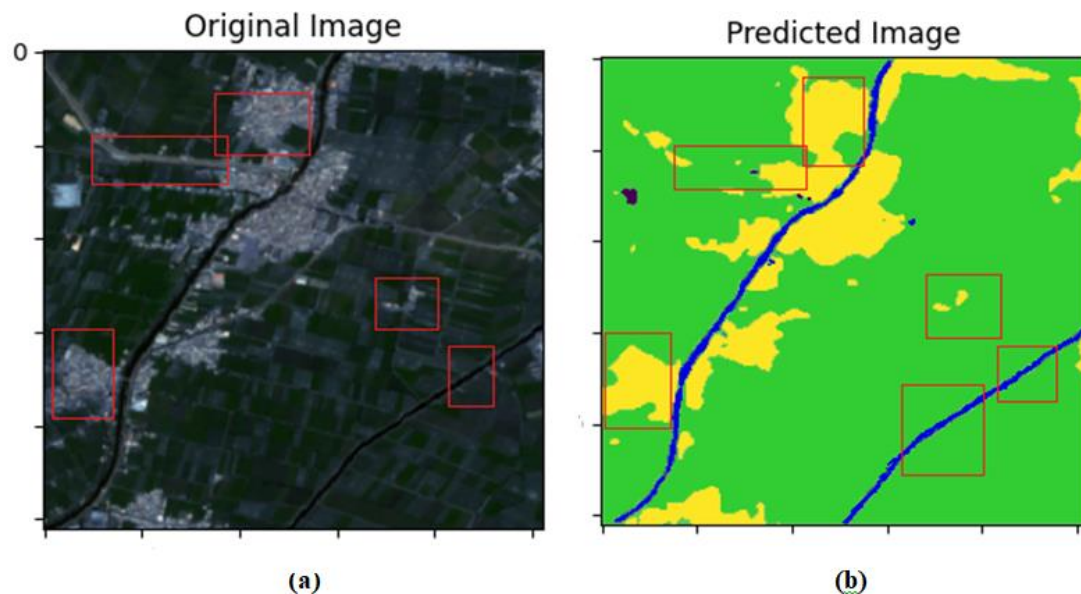


FIGURE 8. segmentation of Sentinel-2 satellite image having a spatial resolution of 10 meters

5.2 Balancing Accuracy and Computational Efficiency:

Computational Efficiency: U-Net provide an acceptable balance between accuracy and computational cost, which makes it more applicable than more recent architectures like SegVit or other transformer-based models, that could lead to higher accuracy in multi-class classification but require meaningfully more resources.

Image Resolution: The model achieves good performance using medium-resolution Sentinel-2 imagery, showing its ability to detect varying data quality. Although, Performance is expected to improve with use very high-resolution (VHR) data. Tong et al [28]

5.3 Limitations:

Geographic Bias: The model was trained on data from Egypt's Delta and valley region, which may reduce its generalization ability to other geographic areas with different environmental conditions unless retrained or improved by training on varied data.

Mask Quality Dependency: The accuracy of the model depends on masks designed manually using Image Annotator software. If the masks are not designed well, this will give completely opposite results.

6. CONCLUSION

This research highlighted the capability of a U-Net-based CNNs for performing classification and segmentation on Sentinel-2 satellite data. The model obtains high overall accuracy (94.14%) and performed well across most LC classes, particularly agricultural regions and water boundaries. These results affirm the suitability of U-Net for remote sensing tasks, especially when working with medium-resolution optical data despite its strong performance, the model showed limitations in accurately classifying narrow and spectrally ambiguous features such as roads. This highlights the potential benefit of incorporating additional data sources, such as SAR imagery, or enhancing the model architecture with attention mechanisms or multi-scale learning overall, the proposed approach offers a

reliable and scalable solution for land cover mapping, with practical implications for sustainable urban planning, water resource monitoring, and environmental management. Future work should focus on improving generalizability across diverse geographic regions and addressing class imbalance through data augmentation or customized loss functions.

6.1 Future Research Directions:

Hybrid Architectures: Performance could be enhanced by integrating U-Net with attention mechanisms or transformer-based components to capture both local and global features.

Edge Deployment Optimization: Techniques like pruning and quantization could be used to reduce model size and memory usage, facilitating deployment on low-resource platforms such as UAVs for flood monitoring.

Multisource Data Integration: Combining optical imagery with radar data (e.g., Sentinel-1 SAR) may improve classification performance in challenging conditions like cloud cover or shadowed areas

Self-Supervised Learning: Leveraging contrastive pretraining methods such as SimCLR on unlabeled data can reduce annotation costs and enhance model generalization across regions.

Acknowledgements

The author aims to express sincere gratitude to the Faculty of Engineering at Shoubra, Benha University, for providing academic support and resources throughout this research. Special thanks are extended to the academic supervisors for their valuable guidance, constructive feedback, and continuous encouragement. Appreciation is also given to the developers of open-source tools and satellite image services such as Google Earth Engine and Sentinel Hub, which played a crucial role in data acquisition and preprocessing. Finally, heartfelt thanks go to colleagues and peers who contributed through insightful discussions and technical support.

REFERENCES

- [1] Khelifi, L. and Mignotte, M. (2020). Deep learning for change detection in remote sensing images: Comprehensive review and meta-analysis. *IEEE Access*, 8: 126385–126400.
- [2] Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F. and Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine*, 5(4): 8–36
- [3] Dritsas, E. and Trigka, M. (2025). Remote sensing and geospatial analysis in the big data era: A survey. *Remote Sensing*, 17(3): 550.
- [4] Rahman, A., Abdullah, H. M., Tanzir, M. T., Hossain, M. J., Khan, B. M., Miah, M. G. and Islam, I. (2020). Performance of different machine learning algorithms on satellite image classification in rural and urban setup. *Remote Sensing Applications: Society and Environment*, 20: 100410.
- [5] Zhang, T., Su, J., Xu, Z., Luo, Y and Li, J. (2021). Urban land cover classification using Sentinel-2 data and machine learning in Beijing. *Remote Sensing Letters*, 12(4): 345–356
- [6] Abburu, S. and Golla, S. B. (2015). Satellite image classification methods and techniques: A review. *International Journal of Computer Applications*, 119(8): 20–25.
- [7] Salah, M (2017). Performance enhancement of standard fuzzy majority voting-based fusion of probabilistic classifier. *Journal of Geomatics*, 13(1): 24–33
- [8] Abaidoo, C. A., Osei Jnr, E. M., Arko-Adjei, A. and Prah, B. E. K. (2019). Monitoring the extent of reclamation of small-scale mining areas using artificial neural networks. *Heliyon*, 5(4): e01445.
- [9] Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F. and Bargellini, P. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, 120:2 5–36.
- [10] Kussul, N., Lavreniuk, M., Skakun, S. and Shelestov, A. (2017). Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, 14(5): 778–782.
- [11] Ienco, D., Gaetano, R., Dupaquier, C. and Maurel, P. (2017). Land cover classification via multitemporal spatial data by deep recurrent neural networks. *IEEE Geoscience and Remote Sensing Letters*, 14(10): 1685–1689.
- [12] Zhang, C. and Roy, D. P. (2017). Sentinel-2 multi-spectral cloud detection using a deep convolutional neural network. *Remote Sensing of Environment*, 204: 3–16.
- [13] Ayush, K., UzKent, B., Meng, C., Tanmay, K., Burke, M. and Lobell, D. (2021). Generating interpretable agricultural insights from satellite imagery with deep learning and Google Earth Engine. *Remote Sensing*, 13(8): 1552.
- [14] Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G. and Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152: 166-177.
- [15] Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J. and Atkinson, P. M. (2020). Joint deep learning for land cover and land use classification. *Remote Sensing of Environment*, 221: 173-187.
- [16] Ronneberger, O., Fischer, P. and Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015)* (pp. 234–241). Springer, Cham.
- [17] Khan, B. A., and Jung, J. W. (2024). Semantic segmentation of aerial imagery using u-net with self-attention and separable convolutions. *Applied Sciences*, 14(9), 3712.
- [18] Zhang, Z., Liu, Q. and Wang, Y. (2018). Road extraction by deep residual U-Net. *IEEE Geoscience and Remote Sensing Letters*, 15(5): 749–753.
- [19] Li, Y., Zhang, H., Zhang, Y. and Zhang, L. (2020). DeepUNet: A deep fully convolutional network for pixel-level sea-land segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13: 3032–3044.
- [20] Jiang, Y. (2025) A Comparative Study of Deep Learning-Based Semantic Segmentation Methods for High-Resolution Remote Sensing Imagery. *Advances in Computer, Signals and Systems*, 9(1): 8–13.
- [21] Hammam, A., Gaber, A., Abdelwahed, M. and Hamed, M. (2020). Geological mapping of the Central Cairo-Suez District of Egypt, using space-borne optical and radar dataset. *The Egyptian Journal of Remote Sensing and Space Science*, 23(3): 275–285.
- [22] Goodfellow, I., Bengio, Y. and Courville, A. (2016). *Deep Learning*. MIT Press. Retrieved from
- [23] Google Earth Engine. (n.d.). COPERNICUS/S2_SR_HARMONIZED – Sentinel-2 Surface Reflectance Harmonized. Google Developers. Retrieved May 22, 2025, from https://developers.google.com/earthengine/datasets/catalog/COPERNICUS_S2_SR_HARMONIZED.
- [24] Yele, V. P., Badhe, N. B., Alegavi, S., and Sedamkar, R. R. (2025). Multi Attention Convolutional Sparse Coding U-Net for Enhanced Land-Use and Land-Cover Segmentation Using Hyperspectral Images. *Sensing and Imaging*, 26(1), 1-32.
- [25] Upadhyay, A., Chandel, N. S., Singh, K. P., Chakraborty, S. K., Nandede, B. M., Kumar, M., and Elbeltagi, A. (2025). Deep learning and computer vision in plant disease detection: a comprehensive review of techniques, models, and trends in precision agriculture. *Artificial Intelligence Review*, 58(3), 1-64.
- [26] He, L. H., Zhou, Y. Z., Liu, L., Cao, W., and Ma, J. H. (2025). Research on object detection and recognition in remote sensing images based on YOLOv11. *Scientific Reports*, 15(1), 14032.
- [27] Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N. and Terzopoulos, D. (2022). Image segmentation using deep learning: A survey. *IEEE TPAMI*, 44(7): 3523-3542.
- [28] Tong, X. Y., Xia, G. S., Lu, Q., Shen, H., Li, S., You, S., & Zhang, L. (2020). Land-cover classification with high-resolution remote sensing images using transferable deep models. *Remote Sensing of Environment*, 237, 111322.