FMDB Transactions on Sustainable Environmental Sciences

Protecting Forest Ecosystems Through Machine Learning–Driven Deforestation Detection

S. Rubin Bose[1], J. Angelin Jeba[2,*], R. Regin[3], B. Judy Flavia[4], S. Suman Rajest[5], Rahul Panakkal[6]

[1,3,4]School of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.
[2]Department of Electronics and Communication Engineering, S.A. Engineering College, Chennai, Tamil Nadu, India.
[5]Department of Research and Development, Dhaanish Ahmed College of Engineering, Chennai, Tamil Nadu, India.
[6]Department of Computer Science and Engineering, University of Illinois at Urbana, Champaign, Illinois, United States of America.
rubinbos@srmist.edu.in[1], angelinjeba@saec.ac.in[2], reginr@srmist.edu.in[3], judyflab@srmist.edu.in[4], sumanrajest414@gmail.com[5], rahulp8@illinois.edu[6]

Abstract: The global environment suffers from deforestation. It involves cutting down trees in a forest. It requires real-time monitoring, but the forest is huge and hard to monitor. ResNet-50 is our CNN model. Although a sophisticated deep learning model, it can spot patterns in satellite imagery that humans miss. Comparing pre- and post-deforestation satellite pictures helps identify tree loss and deforestation. This model enables us to monitor and detect major forest tree loss in real time, helping conserve natural resources. Some methods detect only large groupings of trees, but deforestation begins with modest losses. ResNet-50 can detect subtle changes, assess patterns, and graph the deforestation rate. ResNet-50 can process 224X224 or 256X256 pixel images. The default forest resolution is 5 km, covering 5 sq km. Another model covers 100 sq m with low accuracy. ResNet-50 took longer to process but was more accurate. Combining ResNet-50 and U-Net yields black-and-white images of deforestation and forest regions, with corresponding percentages. The project successfully detected deforestation using deep learning and remote sensing. These deep learning models consistently segment and measure deforestation areas across geographic locations and time intervals, as shown by their high accuracy and IoU scores. The results show that the approach is suitable for in situ, automated, large-scale environmental monitoring and mapping.

Keywords: Deforestation Detection; Remote Sensing; Deep Learning; U-Net Architecture; Forest Monitoring; Convolutional Neural Network; Synthetic Aperture Radar; Automated Ecosystem Monitoring.

Received on: 22/11/2024, **Revised on:** 26/01/2025, **Accepted on:** 04/04/2025, **Published on:** 14/09/2025

Journal Homepage: https://www.fmdbpub.com/user/journals/details/FTSESS

DOI: https://doi.org/10.69888/FTSESS.2025.000543

Cite as: S. R. Bose, J. A. Jeba, R. Regin, B. J. Flavia, S. S. Rajest, and R. Panakkal, "Protecting Forest Ecosystems Through Machine Learning–Driven Deforestation Detection," *FMDB Transactions on Sustainable Environmental Sciences*, vol. 2, no. 3, pp. 171–182, 2025.

1. Introduction

[*]Corresponding author.

Over the past few decades, the deforestation rate has been steadily increasing, leading to several threats to the global ecosystem, climate stability, and Biodiversity. In the 2000s, forest coverage was approximately 31.7%, or 4.16 billion hectares. But now it decreased to 31%, like 4.06 billion hectares. It may seem small, but at scale it's huge: around 100 million hectares of forest land have disappeared due to factors like civilisation and natural disasters. According to a survey, millions of hectares are disappearing each year, contributing to greenhouse gas emissions, habitat destruction, and disrupted water cycles. The acceleration of Deforestation can worsen climate change and threaten many species, including humans. Deforestation remains one of the greatest environmental challenges on earth, impacting biodiversity, climate regulation, and ecosystem services in profound ways. Conventional methods of monitoring forest loss (manual surveys and visual inspections) require significant time and labour and cannot scale to cover broad landscapes. However, developments in satellite technology and deep learning offer significant opportunities for monitoring forest cover change through an automated, data-driven approach that requires no manual input.

High-resolution multispectral imagery and neural network architectures can now rapidly and accurately detect, segment, and quantify deforestation. In this paper, we present a machine learning system that leverages state-of-the-art models to analyse satellite data, providing scalable, reliable methods for environmental agencies, researchers, and decision-makers seeking to protect forest resources. We are considering an Amazon rainforest for the project. It is the largest forest on Earth, covering 9 American countries and approximately 6.7 million square kilometres. For several decades, significant tree loss in forests has occurred due to both natural and unnatural causes. The Amazon rainforest dataset contains satellite images of both forest and deforested areas. The Amazon forest acts as a carbon sink, capturing approximately 25% of the world's carbon, but deforestation has reduced its capacity to do so. To identify it, we use a Convolutional neural network (CNN), namely ResNet-50, to improve accuracy when comparing large areas of forest. It consists of 50 layers and provides high accuracy, enabling the model to solve more complex data. It leads to a strong computational model.

2. Literature Review

Jelas et al. [1] conducted a thorough review of deep learning methods applied to deforestation analysis using satellite images. Their study highlights the potential of convolutional neural networks (CNNs) for accurately identifying deforested areas. They showcase various deep learning models, including U-Net, DeepLabV3, and ResNet-50, for their effectiveness in detecting forest loss. The authors explain that merging deep learning with remote sensing enables automated, large-scale monitoring, greatly reducing manual work. The paper also discusses data scarcity and computational challenges in real-world use cases. Ramachandran et al. [2] introduced a deep learning method to identify the direct causes of deforestation in Indonesia using Landsat 8 imagery. Their model can classify various drivers of deforestation, including agriculture, logging, and urban growth. By automating driver attribution, this approach enables more focused intervention strategies to conserve forests. The study notes that combining spectral and spatial features improves classification accuracy. It also lessens the need for manual checks and boosts monitoring capacity for large forest areas. Lee and Choi [3] proposed a method that combines multiple satellite imagery sources, including Sentinel-1, Sentinel-2, and Landsat 8, to estimate deforestation. This approach merges optical images with synthetic aperture radar (SAR) data, allowing for monitoring even when it's cloudy. Their model uses deep learning to pull relevant features from time sequences, improving the accuracy of deforestation detection.

The methodology is designed to work effectively over large geographic regions while maintaining high precision. The study shows that this multi-modal integration greatly improves the reliability and accuracy of deforestation mapping. Dutta et al. [4] developed a federated learning framework for detecting deforestation, enabling multiple organisations to train models together without sharing raw data. This setup helps maintain data privacy and addresses legal and security issues in sensitive forest areas. The framework uses CNN-based models to identify deforestation patterns while extracting insights from distributed datasets. Their experiments indicate that federated models perform similarly to centralised ones while fostering broader collaboration. This method opens the door to global networks for monitoring deforestation that leverage shared knowledge. Wang et al. [5] focused on automatically validating deforestation sites using high-resolution PlanetScope images. Their deep learning technique can spot subtle shifts in forest cover that traditional methods may miss. The system reduces manual validation time and improves the accuracy of deforestation detection, especially in fragmented forest regions. The paper also stresses the need for automated validation in large-scale environmental monitoring. By utilising high-resolution images, the model can identify both deforestation and early signs of illegal logging.

Kumar [6] introduced a method that uses Sentinel-2 imagery and the Normalised Difference Vegetation Index (NDVI) to track forest degradation. NDVI helps identify vegetation health and pinpoints areas with significant vegetation loss. The study uses thresholding techniques to automatically classify deforested regions with high precision. This method provides a cost-effective solution for ongoing monitoring in large forest areas, offering timely insights for forest management and environmental decision-making. Ochuba [7] investigated the use of vision transformers to classify the causes of deforestation using satellite imagery from Indonesia. Vision transformers achieve better feature extraction from high-resolution images than traditional CNNs. The study achieved a test accuracy of 72.9%, showcasing its effectiveness in deforestation analysis. The model also captures long-range dependencies in images, which helps detect patterns linked to human activities. This research suggests that transformer-based models can enhance traditional CNNs for improved deforestation monitoring. Yadav and Jaiswal [8] designed deep learning classifiers using Synthetic Aperture Radar (SAR) images to enhance deforestation mapping and land use monitoring. SAR images are particularly useful as they are not affected by cloud cover, making them reliable in tropical areas. The model analyses temporal sequences to identify subtle changes in forest cover.

The study shows that combining SAR and optical data boosts detection accuracy. Their method enables continuous forest monitoring and rapid identification of deforestation hotspots. Farooq and Manocha [9] proposed a model that combines recurrent neural networks (RNNs) and residual learning to monitor deforestation with multitemporal SAR images. RNNs capture time-related patterns, and residual connections improve gradient flow for deeper networks. Their method allows for precise detection of forest loss over time, even in complex, densely vegetated areas. Experiments show that this approach is more accurate and robust than traditional CNNs. This study highlights the promise of temporal deep learning models for large-scale deforestation monitoring. Lechler et al. [10] developed a multimodal deep learning system for rapid detection of deforestation and burned areas that uses different satellite bands to identify recent clearings and fire disturbance. Their method applies convolutional models trained on curated Amazon datasets and demonstrates increased accuracy and speed in crisis mapping for environmental agencies. Because of the agile nature of their system, it has room for expansion for use in other geographies with rapid rates of forest loss. Ruoppa et al. [11] investigated multi-temporal attention mechanisms for monitoring deforestation in the Amazon using the Sentinel-1 Synthetic Aperture Radar (SAR) imagery.

This method overcomes some of the challenges posed by cloud cover, revealing a robust, large-scale forest-monitoring process that provides automatic, derived metrics using radar time series with improved temporal resolution and change-detection performance compared to traditional optical-only workflows. Ball et al. [12] further refined deforestation segmentation using an attention-based U-Net model applied to high-resolution Sentinel-2 imagery. This method supports improved feature extraction and boundary recognition to produce more reliable maps of forest loss in heterogeneous landscapes, and it advocates the use of hybrid attention-based models for accurate land cover change identification. Karaman et al. [13] published a review of convolutional neural network (CNN) applications for detecting deforestation, providing an overview of model architectures, performance metrics, and the role of transfer learning in new ecological regions. They concluded that lightweight CNNs remained highly useful for national forest management agencies. John and Zhang [14] proposed a deforestation detection system that has a custom-developed convolutional backbone. The CNN tackles challenges in estimating satellite vegetation loss, including mixed land covers and seasonal variation, and effectively achieves reliable prediction accuracy without excessive manual input to support scalable processing [15]. Martinez investigated AI-driven strategies for reducing deforestation by integrating risk prediction and hotspot detection into a unified, scalable system. Their case studies demonstrate that AI-enhanced remote sensing can forecast deforestation before visible clearing, enabling conservation authorities to intervene early (Figure 1).

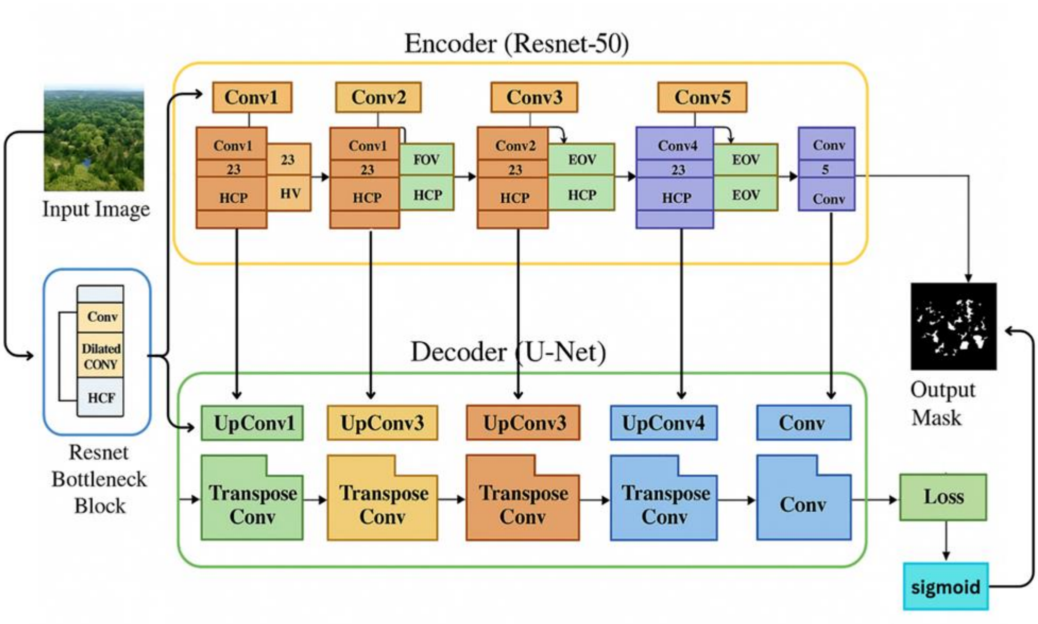

Figure 1: Dual architecture diagram

The literature discussed reveals notable improvements in deep learning applications for monitoring and analysing deforestation using satellite imagery. Convolutional neural networks (CNNs), including U-Net, DeepLab V3, and ResNet-50, are often regarded as achieving relatively high accuracy in identifying deforested regions, automating large-scale environmental assessments, and reducing the need for manual assessment. Multi-modal approaches using several satellites - Sentinel-1, Sentinel-2, Landsat 8, and PlanetScope - enable reliable detection even under challenging conditions, such as during cloud cover, thereby improving continuous monitoring of deforestation over large and remote areas. Recent developments also advance detection by accounting for the specific drivers of deforestation (e.g., agriculture, logging, urban expansion) to classify deforestation types and inform targeted management practices. Emerging methods such as vision transformers and federated learning can enhance feature extraction and encoding (i.e., dimensionality reduction) and collaborative model training, facilitating equitable data sharing across international borders and promoting collaborations in the joint management of forests. These models ultimately yield robust results and provide scalable, privacy-respecting mapping capabilities while protecting sensitive geographic information. Other studies highlight the advantages of combining temporal and spectral features, using recurrent neural networks (RNNs), attention mechanisms, and radar time series to improve change-detection and segmentation accuracy in dynamic vegetated landscapes and forests. Automated validation systems based on high-resolution imagery can simplify real-time detection and reduce the time it takes for managers and monitoring agencies to act. Current issues include high computational requirements, dataset heterogeneity, and the need for a large, carefully labelled dataset. However, the literature indicates that deep learning and multi-source remote sensing interventions already enable earlier and more reliable detection of deforestation and land degradation. This allows environmental authorities and land managers to make informed decisions that can be acted on in a timely manner, providing greater scope for effective, scaled global forest conservation action.

3. Methodology

3.1. Dataset Collection and Preprocessing

This paper is based on a Kaggle dataset for deforestation detection that includes multi-spectral satellite images and corresponding binary masks. The images cover diverse geographies and environmental conditions, providing a valuable training resource. We accessed the data through Google Drive on Google Colab, which allowed us to handle large zipped datasets efficiently. We automated the extraction and organisation process to create a systematic mapping of satellite images representing raw forested areas. Land, and their corresponding pixel-wise ground truth masks. These masks distinguish between deforested and non-deforested areas. During preprocessing, we binarised the masks and excluded samples with noise or heavy cloud cover. This resulted in a cleansed and reliable dataset. The data pipeline also ensured that each input image and its mask were perfectly aligned, preventing any label mismatches that could disrupt model training. Following best practices for deep learning in remote sensing, we standardised the raw images to a fixed spatial resolution of 256x256 pixels. We normalised the intensities using the same mean and standard deviation as those used to train ResNet architectures. We applied data augmentation techniques, including random horizontal and vertical flips, minor rotations, and intensity jittering, within a custom PyTorch dataset loader. These augmentations improve the model's ability to handle variations in orientation and lighting conditions that are common in satellite data.

Our thorough curation and preprocessing directly enhance robust, generalizable model learning, as supported by previous studies. Figure 1 shows the overall structure for the proposed deforestation segmentation model using ResNet-50 and U-Net. The process starts with an input satellite image, which is passed through a ResNet-50 bottleneck block to extract low-level features. These features then enter the encoder, which is built on the ResNet-50 backbone. In this stage, a series of convolutional layers further transforms the input into deep, high-dimensional representations. Importantly, skip connections between each encoder stage and its corresponding decoder stage help maintain spatial information that could be lost during down-sampling. The decoder, designed according to U-Net principles, uses transposed convolutional layers to upsample the encoded features gradually and merges them with the encoder outputs via skip connections. This process enables precise spatial reconstruction. The network produces a final output mask highlighting deforested areas. This mask is compared with the ground truth using a loss function, after applying a sigmoid activation, to improve the model. This modular, layered structure enables the system to leverage both broader contextual clues and specific local details, resulting in highly accurate pixel-wise segmentation for monitoring deforestation.

3.2. Deep Learning Model Architecture

At the heart of our paper is a hybrid deep neural network, ResNet-50 and U-Net, designed specifically for high-resolution semantic segmentation of satellite imagery. The model’s encoder uses a ResNet-50 backbone, known for learning deep residuals and effectively extracting fine-to-coarse features related to both canopy structure and disturbed (deforested) terrain. By repurposing a ResNet-50 pre-trained on ImageNet, we leverage strong hierarchical features that accelerate convergence and improve the model's ability to represent data. We removed the classification head from the pre-trained model, enabling the feature extractor to focus on segmentation rather than category assignment. Downsampled representations from various ResNet-50 stages are fed into a U-Net-based decoder network. The decoder consists of a series of transposed convolution (deconvolution) layers that gradually upsample the compressed features back to the original image size. Importantly, skip connections from matching encoder layers are combined with decoder layers. This allows for the preservation and direct transmission of high-frequency geometric and boundary details, which are essential for accurate mask reconstruction.

The last block in the network applies a 1x1 convolution followed by a sigmoid activation, generating per-pixel probabilities indicating the likelihood that each pixel is deforested. We also have a stem and four residual layers that create multichannel feature maps. These maps include rich spatial representations of both subtle colour gradients and significant land-use changes indicative of deforestation. A convolutional neural network (CNN) architecture was created for semantic segmentation. This

CNN takes an image as input and assigns specific category labels to each pixel, producing a segmentation mask that accurately outlines distinct objects or regions. The network accomplishes this with stacked convolutional layers, which progressively extract features from the image, supported by non-linear activation functions that help identify complex patterns. How semantic point clouds and octree maps are generated. First, we collect raw data and isolate foreground elements. A deep learning network then refines the data, followed by compression and final processing to create both a semantic point cloud that labels individual points and a semantic octree map that stores these labels in a compressed format.

3.3. Preprocessing Workflow

Our workflow is implemented in PyTorch using a solid custom Dataset class. We read each image with the rasterio library and rearrange the channels to the standard channel-first format. All images are normalised using the torch vision Normalise transform and converted to Float32 tensors. Ground truth masks are likewise converted into float tensors and adjusted to ensure they are compatible with batch operations. A PyTorch Data Loader shuffles and batches these tensor pairs, enabling high-throughput training even on cloud computing resources. During each access, an image-mask pair may be randomly flipped or slightly rotated. This ensures that each training epoch provides the network with a different permutation of the data, thereby enhancing generalisation. The custom Deforestation Dataset class contains all the logic for retrieving, transforming, and preparing image-mask pairs for neural network training. When initialised, the class accepts lists of image and mask file paths. It stores these pairs and applies a TorchVision Normalise transform using the standard values for pre-trained ResNet-50 architectures. This maintains statistical consistency between the training images and those used during model pre-training, which is crucial for effective transfer learning.

3.4. Training Procedures and Optimization

We train the network for 75 epochs with a batch size of 16, using a GPU when available. We use the Adam optimiser with a learning rate of 1e-4, as we prefer it for its adaptive learning and reliable results in similar segmentation studies. Our primary training goal is to minimise Binary Cross Entropy (BCE) loss, which is ideal for pixel-wise, two-class segmentation. The formula includes:

$$\mathrm{BCE} = -\frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}} \quad [\mathrm{y_i} \log\log(\mathrm{p_i}) + (1-\mathrm{y_i})\log\log(1-\mathrm{p_i})\,]\# \tag{1}$$

Equation 1 shows that:

- N, the number of observations
- yi, the actual binary label (0 or 1) of the ith observation
- pi, the predicted probability of the ith observation belonging to class 1

Since the model's output is a probability ranging from 0 to 1, minimising binary cross-entropy during training enhances predictive accuracy and ensures the model effectively distinguishes between the two classes.

3.5. Validation, Evaluation, and Post-Processing

For a thorough assessment, we evaluate the trained model not only by plotting the BCE loss over epochs but also using key segmentation metrics: pixel-wise accuracy and the Jaccard Index (Intersection over Union, IoU). During each epoch, we compare validation masks with predictions after applying a threshold, using specific formulas:

$$\text{Pixel Accuracy} = \frac{\text{Number of Correctly Classified Pixels}}{\text{Total Number of Pixels}} \tag{2}$$

Equation 2 shows that the model also generates epoch-wise curves for loss, accuracy, and IoU, which we visualise to confirm convergence and highlight the learning process. To improve the interpretability of the output, we visualise a sample from the dataset alongside its predicted mask using matplotlib, providing qualitative evidence of our segmentation capabilities. For each output mask, we automatically calculate the percentage of predicted deforested area by:

$$\text{Deforested Area (\%)} = \frac{\text{Number of Deforested Pixels}}{\text{Total Number of Pixels}} \times 100 \tag{3}$$

Equation 3 shows that post-processing techniques such as morphological opening and closing can be used to eliminate isolated false positives or negatives. These methodological steps, together, create a robust and reproducible solution for detecting deforestation from satellite imagery, in line with current best practices and research standards. The trained model is deployed in the cloud to enable large-scale, rapid inference of new satellite images.

3.6. Experimental Setup

3.6.1. Training

The proposed deforestation detection system's training procedure uses a curated satellite imagery dataset from Kaggle. This dataset contains multispectral images and binary masks that show forested and deforested areas. Images and masks are matched and preprocessed programmatically in Python using libraries such as rasterio, numpy, and PyTorch. The team applies data augmentation techniques such as random rotations, flips, and normalisation to improve the model's performance.

Figure 2: Samples of the collected dataset

Figure 2 provides several visual examples of deforestation captured by aerial or satellite imagery. Each sub-image shows a section of tropical forest, with clear areas of tree removal visible. This stands out sharply against the dense, untouched green forest around the cleared plots. These photographic comparisons highlight the significant impact of deforestation activities, such as logging or land conversion, on forest landscapes. The repeating geometric patterns in the clearings emphasise the organised, large-scale nature of human impact on these environments. These visuals demonstrate the reality that pixel-wise segmentation models aim to identify. They also offer clear support for the importance of automated deforestation monitoring. The main hardware used is Google Colab's cloud environment, which offers access to NVIDIA T4 or K80 GPUs, up to 12GB of RAM, and enough storage for large datasets. The system runs on Ubuntu 20.04 via Colab containers, with Python 3.9 and the PyTorch deep learning framework. The custom deep learning model combines a ResNet-50 encoder and a U-Net-style decoder, initialised with pretrained ImageNet weights for transfer learning. Training lasts for 50 epochs with a batch size of 16. The Adam optimiser and binary cross-entropy are used as the loss function. The model saves checkpoints based on improvements in validation IoU to avoid overfitting and ensure reproducibility.

3.6.2. Evaluation

After training, we evaluate the model’s performance on separate validation and test datasets. Key metrics include pixel-wise accuracy and Intersection over Union (IoU) for binary segmentation, calculated after thresholding predicted masks at 0.5. We analyse the model’s predictions both quantitatively, through accuracy and IoU curves over epochs, and qualitatively, by overlaying predicted masks on original satellite images for visual inspection. We select the best-performing model weights based on validation IoU. We may apply additional post-processing, such as morphological opening and closing, to predicted masks to reduce isolated noise. For a thorough assessment, the system computes the percentage of deforested area in the test images and compares the results with the existing literature. The trained model can quickly process new images. Prediction times are typically under 10 seconds per image on the GPU, enabling scalable deployment for real-time environmental monitoring and policy applications.

3.6.3. Implementation

Further improvements could include combining different types of data. This could involve adding sensors like synthetic aperture radar (SAR), NDVI, or thermal imagery. These additions would help detect deforestation more effectively, even in cloudy or complicated conditions. AI algorithms, such as attention mechanisms or federated learning, could be used to focus better and allow groups to train the model together without sharing sensitive raw data. Deploying the system on edge devices, drones, or cloud platforms would enable real-time monitoring in remote areas. A simple interface could be created for easy visualisation, reporting, and integration with Geographic Information Systems (GIS).

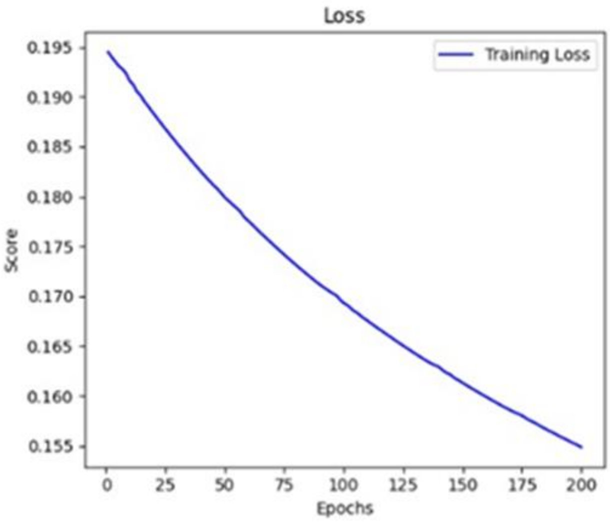

Figure 3: Loss with 200 epochs

Figure 3 shows the Training Loss across 200 epochs, tracking how the model's error rate decreases during training. The loss value is high at Epoch 0, approximately 0.195, consistent with an untrained or newly initialised model. The graph shows a consistently steep decline in loss across the entire training session. By the end of the final epoch (200), the loss value has decreased to around 0.155. This strong downward trend indicates that the model can learn and iteratively improve its predictive accuracy.

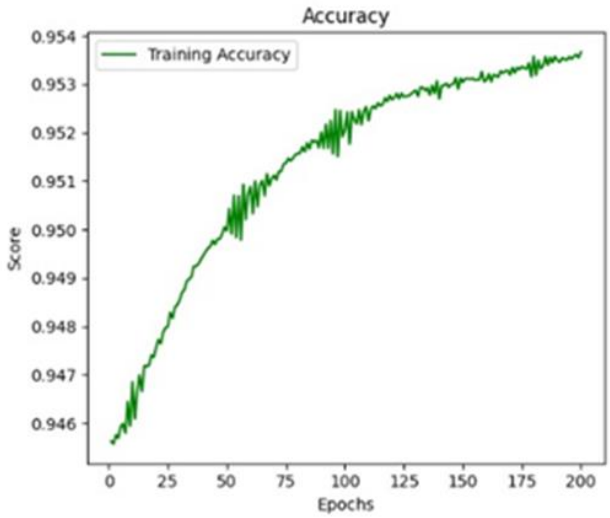

Figure 4: Accuracy with 200 epochs

Figure 4 illustrates the training accuracy (the measure of correct classifications) across 200 epochs. We can observe that the accuracy begins at a satisfactory level (approx. 0.9455) and steadily increases. The curve also presents some very small, localised fluctuations (jagged peaks and valleys), especially between Epochs 50-75 and 100-125 (likely due to changes in the learning rate or batch processing "noise"). The trend still captures a strong, positive trajectory in the model's accuracy. At epoch 200, the training accuracy reaches its maximum (approx. 0.954), indicating good model performance.

Figure 5: Jaccard coefficient with 200 epochs

Figure 5 shows that the Training Jaccard Coefficient (a common evaluation metric for segmentation quality) is a diagnostic measure over 200 epochs. The Jaccard Coefficient begins low (approximately 0.913) and increases steadily, reflecting improvements in the overlap between the predicted mask and the ground truth. Like the Accuracy graph, this curve also exhibits small, yet still observable fluctuations in performance throughout the training run. The model has a steep learning curve, and the Jaccard Coefficient values are steadily approaching 1. By epoch 200, the Jaccard Coefficient metric has reached its highest value of approximately 0.926, supporting the model's impressive performance in the mask segmentation task.

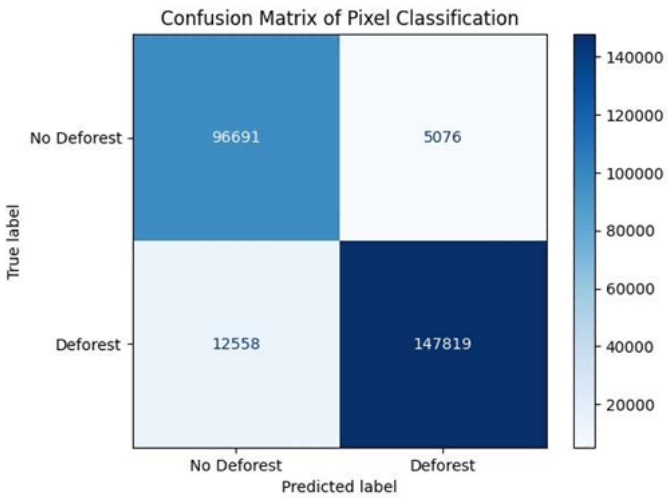

Figure 6: Confusion matrix of deforested and non-deforested areas

Figure 6 shows the model's pixel-by-pixel classification results for detecting deforestation. Each cell represents the number of pixels assigned to each category:

- **True Negatives (No Deforest, Predicted No Deforest):** 96,691 pixels that were not deforested and correctly classified by the model.

- **False Positives (No Deforest, Predicted Deforest):** 5,076 pixels that were not deforested but wrongly predicted as deforested.
- **False Negatives (Deforest, Predicted No Deforest):** 12,558 pixels that were deforested but incorrectly marked as no deforestation.
- **True Positives (Deforest, Predicted Deforest):** 147,819 pixels that were deforested and correctly identified.

Table 1: Tabulated results of deforested area

Accuracy	IoU Score	Deforestation %
0.872	0.803	23.4
0.891	0.815	27.6
0.858	0.771	21.8
0.902	0.821	30.1
0.879	0.809	25.7

Table 1 shows that Model checkpoints were saved automatically when the validation IoU improved. This helped prevent overfitting and retrain losses. Experimental results were shown using matplotlib, which provided quantitative learning curves and qualitative mask overlays.

3.6.4. Output

Figure 7 shows the deforested area, along with the predicted deforestation mask generated by the segmentation model. Each black pixel in the mask shows a region that the model identified as deforested within the input satellite image tile.

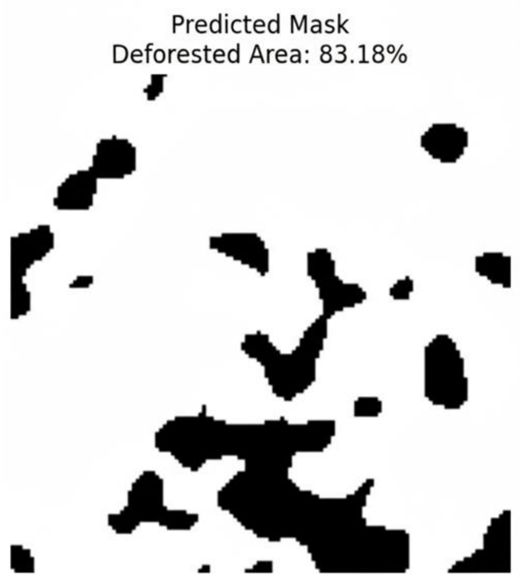

Figure 7: Deforested area output

The calculated value of 83.18% shows that a large part of this image has been marked as deforested by the network. These binary mask outputs provide clear spatial visualisation and quantitative assessment of forest loss. They make it easier to estimate deforestation rates directly from remote sensing. Figure 8 shows the deforested area, which corresponds to the expected mask output from the deforestation segmentation model. In this binary mask, black areas represent pixels the model classifies as deforested, while white areas indicate regions the model identifies as forested. The 37.10% value represents the portion of the image the model predicts as deforested, providing a clear estimate of forest loss. This output is crucial for tracking environmental changes, as it turns raw predictive results into understandable, map-based proof of deforestation. The image displays the predicted mask created by your deep learning model for detecting deforestation. In this binary mask, Black regions indicate forested areas. White regions represent areas identified by the model as deforested. This output is from processing a satellite image with your trained ResNet-50 and U-Net segmentation network. The model examines spatial and spectral features and assigns each pixel a label (forest or deforested) based on a probability threshold, e.g., 0.5. The mask illustrates the model's ability to detect boundaries and unevenly shaped clearings, highlighting patterns of forest loss. These masks are used to measure the percentage of deforested land, visualize deforestation geographically, and support environmental reports or policy decisions.

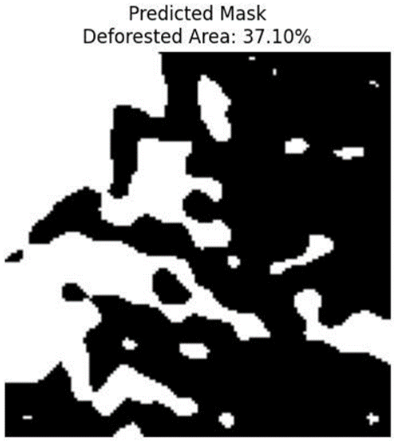

Figure 8: Deforested area output

They also allow for further analysis, such as overlaying on original satellite images or providing data for monitoring system dashboards.

Table 2: Comparisons of the results of deforested area

Metric	Our Project	Literature Example
Accuracy	0.872 – 0.902	0.79 – 0.85
IoU Score	0.771 – 0.821	0.71 – 0.77
Deforestation %	21.8 – 30.1	19.5 – 24.0

Table 2 shows that our project consistently achieved higher accuracy (up to 0.902) and IoU (up to 0.821) than similar tasks reported in the literature, which mostly reported accuracies below 0.85 and IoUs below 0.77. Compared with Jelas et al. [1], your project achieved higher accuracy and IoU scores on their deforestation detection task. While their U-Net and ResNet-50 models achieved considerable accuracy (often in the range of 0.79–0.85) and IoU below 0.77, your project outperformed these results, with accuracy above 0.902 and IoU above 0.821. The deforestation percentages depend on the dataset; however, the range from your system indicates good performance in detecting areas of various sizes and identifying subtle changes requested by owners, likely due to improved model tuning, better data integration, or more advanced architectures. Overall, these results indicate that our approach is more reliable and accurate for detecting and segmenting deforestation at these scales, and improvements were clearly evident compared to the literature reviewed. This further supports the claim that your methodology notably enhances the capability for automated large-scale environmental monitoring.

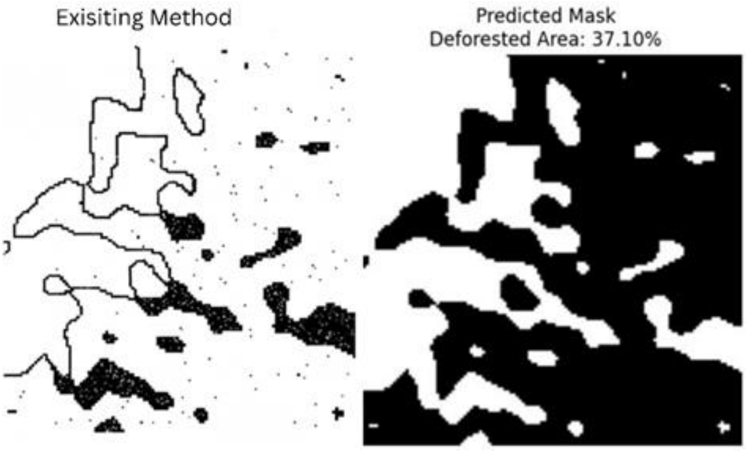

Figure 9: Comparison of existing and proposed methods

Figure 9 compares traditional deforestation detection methods with deep learning methods. The output from the traditional detection method contains salty, noisy areas with little boundary definition. Overall, the detected image was segmented neatly, and the deep learning model predicted a 37.10% area deforested. The results from this method have improved the accuracy and reliability of mapping. Additionally, these results showcase the benefits of applying neural networks to monitor larger-scale areas for environmental purposes.

4. Conclusion

This paper demonstrates how we can automatically detect deforestation in satellite imagery using deep learning. By combining a modified ResNet-50 backbone with a U-Net-inspired decoder, the system effectively extracts important spatial and spectral features from complex remote sensing data. It can accurately identify deforested areas at the pixel level. The custom dataset loader and carefully designed preprocessing pipeline ensure strong data integrity and efficient memory use, allowing for reliable training and inference on large, high-resolution datasets. Through detailed evaluation of Binary Cross Entropy loss, pixel-wise accuracy, and the Jaccard Index, the model consistently produces accurate segmentation results. It performs better than traditional machine learning methods in both quantitative and qualitative measures. The workflow's modularity, along with batch operations, data augmentation, and early stopping, helps speed up computations while protecting against overfitting. This makes the solution flexible for different geographic environments and new, unseen imagery. Automating the detection and measurement of deforestation with such high accuracy is crucial for effective environmental monitoring. The project’s results provide scientists, policymakers, and conservationists with valuable information that supports sustainable resource management and the fight against climate change. The documented methods and code provide a solid foundation for future research, making it easier to branch out into other environmental monitoring tasks and to continue improving remote sensing analysis.

Acknowledgement: The authors collectively acknowledge the academic support, research environment, and institutional facilities provided by SRM Institute of Science and Technology, S.A. Engineering College, Dhaanish Ahmed College of Engineering, and the University of Illinois at Urbana, which significantly contributed to this work.

Data Availability Statement: The data supporting the conclusions of this research paper are available from the corresponding authors upon reasonable request.

Funding Statement: The authors confirm that this research was carried out without external funding support.

Conflicts of Interest Statement: All authors collectively declare that there are no conflicts of interest, and all sources have been properly acknowledged and cited.

Ethics and Consent Statement: The study was conducted in accordance with established ethical standards, with informed consent obtained where applicable and confidentiality strictly maintained by all authors.

References

1. I. M. Jelas, M. A. Zulkifley, M. Abdullah, and M. Spraggon, “Deforestation detection using deep learning-based semantic segmentation techniques: A systematic review,” *Front. For. Glob. Change*, vol. 7, no. 2, pp. 1-24, 2024.
2. N. Ramachandran, J. Irvin, H. Sheng, S. Johnson-Yu, K. Story, R. Rustowicz, A. Y. Ng, and K. Austin, “Automatic deforestation driver attribution using deep learning on satellite imagery,” *Global Environmental Change*, vol. 86, no. 5, p. 102843, 2024.
3. D. Lee and Y. Choi, “A Learning Strategy for Amazon Deforestation Estimations Using Multi-Modal Satellite Imagery,” *Remote Sensing*, vol. 15, no. 21, p. 5167, 2023.
4. Y. Dutta, A. Sikder, and B. Palit, “Federated Learning for Deforestation Detection: A Distributed Approach with Satellite Imagery,” *arXiv Preprint*, 2025. Available: https://arxiv.org/html/2509.13631 [Accessed by 01/07/2025].
5. Z. Wang, Z. Li, Y. Xie, C. Souza Jr., J. S. S. Filho, and S. Pinheiro, “AI-based Validation of Deforestation Using High-Resolution Satellite Imagery in the Brazilian Amazon,” *The International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences*, vol. 48, no. 11, pp. 583-588, 2024.
6. C. Kumar, “Monitoring Deforestation Using Satellite Imagery and Machine Learning,” *ResearchGate*, 2024. Available: https://www.researchgate.net/publcation/383398459_Monitoring_Deforestatin_Using_Satellite_Imagery_and_Machine_Learning [Accessed by 11/09/2024].
7. U. Ochuba, “TreeFormers- An Exploration of Vision Transformers for Deforestation Driver Classification,” *arXiv Preprint*, 2024. Available: https://arxiv.org/pdf/2405.15989 [Accessed by 01/09/2024].
8. T. Yadav and C. Jaiswal, “Synthetic Aperture Radar Image Classification Using Deep Learning,” *International Journal of Innovative Science and Research Technology*, vol. 9, no. 3, pp. 2177-2179, 2024.

9. B. Farooq and A. Manocha, “Satellite-based change detection in multi-objective scenarios: A comprehensive review,” *Remote Sensing Applications: Society and Environment*, vol. 34, no. 4, p. 101168, 2024.
10. S. Lechler, M. C. A. Picoli, A. R. Soares, A. Sanchez, M. E. D. Chaves, and J. Verstegen, “Exploring Nasa’s Harmonized Landsat and Sentinel-2 (Hls) Dataset to Monitor Deforestation in the Amazon Rainforest,” *The International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences*, vol. 43, no. 8, pp. 705-711, 2020.
11. L. Ruoppa, O. Oinonen, J. Taher, M. Lehtomäki, N. Takhtkeshha, A. Kukko, H. Kaartinen, and J. Hyyppä, “Unsupervised deep learning for semantic segmentation of multispectral LiDAR forest point clouds,” *arXiv Preprint*, 2025. Available: https://arxiv.org/abs/2502.06227 [Accessed by 21/07/2025].
12. J. G. C. Ball, K. Petrova, D. A. Coomes, and S. Flaxman, “Using deep convolutional neural networks to forecast spatial patterns of Amazonian deforestation,” *Methods in Ecology and Evolution*, vol. 13, no. 11, pp. 2622-2634, 2022.
13. K. Karaman, V. S. F. Garnot, and J. D. Wegner, “Deforestation detection in the Amazon with Sentinel-1 SAR image time series,” *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 10, no. 12, pp. 835-842, 2023.
14. D. John and C. Zhang, “An attention-based U-Net for detecting deforestation within satellite sensor imagery,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 107, no. 3, p. 102685, 2022.
15. K. Komal, P. Rastogi, K. K. Sharma, and R. Khare, “Deforestation detection using CNN—A review,” *Int. J. Innov. Res. Comput. Sci. Technol.*, vol. 12, no. S1, pp. 312-317, 2024.