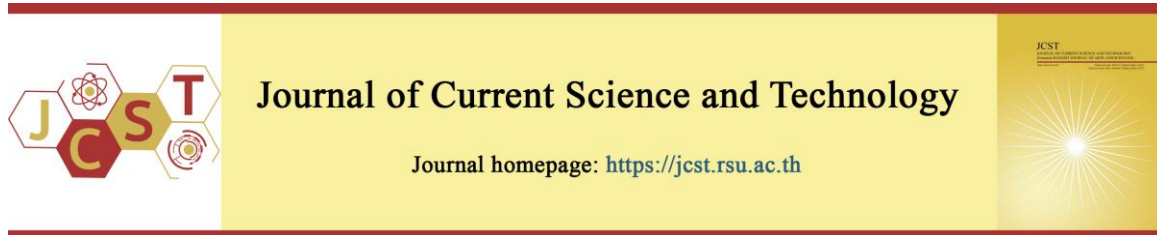


Cite this article: Shobha, P. (2026). Protecting forests through intelligent edge computing: CBAM-UNet for real-time deforestation detection. *Journal of Current Science and Technology*, 16(4), Article 207. <https://doi.org/10.59796/jcst.V16N4.2026.207>



Protecting Forests Through Intelligent Edge Computing: CBAM-UNet for Real-Time Deforestation Detection

Shobha, P, 

Department of Computer Science and Engineering, Nitte (Deemed to be University), Nitte Meenakshi Institute of Technology (NMIT), Bengaluru, Karnataka, India.

E-mail: shobha.p@nmit.ac.in

Received 29 October 2025; Revised 30 March 2026; Accepted 30 July 2026; Published 15 September 2026

Abstract

The accelerating loss and degradation of forest ecosystems necessitate accurate and efficient monitoring approaches. This study presents a lightweight, attention-enhanced semantic segmentation framework for binary forest canopy segmentation using satellite imagery. The proposed model integrates a U-Net architecture with a MobileNetV2 encoder and a Convolutional Block Attention Module (CBAM) to improve feature representation while maintaining computational efficiency. The model is evaluated on the Bragagnolo Amazon Rainforest dataset for forest-versus-non-forest segmentation. The CBAM-enhanced model achieved a mean Intersection over Union (mIoU) of 0.6339 compared with 0.6282 for the baseline U-Net–MobileNetV2, while maintaining a comparable Dice score (0.6141 vs. 0.6187), indicating more consistent class-wise segmentation. The exported ONNX model sustained approximately 5.9 frames per second on an NVIDIA Jetson Nano while consuming under 80 MB of memory, confirming its feasibility for resource-constrained edge deployment. This work focuses strictly on binary canopy segmentation and provides insights into the trade-offs between accuracy and efficiency in lightweight deep-learning models for remote-sensing applications.

Keywords: attention, CBAM, deforestation, deep learning, edge computing, MobileNetV2, remote sensing, satellite imagery, semantic segmentation, UNet.

1. Introduction

Accurate forest canopy segmentation from remote sensing imagery plays a crucial role in land-cover mapping, forest inventory, and environmental monitoring. In tropical regions such as the Amazon rainforest, this task is particularly challenging because of heterogeneous canopy structures, atmospheric effects, and the limited availability of annotated data (Saikhammoon et al., 2023). Traditional remote sensing techniques based on spectral indices, such as the Normalized Difference Vegetation Index (NDVI), provide coarse vegetation cues but are insufficient for the precise pixel-level delineation of forested and non-forested areas. These limitations have motivated the adoption of deep learning-based semantic

segmentation approaches, which have demonstrated superior performance in extracting spatial and contextual information from remotely sensed imagery (Durga & Karthikeyan, 2023).

Although U-Net-based models, deeper encoders, attention mechanisms, and multiscale feature-fusion techniques have improved the accuracy of forest and land-cover segmentation (Soni & Rai, 2023), these gains are commonly achieved at the cost of increased parameter counts, memory consumption, and inference latency (Zhang et al., 2021). As a result, many high-performing models remain unsuitable for deployment on resource-constrained edge devices, including unmanned aerial vehicles, embedded platforms, and low-power forest-monitoring systems.

Lightweight encoders such as MobileNetV2 can reduce computational demand (Phoophuangpairaj & Charnkeitkong, 2026), while attention modules such as CBAM can improve feature discrimination. However, previous studies have not adequately examined whether selectively integrating CBAM into a lightweight MobileNetV2-based U-Net can improve binary forest-canopy segmentation without substantially increasing computational cost (Zhao et al., 2025). Moreover, many studies report segmentation accuracy alone, without jointly evaluating model size, inference speed, and the contribution of attention-module placement (Tang et al., 2024). Therefore, a clear research gap remains in the development and systematic evaluation of an attention-guided, lightweight segmentation architecture that balances pixel-level accuracy with computational efficiency and edge deployability.

Another challenge in remote sensing applications is the limited availability of labeled datasets. The Bragagnolo Amazon rainforest dataset provides high-quality annotations but contains a relatively small number of samples, thereby restricting generalization. Additionally, the use of downscaled input resolutions (e.g., 128×128) introduces a trade-off between computational efficiency and the preservation of fine-grained canopy structures, limiting the scope of the claims that can be made.

To address these gaps, this study proposes a lightweight UNet architecture that integrates a MobileNetV2 encoder with selectively placed CBAM modules in the decoder. The design prioritizes computational efficiency and edge deployability while introducing minimal architectural modifications to evaluate the incremental benefits of attention mechanisms. The model is evaluated on the Bragagnolo Amazon dataset for binary forest canopy segmentation using standard metrics, including the Dice coefficient and mean Intersection over Union, together with efficiency measures such as parameter count and inference speed. The evaluation explicitly acknowledges the dataset and resolution constraints, and the results are interpreted conservatively within this scope. The study is strictly limited to binary segmentation and does not address tasks such as change detection or selective logging.

1.1 Literature Review

Forest monitoring and deforestation detection have increasingly benefited from advances in artificial intelligence, remote sensing, and multimodal data processing. Zhao et al. (2023) conducted a

comprehensive review of individual tree crown (ITC) detection using convolutional neural networks, highlighting both the rapid progress in crown delineation and the persistent challenges posed by heterogeneous forest environments. Extending this body of research, Zheng et al. (2024) synthesized recent optical remote sensing-based ITC detection methods and emphasized the need for biologically informed models and multimodal fusion.

Beyond crown detection, Li et al. (2023) proposed a multiscale attention-based approach for trunk detection in urban scenes, demonstrating the value of attention modules in isolating fine structural details. Related research in orchard systems has also advanced tree-level analysis. Hai et al. (2025) developed a multimodal deep learning model for apple tree pruning assessment, while Wang et al. (2023a) fused machine learning and 3D image-processing techniques for diameter-at-breast-height estimation. In a related methodological vein, Pechprasarn et al. (2023) demonstrated the value of principal component analysis in identifying discriminative features for machine learning classification. Zhou et al. (2025) introduced a multimodal transformer-based architecture for apple disease detection and severity classification, while Bhola et al. (2023) compared transfer-learning models for cucumber disease classification, together illustrating the shift toward lightweight yet accurate field-ready models. Chau et al. (2024) improved tree instance segmentation for urban inventory applications by incorporating moving-object removal and chord-length-based DBH estimation. Complementing these efforts, Rianthakool et al. (2025) assess the eucalyptus stem diameter estimation using terrestrial laser scanning. Shaladi et al. (2025) introduced a hierarchical AI SDN firewall for securing forest monitoring networks.

At the individual-tree level, Liu et al. (2025) developed a graph optimization method for tree segmentation in terrestrial laser scanning (TLS) point clouds, achieving significant improvements in trunk localization and crown segmentation accuracy. Similarly, Jiang et al. (2025) created an organ-level PointNeXt-based segmentation pipeline for apple tree phenotyping, enabling the accurate extraction of multiple structural traits. Yin et al. (2025) addressed crop identification in complex mountainous landscapes through multidimensional features and rule-based object-oriented classification. Thakur et al. (2022) provided a systematic review of computer vision-based plant disease identification, illustrating

the transition from classical machine learning to deep learning approaches. Wu et al. (2025) reviewed intelligent orchard sprayer technologies, emphasizing the importance of sensor integration and real-time control in precision agriculture.

As data availability increases, label generation has become increasingly important. Albrecht et al. (2021) presented Autogeolabel, an automated geospatial label-generation pipeline that demonstrated strong accuracy even when using simple statistical features. In the broader context of spatial change analysis, Xiao et al. (2023) provided a comprehensive review of 3D urban object change detection using aerial and terrestrial point clouds. Gharineiat et al. (2022) reviewed machine learning applications in LiDAR-based topographic and surface-feature extraction. Ghazal et al. (2024) discussed computer vision applications across the crop life cycle in smart agriculture, while Gibril et al. (2021) developed a CNN-based framework for large-scale date palm mapping from UAV imagery. Zhao et al. (2025) explored trends in deep reinforcement learning (DRL) for agricultural machinery, emphasizing algorithm-hardware co-optimization for field deployment.

Deforestation detection has remained a major focus of remote sensing research. Torres et al. (2021) demonstrated the effectiveness of fully convolutional networks, including U-Net and DeepLab variants, for detecting forest loss in Landsat and Sentinel imagery. Solórzano et al. (2023) applied a 3D U-Net model to fuse SAR and multispectral data, achieving high accuracy in the spatiotemporal modelling of forest change. Ortega Adarme et al. (2022) improved deforestation detection using Sentinel-1 data by integrating distance-to-deforestation maps into a CNN framework. Expanding on this architecture, de Andrade et al. (2022) proposed enhanced DeepLabv3+ variants incorporating weighted focal loss to address class imbalance. Karaman et al. (2023) developed an attention-based network using Sentinel-1 SAR time series and introduced the BraDD SITS dataset, reporting substantial improvements in IoU.

John and Zhang (2022) designed an attention-based U-Net that outperformed multiple baselines across Amazon and Atlantic Forest datasets. Wang et al. (2023b) introduced SiamHRNet OCR, a high-resolution deforestation detection model capable of capturing fine boundary details. Vorotyntsev et al. (2021) applied deep semantic segmentation to deforestation mapping in Ukrainian forest regions, using TPU-based training to improve stability. Alshehri et al. (2024) adapted a transformer-based

Change Former architecture for bitemporal Sentinel-2 imagery, reporting strong accuracy across conservation areas. Silva et al. (2022) demonstrated that Sentinel-1 SAR data can achieve highly accurate near-real-time deforestation detection, even under persistent cloud cover in the Amazon.

Scattering-based SAR methods have also shown strong potential. Sugimoto et al. (2022) used scattering power decomposition with PALSAR-2 data to identify deforestation in tropical forests with high precision. Zhang et al. (2021) combined U-Net and LSTM for temporal forest change modelling, achieving competitive F1 performance. Haq et al. (2024) reviewed the integration of IoT, AI, and satellite-based systems for forest conservation, highlighting the role of real-time monitoring. Noa et al. (2021) introduced an adversarial domain adaptation framework for cross-biome deforestation detection, thereby addressing generalization challenges. Dascălu et al. (2023) modelled deforestation events using logistic analysis of Sentinel-1 time-series backscatter. Kaselimi et al. (2022) proposed a vision transformer for multilabel satellite image classification in forest monitoring. Tarazona et al. (2021) fused Landsat and SAR data using machine learning and the PVts β non-seasonal detection method, achieving high accuracy. Vega et al. (2021) designed a CycleGAN-based unsupervised domain adaptation framework for deforestation mapping in tropical biomes. Fodor and Conde (2023) created a deep multimodal model integrating Sentinel, Landsat, VIIRS, and MODIS imagery for burned-area and deforestation detection. Hasan et al. (2024) compared AI models for global deforestation monitoring and identified key challenges in hotspot detection. Bragagnolo et al. (2019) contributed an important Amazon Rainforest semantic segmentation dataset that is widely used for benchmarking. Finally, Kar and Singh (2025) developed MBTC Net, a multimodal attention-based architecture that illustrates the broader applicability of multimodal attention methods to environmental classification problems.

2. Objectives

The primary objective of this study is to design and evaluate a lightweight, attention-enhanced semantic segmentation framework tailored for binary forest canopy mapping on edge computing platforms. To fulfill this overarching goal, the study pursues four key specific objectives. First, it aims to develop a parameter-efficient UNet variant by coupling a MobileNetV2 feature extractor with a Convolutional

Block Attention Module to balance low computational complexity with high spatial accuracy. Second, it seeks to demonstrate accurate pixel-level binary classification of forest and non-forest regions using Amazon Rainforest satellite imagery. Third, it aims to systematically evaluate the performance–efficiency trade-offs by analyzing segmentation accuracy alongside parameter count, memory footprint, and inference speed, specifically isolating the empirical impact of the attention mechanism. Finally, the study aims to validate the operational feasibility of deploying this lightweight architecture on low-power, resource-constrained edge hardware under practical dataset limitations.

3. Methodology

The end-to-end methodology encompasses the design, training, and evaluation of an efficient, attention-guided semantic segmentation model for high-resolution deforestation and forest-cover detection. The proposed system aims to achieve high accuracy while remaining deployable on edge devices with limited computational resources. It integrates

three key components: a U-Net segmentation backbone, MobileNetV2 as a lightweight encoder, and the Convolutional Block Attention Module (CBAM) for adaptive feature refinement.

3.1 System Overview

The proposed segmentation pipeline, illustrated in Figure 1, is designed to process satellite imagery for forest classification tasks. The pipeline begins with the ingestion of RGB TIFF images and their corresponding grayscale segmentation masks from the Amazon Rainforest Dataset. The images undergo preprocessing and augmentation before being passed to the segmentation model. The encoder, based on MobileNetV2 pretrained on ImageNet, efficiently extracts spatial features. The decoder follows the UNet architecture and reconstructs the segmentation map using skip connections. A CBAM is incorporated into the final stage of the decoder to refine attention across the spatial and channel dimensions. The trained model is exported in ONNX format to support cross-platform deployment.

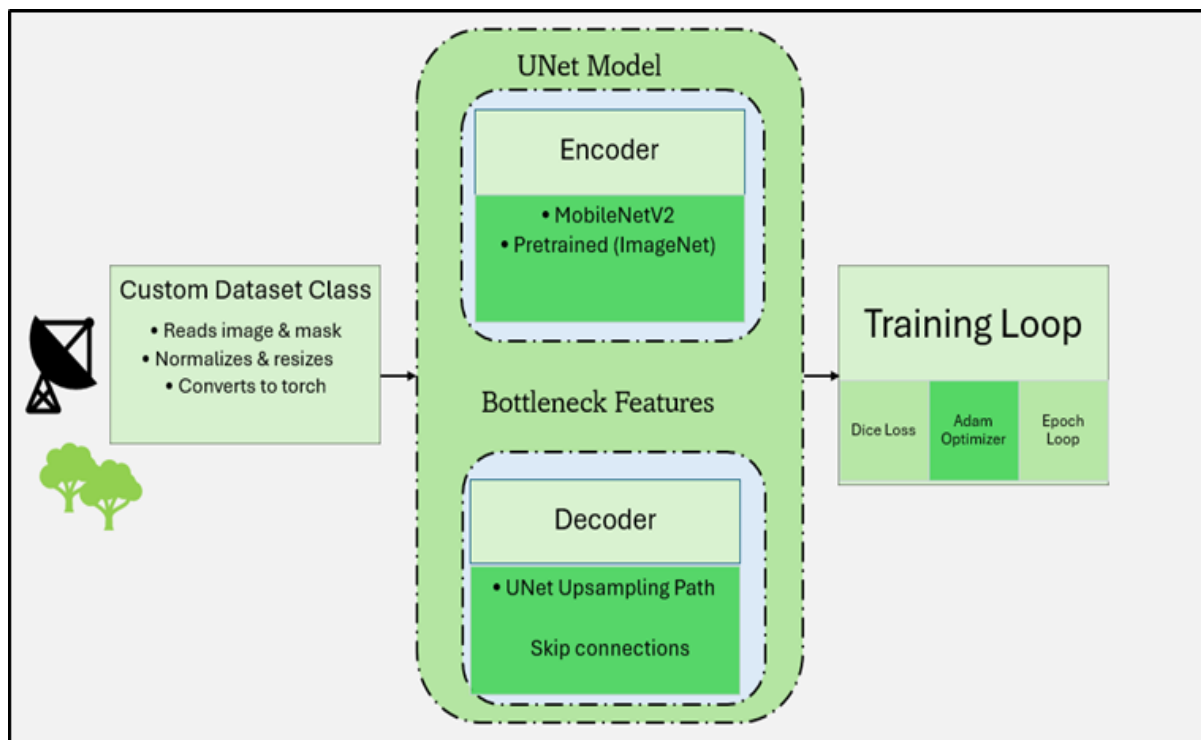


Figure 1 Architecture of the Amazon Forest Segmentation Pipeline

3.2 Dataset

The experiments were conducted using the publicly available Amazon Rainforest Satellite Image Dataset (Bragagnolo et al., 2019), which consists of 60 RGB images with dimensions of 512×512 pixels and corresponding binary masks indicating forested and non-forested regions. These high-resolution images capture detailed canopy structures and are suitable for evaluating lightweight semantic segmentation models. Although the dataset is relatively modest in size, its spatial fidelity and ground-truth annotations make it well suited for prototyping models intended for field-level deployment. To ensure transparent reporting and reproducibility, the dataset was divided into 70% training data (42 images), 20% validation data (12 images), and 10% testing data (6 images). This structured split supports appropriate performance evaluation despite the modest dataset size. Data augmentation techniques were applied to increase sample variability and reduce overfitting.

3.3 Data Preprocessing and Augmentation

To improve computational efficiency and promote generalization, all images and masks were resized to 128×128 pixels. The Albumentations library was used to implement a comprehensive augmentation pipeline comprising random horizontal and vertical flips, 90° rotations, elastic transformations, brightness and contrast adjustments, and image normalization. These augmentations were intended to help the model learn robust features that are invariant to orientation, lighting conditions, and seasonal variation, thereby improving its performance across diverse forest structures and canopy types.

3.4 Model Architecture

The model architecture builds on the UNet framework, which is characterized by its symmetric encoder-decoder structure and the use of skip connections to preserve spatial features. It is enhanced in two keyways: first, by replacing a computationally intensive encoder (e.g., ResNet) with the lightweight MobileNetV2 architecture; and second, by integrating a Convolutional Block Attention Module (CBAM) into the decoder path to improve feature selection.

3.4.1 Encoder: MobileNetV2

MobileNetV2 is used as the encoder because of its computational efficiency and low parameter count. Designed for mobile and embedded vision applications, it employs depth wise separable

convolutions and inverted residual blocks, thereby significantly reducing computational cost while preserving representational capacity. The encoder is initialized with weights pretrained on ImageNet, enabling faster convergence during training and strong generalization despite the limited amount of data.

3.4.2 Decoder: UNet with Skip Connections

The decoder mirrors the encoder through a symmetric series of up sampling layers. It uses skip connections to combine high-resolution spatial features from the encoder with high-level semantic features. This architecture enables precise localization of object boundaries, which is essential for delineating tree crowns and forest edges in satellite imagery.

3.4.3 Attention Mechanism: CBAM

To further refine the extracted features, the proposed model integrates a CBAM into the final decoding block. The CBAM sequentially applies channel- and spatial-attention mechanisms. Channel attention aggregates global spatial context through average and max pooling, followed by a shared multilayer perception (MLP) and a sigmoid activation function to emphasize relevant channels. Spatial attention, in turn, identifies important regions within the feature map by passing pooled feature maps through a 7×7 convolution. This dual-attention strategy helps the network focus on discriminative regions and reduce background noise, which is particularly valuable for identifying fragmented canopy boundaries.

3.4.4 Training Configuration

The model (Figure 2) is implemented in PyTorch and built using the segmentation_models_pytorch library to support architectural modularity. Training employs a hybrid loss function that combines Dice loss (Eq. 1) and binary cross-entropy (BCE; Eq. 2), with both components weighed equally to leverage their respective strengths.

The Dice Loss is given by (Eq. 1)

$$L_{\text{Dice}} = 1 - \frac{2 \sum_{i=1}^N y_i \hat{y}_i}{\sum_{i=1}^N y_i + \sum_{i=1}^N \hat{y}_i} \quad (\text{Eq.1})$$

Binary Cross Entropy (BCE) Loss is given by

$$L_{\text{BCE}} = -\frac{1}{N} \sum_{i=1}^N [y_i \cdot \log(\hat{y}_i) + (1-y_i) \log(1-\hat{y}_i)] \quad (\text{Eq.2})$$

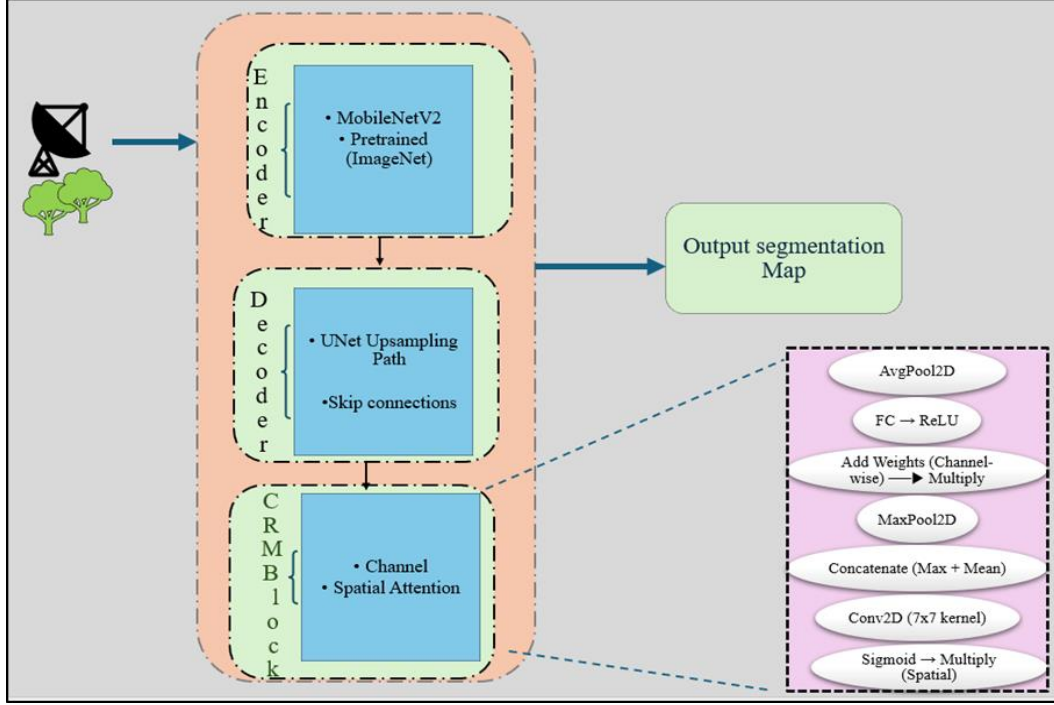


Figure 2 Proposed UNet CBAM architecture with a mobilenetv2 encoder and decoder

Dice Loss effectively handles class imbalance by maximizing the overlap between predicted and ground truth regions, while BCE ensures consistent pixel wise classification:

The combined loss is given by

$$L_{\text{combined}} = 0.5 \times L_{\text{Dice}} + 0.5 \times L_{\text{BCE}} \quad (\text{Eq.3})$$

The model is trained with a batch size of 16 for five epochs. Early stopping is configured on validation mIoU, and the best-performing checkpoint is retained for deployment. This deliberately short schedule reflects the small size of the dataset and the study's emphasis on efficiency; longer training on larger datasets is identified as future work.

3.5 Evaluation Metrics

To comprehensively evaluate the model's effectiveness and deploy ability, four standard metrics are used. Dice Score used to measure the degree of overlap between prediction and ground truth using

$$\text{Dice} = \frac{2 \text{ TP}}{2 \text{ TP} + \text{FP} + \text{FN}} \quad (\text{Eq.4}).$$

The Mean Intersection over Union (mIoU): Assesses segmentation performance across forest and non forest classes.

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

And

$$\text{mIoU} = \frac{1}{C} \sum_{C=1}^C \text{IoU} \quad (\text{Eq.6})$$

where TP is True Positives, FP = False Positives, FN = False Negatives, c = Number of classes.

Parameter Count: Indicates model complexity and memory usage. Frames Per Second (FPS): Evaluates real time inference speed on edge hardware (Jetson Nano). These metrics provide a balanced evaluation of the model's predictive accuracy, efficiency, and practical viability in field applications.

4. Result and Discussion

The proposed CBAM-UNet model was evaluated using the Amazon Rainforest Dataset (Bragagnolo et al., 2019), which supports binary semantic segmentation of forested and non-forested regions. The dataset comprises 60 high-resolution satellite images (512×512 pixels) with corresponding annotated binary masks. Despite its limited size, the dataset contains sufficient visual variability to provide a realistic testbed for assessing lightweight deep

learning models intended for deployment in edge-based ecological monitoring systems. The use of separate training, validation, and testing subsets supports evaluation of unseen data; however, the small sample size limits the extent to which the findings can be generalized. Broader validation across different ecosystems and sensor types will require more diverse training data.

4.1 Visual Evaluation and Ground Truth Comparison

Visual comparison of predicted masks with ground-truth annotations is an essential component of semantic segmentation evaluation. The qualitative results indicate that the CBAM-UNet architecture effectively identifies major forested regions while improving boundary localization, particularly along fragmented canopy edges and transition zones. As shown in Figure 3, both models successfully delineate the principal forested areas; however, the CBAM-UNet produces more precise boundaries than the baseline model. These improvements demonstrate the value of CBAM in enhancing the spatial sensitivity of lightweight segmentation architectures. Nevertheless, minor errors remain in areas characterized by fine-scale canopy fragmentation, indicating that extremely heterogeneous regions continue to present challenges.

These comparisons underscore the importance of integrating attention mechanisms such as CBAM to enhance spatial sensitivity. Although the baseline model performs adequately, the attention-enhanced model provides more precise localization of canopy regions, which is critical for applications such as tree-crown delineation and fine-scale forest monitoring.

4.2 Quantitative Performance Analysis

Model performance was rigorously assessed through a comparative evaluation of two architectures: (1) the proposed CBAM-UNet and (2) a baseline UNet employing the same MobileNetV2 encoder without attention modules. As shown in Table 1, both models were evaluated using standard semantic segmentation metrics, including the Dice score, mean Intersection over Union (mIoU), parameter count, and inference speed, measured in frames per second (FPS) on edge hardware.

- Dice Score quantifies the overlap between predicted and actual forest regions.
- mIoU averages the intersection over union across both classes.
- Parameter Count estimates memory and computational load.
- FPS reflects real time inference capability.

Table 1 Performance comparison of CBAM-UNet vs. Plain UNet

Model	Dice Score ↑	mIoU ↑	Parameters	FPS (↑ = faster)
CBAM-UNet	0.6141	0.6339	6,629,048	5.99
Plain UNet	0.6187	0.6282	6,630,120	6.23

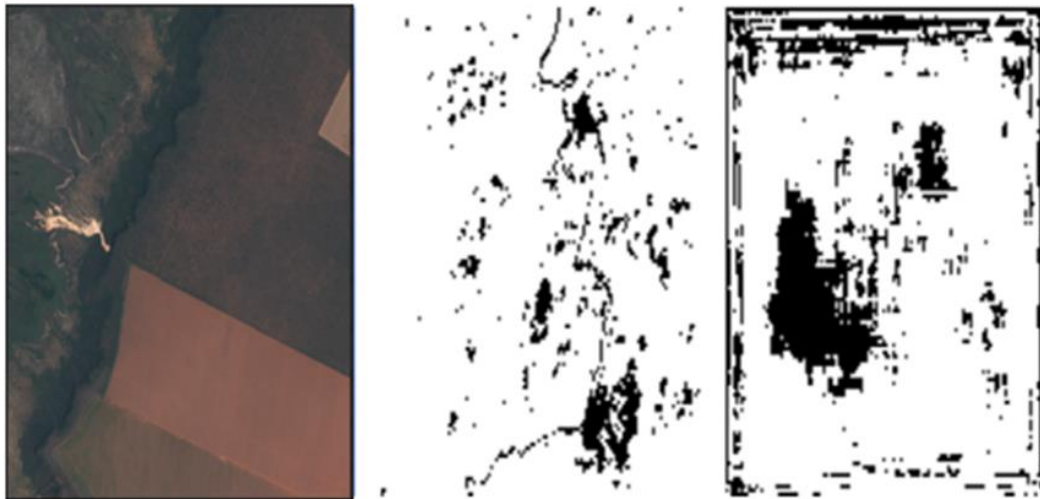


Figure 3 Visual comparison between the predicted segmentation mask and ground truth annotation

ayer (type:depth-idx)	Output Shape	Param #
BAMUNet	[1, 1, 128, 128]	--
-Sequential: 1-1	[1, 1, 128, 128]	--
└Conv2d: 2-1	[1, 16, 128, 128]	448
└ReLU: 2-2	[1, 16, 128, 128]	--
└Conv2d: 2-3	[1, 1, 128, 128]	145
-CBAM: 1-2	[1, 1, 128, 128]	103
=====		
total params: 696		
trainable params: 696		
non-trainable params: 0		
total mult-adds (Units.MEGABYTES): 9.72		
=====		
input size (MB): 0.20		
forward/backward pass size (MB): 2.23		
params size (MB): 0.00		
estimated Total Size (MB): 2.43		
=====		

Figure 4 Summary visualization of lightweight CBAM-UNet architecture

Table 2 Edge device deployment benchmarks

Device	Model Size	FPS	Memory Usage	Platform
Jetson Nano	6.6M	5.9	~78 MB	ONNX
Raspberry Pi 4	6.6M	5.9	~82 MB	ONNX
Intel NUC	6.6M	12.3	~65 MB	ONNX

The CBAM-UNet achieved a Dice score of 0.6141 and an mIoU of 0.6339, whereas the baseline UNet-MobileNetV2 achieved a Dice score of 0.6187 and an mIoU of 0.6282. These findings indicate that the primary improvement associated with CBAM-UNet lies in mIoU, which more effectively captures balanced class-wise overlap. Although the baseline model attained a marginally higher Dice score, CBAM-UNet produced more consistent segmentation across both classes and more accurate boundary delineation.

Although the baseline model achieved a marginally higher Dice score, CBAM-UNet demonstrated stronger mIoU performance and more accurate boundary delineation in the visual outputs. Because mIoU averages performance across both classes, it provides a more balanced assessment of segmentation quality than the Dice score, which is more sensitive to class imbalance.

4.3 Architectural and Computational Overview

The compact architecture of CBAM-UNet is one of its principal strengths. As illustrated in Figure 4, the model processes 128×128 inputs through a series of convolutional layers, producing attention-enhanced segmentation outputs. The encoder is composed of MobileNetV2 blocks that use depth wise separable convolutions to achieve computational efficiency. Following decoding, the CBAM module refines the feature maps through the sequential application of channel and spatial attention.

The compact architecture of CBAM-UNet is one of its principal strengths. As illustrated in Figure 4, the model processes 128×128 inputs through a series of convolutional layers, producing attention-enhanced segmentation outputs. The encoder comprises MobileNetV2 blocks that use depth wise separable convolutions to achieve computational efficiency. Following decoding, the CBAM refines the feature maps through the sequential application of channel and spatial attention. The complete model contains approximately 6.63 million parameters and occupies roughly 25 MB in 32-bit floating-point format, which remains well within the memory budget of the target edge platforms.

4.4 Edge Deployment Performance

To validate real-world deployability, the trained model was exported to ONNX format and deployed on three edge-computing platforms: the NVIDIA Jetson Nano (4 GB RAM), Raspberry Pi 4, and Intel NUC. These platforms represent typical field-deployment devices used in forest monitoring applications involving drones, mobile terminals, or sensor nodes, as summarized in Table 2.

The results affirm that the model delivers real time performance (≥ 5 FPS) on midrange edge devices, with RAM usage consistently under 85 MB. This level of performance is crucial for deploying AI in connectivity constrained or off grid regions, where cloud inference is not feasible.

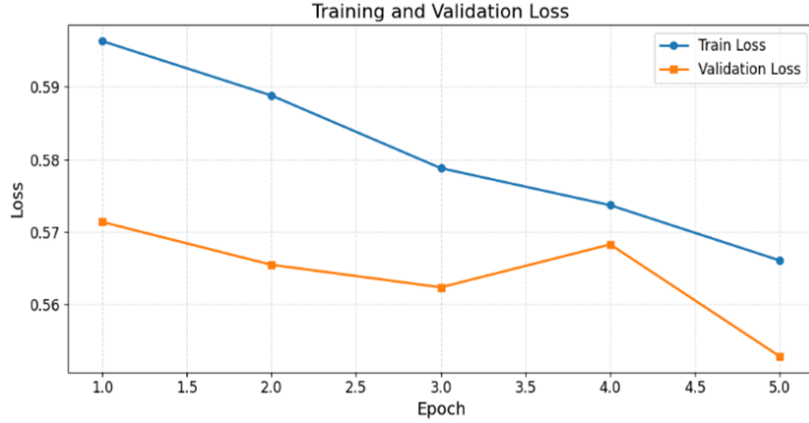


Figure 5 Training and validation loss curves over five epochs using hybrid Dice + BCE loss

4.5 Training Convergence and Learning Behavior

The training trajectory of the CBAM-UNet model was monitored using training and validation loss curves over five epochs. Figure 5 presents the hybrid loss values, comprising Dice loss and binary cross-entropy, across the training epoch. The training loss decreased monotonically across all five epochs, while the validation loss declined overall to 0.5529 by the final epoch, with a transient increase at epoch 4. Given the very short schedule, this fluctuation is more plausibly attributable to batch-level variance than to the onset of overfitting. No divergence between training and validation loss was observed within this short training schedule; however, because training was limited to five epochs (approximately 15 gradient updates), the absence of overfitting should be regarded as a consequence of the abbreviated schedule rather than as evidence of strong generalization. The hybrid loss function helps balance region level segmentation with pixel wise prediction accuracy, while the use of early stopping ensures training efficiency and avoids unnecessary epochs.

4.6 Ablation Study

To isolate the contributions of individual architectural components, an ablation study was conducted comparing three model configurations:

1. Baseline: UNet + MobileNetV2 (no attention)
2. Proposed: UNet + MobileNetV2 + CBAM
3. Heavy Encoder: UNet + ResNet34

The inclusion of CBAM improved mIoU by 0.0057 (from 0.6282 to 0.6339), a relative gain of 0.91%, indicating a modest refinement of spatial focus without substantial computational overhead. By contrast, replacing MobileNetV2 with ResNet34 produced slightly higher accuracy but increased the

model size to 23.5 million parameters, thereby reducing its suitability for edge deployment. These findings indicate that the proposed CBAM-MobileNetV2 configuration achieves a favorable balance between segmentation performance and computational efficiency.

In addition to comparing the baseline UNet-MobileNetV2 model with its CBAM-enhanced variant, preliminary experiments evaluated alternative CBAM placement strategies within the decoder. Applying CBAM modules at all decoder stages or at earlier decoding layers did not produce consistent improvements and, in some cases, resulted in marginal performance degradation because of the additional parameter overhead. CBAM was therefore retained only at selected decoder stages, where the balance between feature refinement and computational cost was most favorable. These results suggest that attention mechanisms provide limited but measurable benefits in lightweight architecture and that indiscriminate attention placement may be counterproductive under edge-deployment constraints.

4.7 Field Deployment Validation

For further validation, the final trained model was exported to ONNX format and evaluated on the Jetson Nano in terms of inference latency and resource consumption. The ONNX model processed 128×128 inputs at approximately 5.9 FPS while consuming less than 80 MB of RAM. These results demonstrate their suitability for integration into field-deployable systems, including UAVs, smartphones, and embedded devices. The ONNX export also provides platform independence, enabling deployment across Windows, Linux, and embedded operating systems.

Table 3 Summary of the CBAM-UNet Pipeline Configuration and Edge Deployment Performance

Component	Configuration Details
Architecture	UNet + MobileNetV2 + CBAM
Dataset	Amazon Rainforest (RGB TIFF + binary PNG masks)
Input Resolution	128×128 pixels
Preprocessing	Albumentations: flips, rotation, elastic transformations, normalize
Loss Function	Hybrid: 50% Dice + 50% BCE
Training Duration	5 epochs with early stopping
Evaluation Metrics	Dice Score, mIoU, FPS, Parameter Count
Export Format	PyTorch (.pth), ONNX (.onnx)

4.8 Pipeline and Deployment Summary

A comprehensive snapshot of the model's configuration, training setup, and deployment profile is provided in Table 3. The architecture is inherently modular, enabling future extension to multiclass segmentation tasks or the integration of multimodal input channels. The efficient training strategy, comprising five epochs, and the compact model footprint make the system accessible to resource-constrained organizations and field researchers. Owing to the limited dataset size, cross-validation and statistical averaging across multiple runs were not feasible within the scope of this study. All reported results were obtained from a single training run using fixed hyperparameters. Consequently, small performance differences, such as the observed improvement of 0.0057 in mIoU, should be interpreted cautiously, as they may fall within the range of normal training variability. Accordingly, this work emphasizes feasibility and efficiency trends rather than statistically definitive performance gains.

5. Conclusion

This study presents CBAM-UNet, a lightweight, edge-deployable architecture designed for forest-cover segmentation in resource-constrained environments. By integrating MobileNetV2 and CBAM within a UNet framework, the model combines computational efficiency, spatial precision, and attention-enhanced feature refinement, thereby addressing key limitations of semantic segmentation approaches that rely on computationally intensive encoders or providing limited interpretability. Experiments conducted on the Amazon Rainforest Dataset (Bragagnolo et al., 2019) showed that CBAM-UNet achieved a higher mIoU than the baseline model, while differences in Dice score remained minimal. These findings indicate that the primary performance gain was associated with improved class-wise segmentation rather than overall region overlap. A major strength of the proposed approach is its attention-driven interpretability, which

helps identify ecologically important regions such as canopy gaps and forest edges features that may be overlooked by conventional CNNs. This capability enhances the model's value in operational settings where transparency and actionable insights are essential. A key limitation of this study is the relatively small dataset size of 60 annotated images, which restricts the range of canopy structures, environmental conditions, and geographic variability represented. Although data augmentation and the lightweight model design help mitigate overfitting, generalization across different forest biomes and satellite platforms may remain limited. Future research should therefore incorporate larger multisession datasets, cross-biome sampling, and multimodal inputs to improve robustness and transferability.

Further development of CBAM-UNet may include multiclass segmentation, temporal change detection, and multimodal fusion with SAR or LiDAR data. Additional optimization techniques, such as pruning, quantization, and explainable AI methods, may further improve its suitability for real-time, field-level forest monitoring.

6. Acknowledgements

I acknowledge the support provided by the Department of Computer Science and Engineering at Nitte Meenakshi Institute of Technology (NMIT), Bengaluru, for providing the necessary computational resources and research environment to carry out this study. The author also expresses gratitude to the developers and contributors of open-source libraries, including PyTorch, segmentation_models_pytorch, and Albumentations, which facilitated the implementation and experimentation of the proposed model. Special thanks are extended to the creators of the Amazon Rainforest Dataset (Bragagnolo et al., 2019) for making high-quality annotated data publicly available for research purposes.

7. Abbreviations

Abbreviation	Full Term
CBAM	Convolutional Block Attention Module
DBH	Diameter at Breast Height
UNet	U-shaped Convolutional Neural Network
DRL	Deep Reinforcement Learning
CNN	Convolutional Neural Network
mIoU	Mean Intersection over Union
BCE	Binary Cross Entropy
FPS	Frames Per Second
RGB	Red, Green, Blue
ONNX	Open Neural Network Exchange
IoU	Intersection over Union
ITC	Individual Tree Crown
LiDAR	Light Detection and Ranging
TP	True Positive
FP	False Positive
FN	False Negative
MLP	Multilayer Perceptron
NDVI	Normalized Difference Vegetation Index
SAR	Synthetic Aperture Radar
TLS	Terrestrial Laser Scanning
UAV	Unmanned Aerial Vehicle

8. CRediT Statement

Shobha P: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Resources, Supervision.

9. References

- Albrecht, C. M., Marianno, F., & Klein, L. J. (2021). *Autogeolabel: Automated label generation for geospatial machine learning* [Conference presentation]. 2021 IEEE International Conference on Big Data (Big Data). IEEE, Orlando, FL, USA.
<https://doi.org/10.1109/BigData52589.2021.9672060>
- Alshehri, M., Ouadou, A., & Scott, G. J. (2024). Deep transformer-based network deforestation detection in the Brazilian Amazon using Sentinel-2 imagery. *IEEE Geoscience and Remote Sensing Letters*, 21, 1-5.
<https://doi.org/10.1109/LGRS.2024.3355104>
- Bhola, A., Verma, S., & Kumar, P. (2023). A comparative analysis of deep learning models for cucumber disease classification using transfer learning. *Journal of Current Science and Technology*, 13(1), 23-35.
<https://doi.org/10.14456/jcst.2023.3>
- Bragagnolo, L., da Silva, R. V., & Grzybowski, J. M. V. (2019). *Amazon rainforest dataset for semantic segmentation* [Data set].
<https://doi.org/10.5281/zenodo.3233081>
- Chau, W. Y., Chow, J. K., Tan, T. J., Wu, J., Leung, M. L., Tan, P. S., ... & Wang, Y. H. (2024). Towards automatic urban tree inventory: Enhancing tree instance segmentation via moving object removal and a chord length-based DBH estimation approach. *Computers and Electronics in Agriculture*, 226, Article 109378.
<https://doi.org/10.1016/j.compag.2024.109378>
- Dascălu, A., Catalão, J., & Navarro, A. (2023). Detecting deforestation using logistic analysis and Sentinel-1 multitemporal backscatter data. *Remote Sensing*, 15(2), Article 290.
<https://doi.org/10.3390/rs15020290>
- de Andrade, R. B., Mota, G. L. A., & da Costa, G. A. O. P. (2022). Deforestation detection in the amazon using DeepLabv3+ semantic segmentation model variants. *Remote Sensing*, 14(19), Article 4694.
<https://doi.org/10.3390/rs14194694>
- Durga, P., & Karthikeyan, S. (2023). Comparative analysis for augmented decision-making applications using deep learning models. *Journal of Current Science and Technology*, 13(3), 791-803.
<https://doi.org/10.59796/jcst.V13N3.2023.2273>
- Fodor, G., & Conde, M. V. (2023). Rapid deforestation and burned area detection using deep multimodal learning on satellite imagery. *arXiv preprint arXiv:2307.04916*.
<https://doi.org/10.48550/arXiv.2307.04916>
- Gharineiat, Z., Tarsha Kurdi, F., & Campbell, G. (2022). Review of automatic processing of topography and surface feature identification LiDAR data using machine learning techniques. *Remote Sensing*, 14(19), Article 4685. <https://doi.org/10.3390/rs14194685>
- Ghazal, S., Munir, A., & Qureshi, W. S. (2024). Computer vision in smart agriculture and precision farming: Techniques and applications. *Artificial Intelligence in Agriculture*, 13, 64-83.
<https://doi.org/10.1016/j.aiaa.2024.06.004>
- Gibril, M. B. A., Shafri, H. Z. M., Shanableh, A., Al-Ruzouq, R., Wayayok, A., & Hashim, S. J. (2021). Deep convolutional neural network for

- large-scale date palm tree mapping from UAV-based images. *Remote Sensing*, 13(14), Article 2787.
<https://doi.org/10.3390/rs13142787>
- Hai, T., Wang, W., Yan, F., Liu, M., Li, C., Li, S., ... & Lv, C. (2025). A deep learning-based approach to apple tree pruning and evaluation with multi-modal data for enhanced accuracy in agricultural practices. *Agronomy*, 15(5), Article 1242.
<https://doi.org/10.3390/agronomy15051242>
- Haq, B., Jamshed, M. A., Ali, K., Kasi, B., Arshad, S., Kasi, M. K., ... & Ur-Rehman, M. (2024). Tech-driven forest conservation: Combating deforestation with internet of things, artificial intelligence, and remote sensing. *IEEE Internet of Things Journal*, 11(14), 24551-24568.
<https://doi.org/10.1109/JIOT.2024.3378671>
- Hasan, R., Farabi, S. F., Kamruzzaman, M., Bhuyan, M. K., Nilima, S. I., & Shahana, A. (2024). AI-driven strategies for reducing deforestation. *American Journal of Engineering and Technology*, 6(6), 6–20.
<https://doi.org/10.37547/tajet/Volume06Issue06-02>
- Jiang, L., Li, C., & Fu, L. (2025). Apple tree architectural trait phenotyping with organ-level instance segmentation from point cloud. *Computers and Electronics in Agriculture*, 229, Article 109708.
<https://doi.org/10.1016/j.compag.2024.109708>
- John, D., & Zhang, C. (2022). An attention-based U-Net for detecting deforestation within satellite sensor imagery. *International Journal of Applied Earth Observation and Geoinformation*, 107, Article 102685.
<https://doi.org/10.1016/j.jag.2022.102685>
- Kar, S., & Singh, P. K. (2025). MBTC-Net: Multimodal brain tumor classification from CT and MRI scans using deep neural network with multi-head attention mechanism. *Medicine in Novel Technology and Devices*, 27, Article 100382.
<https://doi.org/10.1016/j.medntd.2025.100382>
- Karaman, K., Sainte Fare Garnot, V., & Wegner, J. D. (2023). Deforestation detection in the Amazon with Sentinel-1 SAR image time series. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 10, 835-842.
<https://doi.org/10.5194/isprs-annals-X-1-W1-2023-835-2023>
- Kaselimi, M., Voulodimos, A., Daskalopoulos, I., Doulamis, N., & Doulamis, A. (2022). A vision transformer model for convolution-free multilabel classification of satellite imagery in deforestation monitoring. *IEEE Transactions on Neural Networks and Learning Systems*, 34(7), 3299-3307.
<https://doi.org/10.1109/TNNLS.2022.3144791>
- Li, R., Sun, G., Wang, S., Tan, T., & Xu, F. (2023). Tree trunk detection in urban scenes using a multiscale attention-based deep learning method. *Ecological Informatics*, 77, Article 102215.
<https://doi.org/10.1016/j.ecoinf.2023.102215>
- Liu, Y., Chen, D., Na, J., Peethambaran, J., Pfeifer, N., & Zhang, L. (2025). Segmentation of individual trees in TLS point clouds via graph optimization. *IEEE Transactions on Geoscience and Remote Sensing*, 63, 1-21.
<https://doi.org/10.1109/TGRS.2025.3567357>
- Noa, J., Soto, P. J., Costa, G. A. O. P., Wittich, D., Feitosa, R. Q., & Rottensteiner, F. (2021). Adversarial discriminative domain adaptation for deforestation detection. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 3, 151-158.
<https://doi.org/10.15488/16631>
- Ortega Adarme, M., Doblas Prieto, J., Queiroz Feitosa, R., & De Almeida, C. A. (2022). Improving deforestation detection on tropical rainforests using Sentinel-1 data and convolutional neural networks. *Remote Sensing*, 14(14), Article 3290.
<https://doi.org/10.3390/rs14143290>
- Pechprasarn, S., Wattanapermpool, O., Warunlawan, M., Homsud, P., & Akarajarasroj, T. (2023). Identification of important factors in the diagnosis of breast cancer cells using machine learning models and principal component analysis. *Journal of Current Science and Technology*, 13(3), 642-656.
<https://doi.org/10.59796/jcst.V13N3.2023.700>
- Phoophuangpairoj, R., & Charnkeitkong, P. (2026). A hybrid deep learning and machine learning approach for predicting aqueous solution concentrations. *Journal of Current Science and Technology*, 16(2), Article 173.
<https://doi.org/10.59796/jcst.V16N2.2026.173>
- Rianthakool, L., Yingdee, J., Chumsangsri, T., & Sittibal, D. (2025). Assessment of eucalyptus

- stem diameter estimation using terrestrial laser scanning. *Current Applied Science and Technology*, 25(5), Article e0264569. <https://doi.org/10.55003/cast.2025.264569>
- Saikhammoon, R., Sungkaew, S., Thinkampaeng, S., Phumphuang, W., Kamy, T., & Marod, D. (2023). Forest restoration in an abandoned seasonally dry tropical forest in the Mae Klong Watershed, Western Thailand. *Environment and Natural Resources Journal*, 21(5), 443-457. <https://doi.org/10.32526/enrj/21/20230121>
- Shaladi, T. A. H., Nashnosh, M. T. R., & Alkabir, M. M. (2025). Forest tree AI SDN firewall: A hierarchical architecture for adaptive network security. *International Journal of Current Science Research and Review*, 8(5), 2519-2530. <https://doi.org/10.47191/ijcsrr/V8-i5-62>
- Silva, C. A., Guerrisi, G., Del Frate, F., & Sano, E. E. (2022). Near-real time deforestation detection in the Brazilian Amazon with Sentinel-1 and neural networks. *European Journal of Remote Sensing*, 55(1), 129-149. <https://doi.org/10.1080/22797254.2021.2025154>
- Solórzano, J. V., Mas, J. F., Gallardo-Cruz, J. A., Gao, Y., & de Oca, A. F. M. (2023). Deforestation detection using a spatio-temporal deep learning approach with synthetic aperture radar and multispectral images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 199, 87-101. <https://doi.org/10.1016/j.isprsjprs.2023.03.017>
- Soni, A., & Rai, A. (2023). A systematic survey on deep learning techniques for chest disease detection using chest radiographs. *Journal of Current Science and Technology*, 13(2), 267-295. <https://doi.org/10.59796/jcst.V13N2.2023.1744>
- Sugimoto, R., Kato, S., Nakamura, R., Tsutsumi, C., & Yamaguchi, Y. (2022). Deforestation detection using scattering power decomposition and optimal averaging of volume scattering power in tropical rainforest regions. *Remote Sensing of Environment*, 275, Article 113018. <https://doi.org/10.1016/j.rse.2022.113018>
- Tang, J., Weeramongkolkul, M., Suwankesawong, S., Saengtabtim, K., Leelawat, N., & Wongwailikhit, K. (2024). Toward a more resilient Thailand: Developing a machine learning-powered forest fire warning system. *Heliyon*, 10(13), Article e34021. <https://doi.org/10.1016/j.heliyon.2024.e34021>
- Tarazona, Y., Zabala, A., Pons, X., Broquetas, A., Nowosad, J., & Zurqani, H. A. (2021). Fusing Landsat and SAR data for mapping tropical deforestation through machine learning classification and the PVts- β non-seasonal detection approach. *Canadian Journal of Remote Sensing*, 47(5), 677-696. <https://doi.org/10.1080/07038992.2021.1941823>
- Thakur, P. S., Khanna, P., Sheorey, T., & Ojha, A. (2022). Trends in vision-based machine learning techniques for plant disease identification: A systematic review. *Expert Systems with Applications*, 208, Article 118117. <https://doi.org/10.1016/j.eswa.2022.118117>
- Torres, D. L., Turnes, J. N., Soto Vega, P. J., Feitosa, R. Q., Silva, D. E., Marcato Junior, J., & Almeida, C. (2021). Deforestation detection with fully convolutional networks in the Amazon Forest from Landsat-8 and Sentinel-2 images. *Remote Sensing*, 13(24), Article 5084. <https://doi.org/10.3390/rs13245084>
- Vega, P. J. S., da Costa, G. A. O. P., Feitosa, R. Q., Adarme, M. X. O., de Almeida, C. A., Heipke, C., & Rottensteiner, F. (2021). An unsupervised domain adaptation approach for change detection and its application to deforestation mapping in tropical biomes. *ISPRS Journal of Photogrammetry and Remote Sensing*, 181, 113-128. <https://doi.org/10.1016/j.isprsjprs.2021.08.026>
- Vorotyntsev, P., Gordienko, Y., Alienin, O., Rokovyi, O., & Stirenko, S. (2021). *Satellite image segmentation using deep learning for deforestation detection* [Conference presentation]. 2021 IEEE 3rd ukraine conference on electrical and computer engineering (UKRCON). IEEE, Lviv, Ukraine. <https://doi.org/10.1109/UKRCON53503.2021.9575783>
- Wang, Y., Sun, J., & Wang, F. (2023a). Estimation of forest Diameter-at-Breast-Height: A fusion of machine learning and 3D image processing innovations. *Traitement du Signal*, 40(5), Article 2291. <https://doi.org/10.18280/ts.400547>
- Wang, Z., Liu, D., Liao, X., Pu, W., Wang, Z., & Zhang, Q. (2023b). SiamHRnet-OCR: A novel deforestation detection model with

- high-resolution imagery and deep learning. *Remote Sensing*, 15(2), Article 463.
<https://doi.org/10.3390/rs15020463>
- Wu, M., Liu, S., Li, Z., Ou, M., Dai, S., Dong, X., ... & Jia, W. (2025). A review of intelligent orchard sprayer technologies: Perception, control, and system integration. *Horticulturae*, 11(6), Article 668.
<https://doi.org/10.3390/horticulturae11060668>
- Xiao, W., Cao, H., Tang, M., Zhang, Z., & Chen, N. (2023). 3D urban object change detection from aerial and terrestrial point clouds: A review. *International Journal of Applied Earth Observation and Geoinformation*, 118, Article 103258.
<https://doi.org/10.1016/j.jag.2023.103258>
- Yin, L., Zhou, Z., Zhao, W., Liao, Y., Huang, D., & Li, W. (2025). Research on identification and extraction of crop plants in plateau mountainous areas based on multi-dimensional features. *International Journal of Remote Sensing*, 46(1), 77-104.
<https://doi.org/10.1080/01431161.2024.2391080>
- Zhang, J., Wang, Z., Bai, L., Song, G., Tao, J., & Chen, L. (2021). *Deforestation detection based on U-Net and LSTM in optical satellite remote sensing images* [Conference presentation]. 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. IEEE, Brussels, Belgium.
<https://doi.org/10.1109/IGARSS47720.2021.9554689>
- Zhao, H., Morgenroth, J., Pearse, G., & Schindler, J. (2023). A systematic review of individual tree crown detection and delineation with convolutional neural networks (CNN). *Current Forestry Reports*, 9(3), 149-170.
<https://doi.org/10.1007/s40725-023-00184-3>
- Zhao, J., Fan, S., Zhang, B., Wang, A., Zhang, L., & Zhu, Q. (2025). Research status and development trends of deep reinforcement learning in the intelligent transformation of agricultural machinery. *Agriculture*, 15(11), Article 1223.
<https://doi.org/10.3390/agriculture15111223>
- Zheng, J., Yuan, S., Li, W., Fu, H., Yu, L., & Huang, J. (2024). A review of individual tree crown detection and delineation from optical remote sensing images: Current progress and future. *IEEE Geoscience and Remote Sensing Magazine*, 13(1), 209-236.
<https://doi.org/10.1109/MGRS.2024.3479871>
- Zhou, C., Ge, X., Chang, Y., Wang, M., Shi, Z., Ji, M., ... & Lv, C. (2025). A multimodal parallel transformer framework for apple disease detection and severity classification with lightweight optimization. *Agronomy*, 15(5), Article 1246.
<https://doi.org/10.3390/agronomy15051246>