Journal of Computer Science

Research Article

# Deep Learning and Image Analytics for Forest Land Management: Classification, and Object Detection

**Devam Dixit, Aditya Bhavsar, Shreyas Jani and Nilesh Yadav**

*Department of Computer Engineering, K J Somaiya Institute of Technology, Mumbai, India*

*Article history*
Received: 24-06-2025
Revised: 30-12-2025
Accepted: 20-01-2026

**Corresponding Author**:
Devam Dixit
Department of Computer Engineering, K J Somaiya Institute of Technology, Mumbai, India
Email: d.dixit@somaiya.edu

**Abstract:** Rapid urbanization and industrialization have contributed to deforestation which has led to a severe degradation of the environment and biodiversity loss. Traditional field survey methods are not applicable in the monitoring of large forest areas due to the fact that such methods are manual, costly and time consuming as well as subject to human error. The current methods of remote sensing cannot effectively detect various tree species and the number of trees in large areas as well. To address these issues, this paper introduces a computerized system of tree detection and species identification on high-resolution image data and image analysis methods. The suggested system uses the combination of remote sensing data and deep learning objects detection and classification models to identify, count, and classify tree species using large forest cover. The framework has higher accuracy and faster processing in comparison with traditional surveys and simple approaches of remote sensing. This methodology is useful in the mass evaluation of biodiversity and gives accurate information on the monitoring of forests and land-use studies. The system allows making precise and prompt decisions based on the data to achieve sustainable forest management and conservation of natural resources by facilitating precise and timely counting of trees and identifying the species with the use of the system by environmental scientists, conservationists, and policymakers.

**Keywords:** Tree Enumeration, Image Analytics, Land Management, Conservation, Remote Sensing

## Introduction

Fast industrial development zones along with expanding urban growth produce harmful ecological consequences that strongly affect forests. Human settlements operating inside cleared forest areas generated major biodiversity impacts that led to extensive environmental devastation. Traditional manual forest survey methods fail to monitor transformations because they produce high operational expenses and necessitate long-duration work as well as human mistakes in the data collection. Liu et al. (2019) shows modern monitoring solutions must transform into effective systems for forest ecosystem protection and monitoring since there is now substantial market interest in protective forest technology. In response to this challenge, the proposed research displays a framework that combines high-resolution remote sensing data with intelligent detection algorithms to identify and enumerate tree species. The framework supports conservation planning and biodiversity assessment using reliable data on species distribution and density.

Different automated tree enumeration frameworks and tools require proper recognition during image analytics operations to safeguard the environment through expandable solutions. Previous research data provided by Publisher Roboflow (Mohana et al., 2023) allowed for dataset administration in tree enumeration tasks. The project implements a valuable methodology by incorporating satellite imagery from Google Earth Engine but uses Sentinel data specifically. Tool users benefit from increased resolution of remote sensing data because it makes tree enumeration both precise and scalable thus adding substantial worth to the process.

The tree classification needed raw Convolutional Neural Network (CNN) layers instead of relying on pre-trained models as shown in Persson et al. (2023). The customized approach to our model construction makes it unique to the specific nature of the dataset that leads to better tree classification results. Our approach provides fresh distinctions from the current works while maintaining better specialized methods in tree research using remote sensing.

A brief overview of several research papers and other resources, which were analyzed, are given below. Suteri et al. (2024) talked about enumeration and classification using YOLOv4 and CNN respectively and received good results. Koneru et al. (2018) introduced the factor of Cartosat2 High-resolution Imagery from the Satellite Cartosat2 which helped them in Tree Enumeration. Paper Mohana et al. (2023) introduced the CLAHE method for effective enhancement of image quality. Persson et al. (2023) applied various Deep Learning techniques and Linear Discriminant Analysis for Tree Classification and to achieve a higher level of accuracy while classification.

Waleed et al. (2020) used an automatic scheme for detecting and counting olive trees using red band extraction and Sobel edge detection. Ligade et al. (2024) talked about the utilization of Geographical Information System (GIS) data and Satellite Imagery for distinguishing plants depending on their characteristics such as bark, color, and shape, which enables precise monitoring. Djerriri et al. (2018) used regression-based CNN for palm tree counting which resulted in high accuracy using low effort. Aparna et al. (2018) incorporated high resolution UAV images and used CNN to automatically count coconut trees.

The recent literature has extensively covered the application of data-driven and remote sensing methods in monitoring the environment and analysing infrastructures, as well as assessing trees. Some studies have centered on massive environmental recordings, in which the worldwide pattern of deforestation and local benchmarking were determined by FAO information and satellite records to discern how the forested zone had transformed over an extended period (Safonova et al., 2021). In the same way, the population increase of cities and the consequences of this increase on infrastructures have been discussed with the help of the World Bank and UN data sets with the help of such visualizing sources as OurWorldInData.org (Freudenberg et al., 2022).

Several infrastructure-related studies have employed government databases and standardized surveys to assess the patterns of infrastructure investment, supply chain transparency and national development projects (Weinstein et al., 2020; Zhao et al., 2023). Further to that, research that speaks of infrastructure expansion and the economic impetus has also highlighted the contribution of long-term investments in determining the result of regional and national development (Ventura et al., 2024). Portals of government projects have also been used to obtain insights regarding infrastructure projects under different national programs (Huang et al., 2024).

Advanced deep learning and remote sensing have been adopted in the context of forestry applications. The classification of tree species has been done through active and passive remote sensing data and GIS tools, which have given the opportunity to conduct the ecological assessment accurately (Birla et al., 2025). The performance of the tree classification has also been enhanced using UAV-based RGB images with CNN models that enable automatic features extraction when the high-resolution images are processed (Jiang et al., 2025). In tree counting tasks, deep neural network architectures including Faster R-CNN also have been used with a better performance in terms of accuracy and scalability in forest monitoring tasks (Shukla et al., 2025).

Nonetheless, studies done before investigated different approaches to tree counting and classification like the conventional image processing algorithms like Sobel edge detector to the current trends in object detection and specifically trained classification networks, none of them have integrated modern developments in each of the object detection and tailored classification algorithms. Our project is unique because we have utilized YOLOv8 which is a superior object detector algorithm with a high rate and precision that is able to enumerate trees in a highly effective and accurate manner even on multifaceted forest images. We have also used a trained CNN (Convolutional Neural Network) which is specifically trained to classify tree species and a fine-tuned version has been trained on our data to enable a higher accuracy than when using generic off-the-shelf architectures.

## Methods

### *Architecture Design Specification*

The system flow structure represents the entire working process of our tree discovery and categorization system. The system is divided into a sequence of properly distinct phases, each being of considerable importance in providing a smooth and effective pipeline from data collection to ultimate prediction and visualization. Such phases are modeled to work synchronously with each other, both contributing to the overall performance and accuracy of the model but even to its adaptability and scalability.

One of the primary strengths of this architecture is that it is flexible across various forest ecosystems, making the system operate optimally even when ported to new or heterogeneous environments. By designing the system modularly as done here, we make maintenance easy, improve generalizability, and even future integration of other capabilities such as real-time monitoring or multi-species classification. Fig. 1 shows the in-depth system architecture.

The system flow architecture diagram shows the complete workflow process of our tree enumeration and tree species classification system. When the system initialization occurs, the process activates from beginning to end.

**DOI:** 10.3844/jcssp.2026.1100.1112

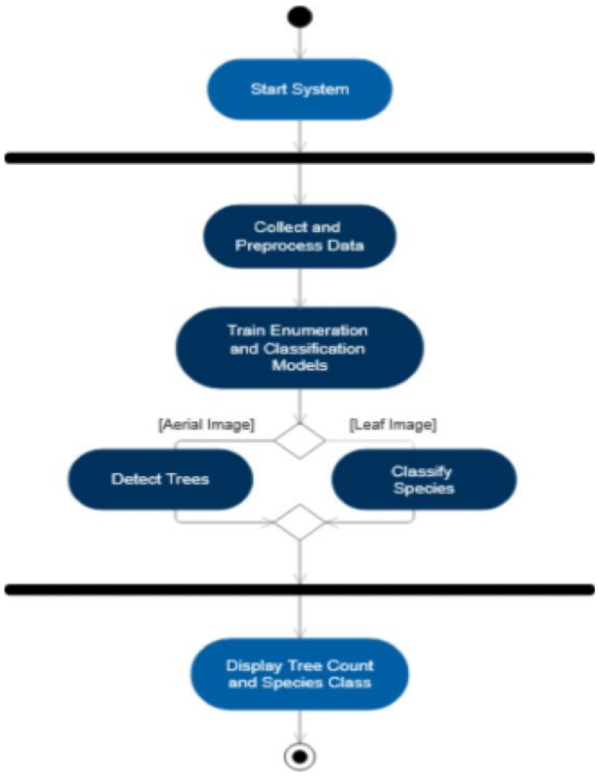


**Fig. 1:** System Design Architecture for Tree Enumeration and Tree Species Classification

The system requires input data collection after starting the initialization process. The data set comprises aerial images dedicated to tree counting and close-up leaf images dedicated to species classification. The preprocessing steps during this stage entail operations for resizing with normalization, noise reduction, and data augmentation techniques that make input images prepared for model training through consistency and clarity.

After data pre-processing, the system goes ahead, to train two independent models, each optimized to handle specific functions: One is for tree enumeration (tree detection) and another for tree species classification. Having a double model strategy enables the system to specialize and become best for every input type. Satellite images are used in the tree detection model to detect and count the number of individual trees found in an input image. Simultaneously, leaf images are fed to the classification model, and it predicts each tree species from exclusive visual characteristics embedded in the leaf patterns. Once both models have made their predictions, the output is passed on to the last step in the pipeline.

The suggested framework is based on the two-step workflow of operations aimed at forest management as opposed to one-to-one identification on the tree level. At the initial phase, the analysis of high-resolution satellite images by the YOLOv8 object detection model allows estimating tree density and spatial distribution of forest areas. This macro level analysis gives fast and scaling data on the number of trees and the possible places of ecological interest.

The second phase involves processing of the ground-level leaf images that have been scanned by a surveyor or field worker with a custom CNN-based classifier to detect tree species. These leaf images are examined on a local scale individually where species level recognition is possible as fine-grained visual information is available.

The conceptual integration of the two modules is done at the decision-support level and not the level of the individual tree. The results of satellite-based enumeration are applied to steer and prioritize the field surveys, whereas the results of species classification are provided by the region-wise aggregation to be used to support the biodiversity assessment and conservation planning. The system does not put the effort to make a given linkage of a specified satellite-detected tree to a specific leaf image as such a connection would need the accurate geolocation labeling as well as the special infrastructure to procure ground-level of leaf image information that would be time-intensive and more intricate.

In the last step, the system prints out the output, presenting the estimated number of trees and the known species for each instance. This user interface enables stakeholders, be they forestry professionals, ecologists, or conservation planners, to take and make use of the results in practical applications.

### *Data Collection*

We only used existing sources for acquiring data throughout the tree enumeration and species classification project. Satellite imagery acts as the fundamental tool for performing our tree enumeration tasks. The data collection took place from Google Satellite Images which provided full visibility of trees across broad forest areas. The primary purpose of this dataset existed to detect separate trees across the distribution space. The satellite images transmitted essential data about both tree distribution density and canopy position along with general health indicators used widely for ecological surveys and resource planning. We manually annotated 500 satellite images and used them for YOLOv8 model training. The tree species classification evaluation required the implementation of Leaf Dataset which contains high-resolution leaf images labeled into different species types. This dataset offered extensive biological diversity because it contained leaf characteristics such as form and texture as well as vein patterns together with different color shades which make species identification possible. We obtained the labeled Leaf Dataset that included in total 3000 images of 10 leaves together with their corresponding species information. The biological features within this dataset were essential for performing tree species identification. We used publicly available datasets to eliminate manual data collection needs while maintaining mentioned levels of diversity and quality which prepared suitable data for model training and data analysis purposes.

The satellite imagery was used to enumerate the trees by using Google Earth Engine (GEE) that was primarily supplied by Sentinel-2 satellite mission. Sentinel-2 has 10 meters spatial resolutions of visible and near-infrared bands in multispectral optical imagery. At a Ground Sample Distance (GSD) of 10 m, individual tree canopies could be identified in the moderately dense forest areas successfully. The selected image was manually inspected and marked in order to generate bounding box labels to train the YOLOv8 object detection model.

The leaf image data that are to be utilized during the classification of the tree species was gathered by the publicly available sources and classified into nine distinct tree species. The data set is composed of approximately 3,000 samples of RGB leaves each of which is a sample of a tree species. The pictures were all brought to the similar resolution and enhanced by the standard image transforming techniques in order to enhance the robustness. The data is not proprietary and it can be reconstructed with the assistance of publicly shared repositories. The data has no known benchmark data such as Flavia or Swedish Leaf.

### *Data Processing*

The model received preprocessed data through two procedures: It first standardized image dimensions through resizing then secondly applied noise reduction to remove artifacts. The data became more suitable for training Deep Learning Models through pixel intensity value normalization procedure. The established methodology allowed rapid tree identification because it extracted important spatial information from space-based image data. The images of the satellites mostly depict tropical and semi-urban forests. Although the dataset includes moderate visual diversity, it is not representative of all biomes of forests in the world.

Through the implementation of data augmentation techniques which included rotating images alongside modifications in flipping and scaling and brightness control adjustments, the model became more robust. The applied transformations expanded the variety of input images thus enabling better generalization abilities when the model confronts previously unknown leaf visuals. The dataset was divided into different subsets for training while validation and testing functions performed error promotion with accuracy determination and precision/recall monitoring.

The bounding box regression loss in YOLOv8 can be calculated using the Generalized Intersection over Union (GIoU) loss function, as shown in Eq. (1):

$$L_{GIoU} = 1 - \left( \frac{|B \cap B_{gt}|}{|B \cup B_{gt}|} - \frac{|C - (B \cup B_{gt})|}{|C|} \right) \quad (1)$$

Where:

$L_{GIoU}$ is Generalized Intersection over Union Loss
$B$ is the predicted bounding box
$B_{gt}$ is the Ground Truth Bounding box
$C$ is the smallest box enclosing both $B$ and $B_{gt}$

The calculation includes *B* for predicted bounding box alongside *Bgt* for ground truth box together with *C* for the smallest box enclosing both *B* and *Bgt.* The first term in Eq. (1) shows the first component in the loss function represents regular Intersection over Union calculation while the second part imposes penalties on predictions which fail to match all together with ground truth annotations:

$$Final\ Test\ Accuracy = \frac{\Sigma\,(TP_i)}{\Sigma\,(TP_i + FP_i + FN_i)} \quad (2)$$

Where:

$TP_i$ are True positives for class $i$
$FP_i$ are the False positives for class $i$
$FN_i$ are the False negatives for class $i$
$i$ is the class index

In Eq. (2), the equation evaluates accuracy through calculation of *TPi* which stands for true positives and *FPi* for false positives as well as *FNi* for false negatives of class *i*. This equation determines the detection accuracy through an analysis of correctly detected objects against all detected and missed objects.

### *Model Selection and Building*

Fig. 2 shows the system diagram of our entire project.

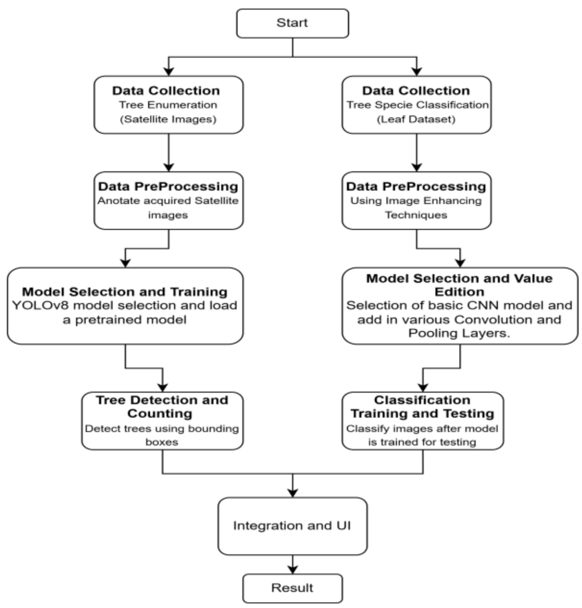


**Fig. 2:** System Diagram for Tree Enumeration and Tree Species Classification using YOLOv8 and Custom CNN

**DOI:** 10.3844/jcssp.2026.1100.1112

The workflow focuses on automating tree enumeration and species classification using various Deep Learning techniques.

The entire process begins with data collection from two sources: Satellite images for tree enumeration and a leaf dataset for species classification. Both these datasets undergo preprocessing - satellite images are annotated to label trees whereas the leaf images are enhanced to improve clarity.

We have selected two very different models for both our tasks after extensive research from previous works and papers, we decided that YOLOv8 would be the best for Tree Enumeration and CNN would excel in classification. Upon selection, the YOLOv8 model architecture was loaded and left to be trained on a blank slate on the annotated satellite dataset to recognize and estimate the number of trees in the form of bounding boxes. For Species Classification, we built a custom CNN model by adding multiple convolution and pooling layers to refine feature extraction.

The Model YOLOv8 that was applied in this study to enumerate trees was trained without the transfer learning. This model was launched on the YOLOv8 network structure file and the only data it was trained on was the manually annotated set of satellite imagery data retrieved through Google Earth Engine. There were no pre-trained weights so that the model would learn task-specific features that are relevant in task canopy tree detection and not pass on generic object representations on large-scale benchmark datasets. This training approach enables equal consideration of the ability of YOLOv8 to fit domain-specific forestry data and it prevents bias that object classes that are pre-trained can introduce.

After training, tree enumeration was performed by detecting and counting trees from satellite images, while the CNN model underwent testing. Once both the models were ready, we integrated them. The Final output provides an efficient and automated method to enumerate trees and classify them, aiding in environmental monitoring, forest management, and conversation efforts. The entire training was done through Google Collaboratory. The training of the YOLOv8 tree enumeration model and the special CNN tree species classification model was performed using an NVIDIA Tesla T4 graphics card having a 16 GB graphics card memory and the benchmark of 12-13 GB system RAM. The YOLOv8 model used has been trained on 100 epochs using the input image size of 640 × 640 and finished the training process in a comparatively short amount of time as it is a lightweight architecture and its use of the GPU is optimized. On the contrary, the custom CNN model used to classify tree species needed more time to compute and took 80 epochs to train, averaging a training time of about 300 seconds per epoch.

### *User Interface Design*

After training and testing our model, we achieved satisfactory results and decided to create a website using React JS and Fast API for backend with integrated model for tree enumeration and tree classification.

For the frontend, in our environment we created a react app with all the required files. Our website contains three primary functions, firstly, the website should have an image uploading feature, secondly, a choice amongst tree enumeration and classification is to be provided and finally, the output from the model needs to be displayed. To handle our uploading feature, we created a function named handle submit that would append a file from the user and pass it to the model. For the second task of choice selection, we use the same handle submit function to pass model type. We saved our model architectures in a separate .py file in a backend folder, also we saved the best points achieved from training the model into a model's folder which contained weight.

The flow of our system in phase 3 is as follows, firstly the user will upload the image and press the submit button which would load the image to our system then the user will select whether the task is of enumeration or classification, from a radio button, which will be passed along with the backend. The backend runs on Fast API which will take in the request from the user and if the user has selected enumeration the image will be passed on the YOLO model which is internally routed and connected to its best point which contains weights that gave us the best testing results.

Fig. 3 shows the User Interface (UI) developed for tree enumeration and tree species classification. The UI allows users to upload images and view the model's results (processed images), including the number of trees detected and the identified species based on leaf images. It provides an easy-to-use platform for applying our models in real-time, making tree monitoring and classification accessible for forest management and environmental studies.

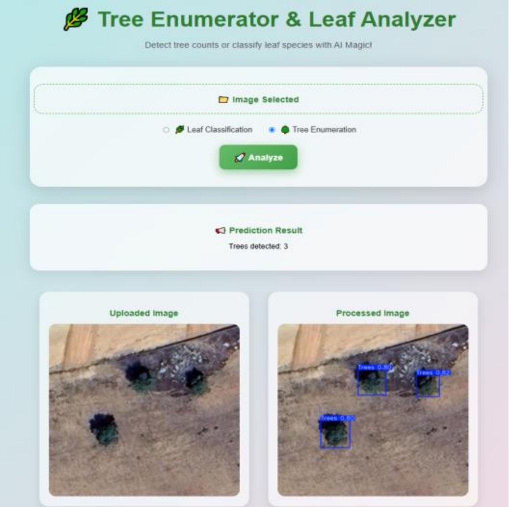


**Fig. 3:** User Interface for tree enumeration and tree classification

## Results

### *Metrics*

It has been verified that the proposed tree enumeration and species classification system operated through YOLOv8 enumeration followed by Custom CNN model classification. The enumeration phase of YOLOv8 delivered high accuracy and efficiency throughout 100 epochs which resulted in testing accuracy reaching 0.8665 and F1 score attaining 0.854. The model operated at 0.9108 precision and 0.803 recall levels when set to an Intersection over Union threshold of 0.5. The model showed strong results in precision evaluation through its mAP values which reached 0.8732 for mAP@50 and 0.4648 for mAP@50-95.

Table 1 shows a detailed overview of the results that were achieved after training our dataset using the YOLOv8 model. The model was trained on a well-curated dataset and fine-tuned in multiple iterations to ensure optimal performance. Regular evaluation and parameter tuning were performed during the training process to improve the model's performance in detecting and classifying instances of trees. As demonstrated in the table, the model exhibits high performance metrics in all the evaluation indicators, which reflects the quality of the training strategy and data.

The model performed outstandingly with 86% test accuracy, an F1 score of 85%, precision of 91%, recall of 80%, and a mAP@50 of 87%. These outcomes demonstrate the efficiency of the model in object detection, with a good balance between accuracy, precision, and recall, and verifying its trustworthiness for practical detection and classification purposes.

Using 80 epochs within the Custom CNN model training process produced 0.88 training accuracy and 0.87 testing accuracy. The F1 score reached 0.87 while precision and recall values measured at 0.88 and 0.87 which show the system maintained a consistent and balanced classification performance for tree species identification. The combined system consisting of YOLOv8 enumeration abilities with Custom CNN classification capabilities provides an efficient method.

The model performed outstandingly with 86% test accuracy, an F1 score of 85%, precision of 91%, recall of 80%, and a mAP@50 of 87%. These outcomes demonstrate the efficiency of the model in object detection, with a good balance between accuracy, precision, and recall, and verifying its trustworthiness for practical detection and classification purposes.

Using 80 epochs within the Custom CNN model training process produced 0.88 training accuracy and 0.87 testing accuracy. The F1 score reached 0.87 while precision and recall values measured at 0.88 and 0.87 which show the system maintained a consistent and balanced classification performance for tree species identification. The combined system consisting of YOLOv8 enumeration abilities with Custom CNN classification capabilities provides an efficient method.

**Table 1:** Dashboard indicating results of YOLOv8 for Tree Enumeration using Satellite Imagery

| Parameter | Value |
|---|---|
| Model | YOLOv8 |
| Epochs | 100 |
| Testing Accuracy | 0.8665 |
| F1 Score | 0.854 |
| IoU Threshold | 0.5 |
| Save Period | 1 |
| Precision | 0.9108 |
| Recall | 0.803 |
| mAP@50 | 0.8732 |
| mAP@50-95 | 0.4648 |
| Inference Speed | 2.1 ms (per image) |
| Post-Processing Speed | 1.1 ms (per image) |

Referring to Table 2, we analysed that 80 epochs were best fit for our epoch training after evaluating results from various other epoch measures. On our best fit epoch, we achieved a 87% testing accuracy without any overfitting, which suggested that the entire model is balanced and performing well.

The identification of the species of trees is conducted with the assistance of a specially designed Convolutional Neural Network (CNN), being fine-tuned to the leaf image analysis. The network captures the RGB leaf images with the size of 128 by 128 pixels. It has three convolutional blocks, which are followed by the batch normalization, ReLU activation and max-pooling to extract hierarchical features sequentially.

According to Table 3, the convolutional layer is the first layer with 16 3 3 filters and padding to maintain the spatial dimension, then there is the batch normalization and 2 2 max-pooling layers. The second convolutional layer localizes the feature to 32 filters whose kernel size and pooling are the same. The third convoluted layer further increases the representation to 64 filters which allows the model to detect complex patterns of leaves, like venation and texture.

The resulting feature maps are flattened and passed through a 128 neuron fully connected layer with ReLU activation. The final output layer that is used does multi-class classification in nine tree species using a softmax activation. The model is optimized on Adam optimizer and the training is grounded on the cross-entropy loss.

An object detection model's training throughout enables observing how Precision, Recall, mAP@0.5 and mAP@0.5:0.95 components evolve during epochs through this graphical representation shown in Fig. 4.

**Table 2:** Dashboard indicating results of Custom CNN model for Tree Classification using Tree Leaf Dataset

| Parameter | Value |
|---|---|
| Model | Custom CNN |
| Epochs | 80 |
| Training Accuracy | 0.88 |
| Testing Accuracy | 0.87 |
| F1 Score | 0.87 |
| Precision | 0.88 |
| Recall | 0.87 |

**DOI:** 10.3844/jcssp.2026.1100.1112

**Table 3:** CNN Model Architecture for Tree Species Classification

| Layer No. | Layer Type | Parameters | Output Size |
|---|---|---|---|
| Input | RGB Image | 128 × 128 × 3 | 128 × 128 × 3 |
| 1 | Conv2D + Batch Normalization + ReLU | 16 filters, 3 × 3 kernel, padding = 1 | 128 × 128 × 16 |
| 2 | Max Pooling | 2 × 2 | 64 × 64 × 16 |
| 3 | Conv2D + Batch Normalization + ReLU | 32 filters, 3 × 3 kernel, padding = 1 | 64 × 64 × 32 |
| 4 | Max Pooling | 2 × 2 | 32 × 32 × 32 |
| 5 | Conv2D + Batch Normalization + ReLU | 64 filters, 3 × 3 kernel, padding = 1 | 32 × 32 × 64 |
| 6 | Max Pooling | 2 × 2 | 16 × 16 × 64 |
| 7 | Flatten | — | 16,384 |
| 8 | Fully Connected (FC) + ReLU | 128 neurons | 128 |
| 9 | Output Fully Connected Layer | 9 neurons | 9 |
| Output | Softmax (via Cross-Entropy Loss) | — | Class probabilities |

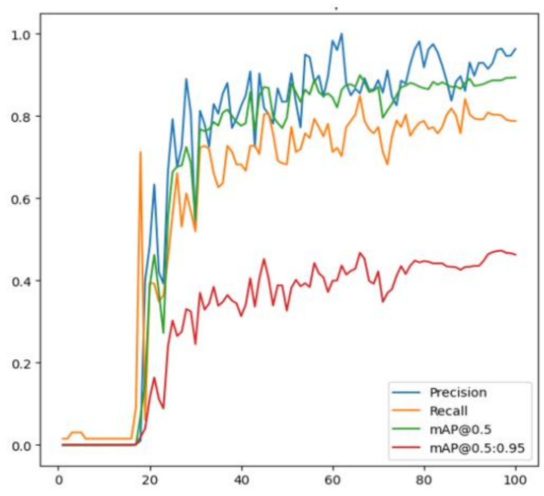


**Fig. 4:** All metrics over Epoch Graph for Tree Enumeration using Yolov8

The model achieves effective learning between epochs 20 and 21 since all metrics rise from their starting points near zero to peak values. The training process guarantees stable growth for both precision and mAP@0.5 since precision achieves values greater than 0.9 which indicates high-quality predictions. mAP@0.5:0.95 shows steady growth before reaching a stable point at 0.45 while the system recall stabilizes at 0.8.

The Precision Recall (PR) curves shown in Fig. 5, assessment reveals the operational behavior of multi-class classifiers between different tree species grouped into 11 categories. The assessment leads to vertical movement of curves toward the top-right section because when precision levels increase, recall levels also increase. The performance curves of "Apta" and "Indian Rubber Tree" graph near perfect levels. The tree identification accuracy for species “Kashid” and “Pimpal” remains problematic or unpredictable due to the difficulty of their classification. The model performs well with certain species but processing difficulties emerge for extra species because the training data has inadequate class distribution and overlapping features among species.

Fig. 6 is a precision recall F1-score bar plot for all ten of the same classes. Precision (blue bars) is the ratio of correctly predicted instances to all predicted instances, recall (orange bars) is the ratio of correctly predicted instances to actual instances, and the F1-score (green bars) is the harmonic mean of the former two. Most of the classes achieve a high score on all three parameters, generally above 0.85 which indicates even and effective classification. For example, "Nilgiri" and "Pimpal" show consistently high scores, indicating faith in the models and good recall in identifying such trees. There are some classes, such as "Karanj" and "Sita Ashok," where there are small variations where precision or recall is below 0.9, indicating the possibility of rare false alarms or missed detections.

Fig. 7 depicts the Per-Class Accuracy of a classification model on various tree types. Each bar represents the accuracy for an individual class, varying from 0.8 to nearly 1.0. The best class is Vilayati Chinch with nearly 98–99% accuracy, and this is followed by Apta, Somnohar, and Kashid, each having accuracies above 0.93. On the other hand, the classes Karanj and Sita Ashok show the lowest accuracy, both falling in the 80% range. This suggests that even though the model is doing perfectly well on a few classes, there is some room for better differentiation between the more challenging or visually confused classes.

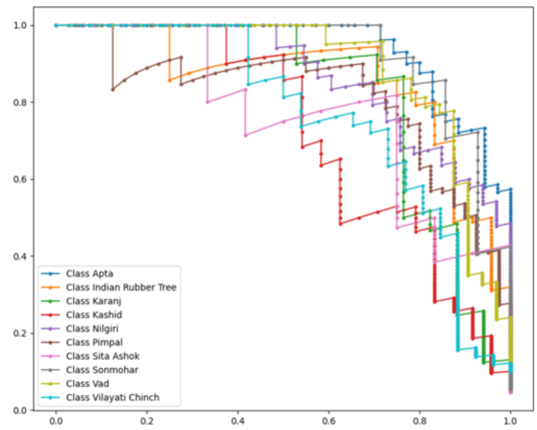


**Fig. 5:** Precision Over Recall Curve Graph for Tree Classification using Custom CNN using Leaf dataset

**DOI:** 10.3844/jcssp.2026.1100.1112

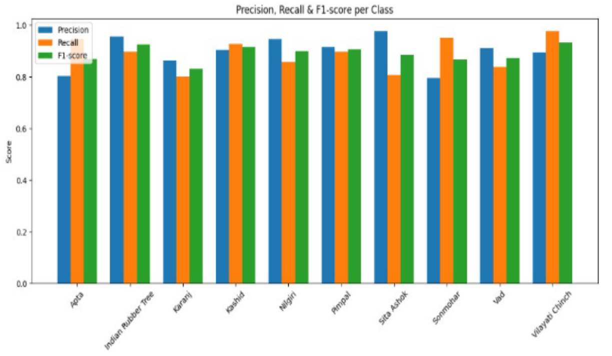


**Fig. 6:** Per Class Classification Graph for Precision, Recall and F1 Score per Tree Species on leaf dataset

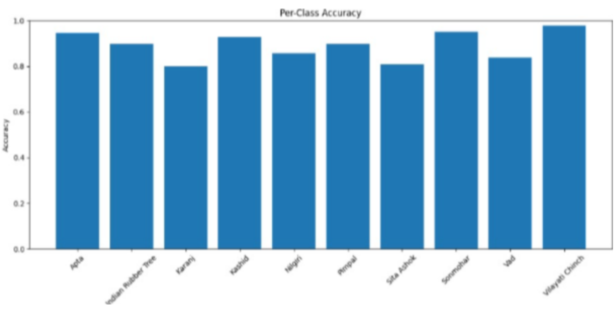


**Fig. 7:** Bar Graph for Testing Accuracy of Prediction for Tree Species Classification

### *Performance*

Two different tests established the tree identification system by using YOLOv8 enumeration along with Custom CNN model classification. The sequence of two processing operations generates complete data including tree numbers along with type details to perform forest land diversion assessments.

During 100 epoch development of the YOLOv8 model researchers obtained testing results with 0.8665 accuracy alongside 0.854 F1 score. The model displayed exceptional ability to detect and position trees accurately when using an IoU threshold of 0.5 because it delivered precision results at 0.9108 and recall results at 0.803. Analyzed precision stats were remarkably important because they reached 0.8732 mAP@50 and 0.4648 mAP@50-95. YOLOv8 demonstrates remarkable efficiency features through its 2.1 ms image inference speed in addition to 1.1 ms post-processing time to enable real-time forest monitoring applications. Fig. 8 shows Bounding boxes for each tree count or enumeration using Yolov8 by our Yolov8 model. After 30 epochs of training the Custom CNN model reached 0.9538 training accuracy which converted into 0.7756 testing accuracy. The evaluation metrics of F1 score precision achieved 0.78 yet recall provided a result of 0.79 but F1 score registered 0.78 at best.

Although the model succeeds in identifying difficult patterns its testing accuracy decrease indicates that it avoids becoming overly specialized in producing precise data predictions for unfamiliar observations. The integration of both phases produces results which confirm the effectiveness of implementing the proposed system.

The system enables quick tree density evaluations and accurate species recognition to boost protection of biodiversity as well as environmental influence study. YOLOv8 acceleration together with detection speed operating alongside the CNN-based classification abilities leads to the comprehensive assessment of land-diverted forest territories. Fig. 9 displays the successful identification of the test Image which was classified as Sita Ashok.

With the results shown in above figs, we can infer that our Tree Enumeration model that is YOLOv8 accurately counts the number of trees in a single satellite image which can help the government bodies or private companies to ensure proper land management with the help of satellite image.

Our Custom made CNN model successfully classifies 10 classes provided in our data which helps in identifying the endangered or rare species in a particular area which can help the government bodies or private companies to ensure proper land management with the help of Tree Species Leaf Images.

The confusion matrix in Fig. 10 shows a high level of diagonal dominance which shows that the model has a good level of classification accuracy on most of the tree species. Apta, Indian Rubber Tree, Pimpal, Sita Ashok and Vilayati chinch are examples of species with high true positive and low false classification. There is some little confusion between visual similar species like Karanj and Somohar and Nilgiri and Kashid which can be understood as a result of similarities in the texture and form of leaves. On the whole, the small off-diagonal values present the strength of the CNN model regarding the separation of various tree species with respect to different visual features.

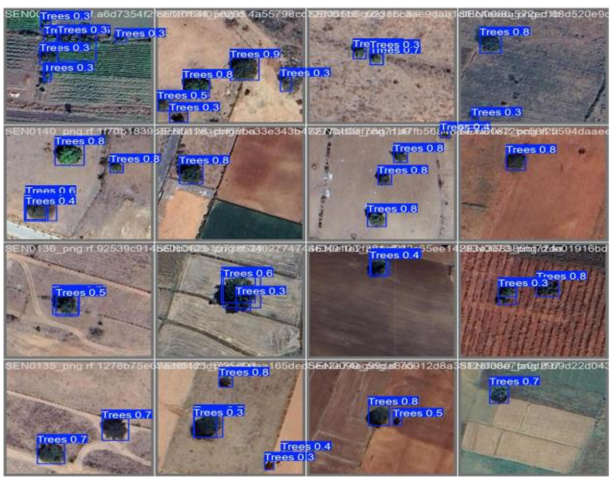

**Fig. 8:** Tree Enumeration using YOLOv8 showcasing Multiple Test images with Bounding boxes provided by our YOLOv8 model

**DOI:** 10.3844/jcssp.2026.1100.1112

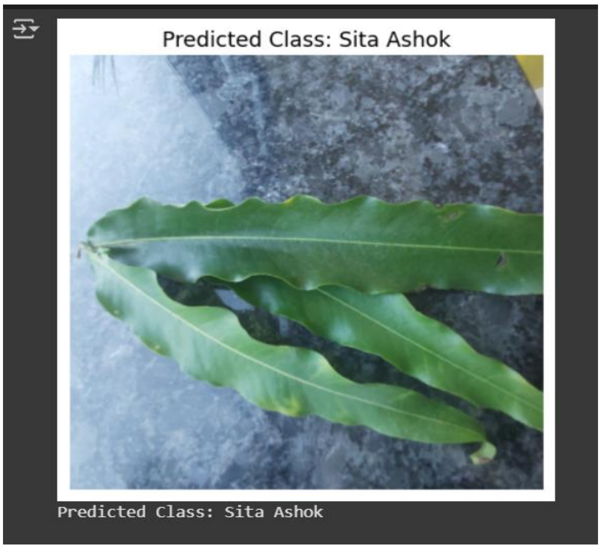


**Fig. 9:** Predicted a test Image on Custom CNN model which successfully classified the image as Sita Ashok

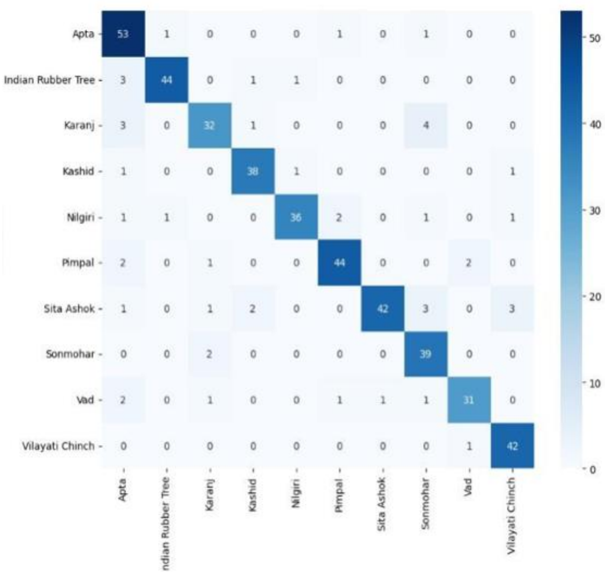


**Fig. 10:** Confusion matrix of the CNN-based tree species classification model showing class-wise prediction performance across ten tree species

## *Comparative Analysis*

Comparative analysis is a systematic process utilized to compare and examine the performance, attributes, or outcomes of two or more entities with similar conditions. In computer vision and machine learning, this approach is particularly effective in comparing different models or algorithms for deciding which one gives better performance in terms of accuracy, efficiency, and reliability. By comparing multiple metrics such as accuracy, precision, recall, F1-score, and mean Average Precision (mAP), researchers can establish the strengths and weaknesses of each model.

Our Comparative Analysis begins with interpreting the results of the YOLOv8 and YOLOv5 models, which were trained for Tree Counting and Tree Detection. With the help of above comparative analysis, we can scrutinize which of the following models performs better and it justifies the selection of the most suitable model for our project. Table 4 shows the Comparative analysis between the YOLOv8 and YOLOv5 models where YOLOv8 always performs better in most of the key areas. With 86.65% testing accuracy compared to 70.7% for YOLOv5, YOLOv8 greatly outshines its overall detection capability. Similarly, YOLOv8 also reports a marginally higher F1-score of 0.867 versus 0.839, reflecting a more balanced trade-off between recall and precision. Accuracy is ever so slightly higher in YOLOv8 (0.9108) versus YOLOv5 (0.902) and is also more effective at recall (0.803 vs. 0.773), reflecting that YOLOv8 is more effective at labeling objects of interest correctly. With regards to localization precision, YOLOv8 is superior with a higher mAP@50 (0.8732) compared to 0.829 for YOLOv5, but YOLOv8's mAP@50-95 is marginally inferior to YOLOv5's (0.4648 vs. 0.474), indicating that YOLOv5 may be better at handling more challenging IoU thresholds. This compromise is negligible, however, and YOLOv8 in general is the more precise and stable model for object detection tasks in this comparison.

Models based on deep learning and YOLO have been used in forestry applications in recent studies starting in 2021. Hani et al. (2023) tested the idea of counting trees with the help of UAV imagery on the example of YOLOv5, Aparna et al. (2018); Djerriri et al. (2018) tested their ideas on the basis of CNN with the purpose of counting trees, but in small areas or with more complex computation problems. Scalable canopy analysis demonstrated by Zurqani (2021) has been applied to Google Earth Engine but without any modern object detection frameworks like YOLOv8. More recent studies by Mahule et al. (2024); Ligade et al. (2024) devoted attention to the tree enumeration of forest land diversion without incorporating the classification on a species level.

**Table 4:** Dashboard indicating comparative results for Tree Counting and Tree Detection

| Parameters | Values | |
|---|---|---|
| Model | YOLOv8 | YOLOv5 |
| Epochs | 100 | 100 |
| Testing Accuracy | 0.8665 | 0.707 |
| F1_Score | 0.867 | 0.839 |
| IoU Threshold | 0.5 | 0.5 |
| Save Period | 1 | 1 |
| Precision | 0.9108 | 0.902 |
| Recall | 0.803 | 0.773 |
| mAP@50 | 0.8732 | 0.829 |
| mAP@50-95 | 0.4648 | 0.474 |

**DOI:** 10.3844/jcssp.2026.1100.1112

The proposed framework, in its turn, uses YOLOv8 with satellite imagery of Sentinel-2 to count the number of trees in large areas and integrates it with a local CNN to label the species, thus taking recent studies using IOLO in forestry to the next level of scalable large-scale biodiversity assessment.

After completing the Tree Detection and Enumeration task, we proceeded with Tree Species Classification by using a Custom CNN model. The Comparative Analysis below represents performance comparison of the Custom CNN model with popular pre-trained models, that is, MobileNet V2 and ResNet-50, in classifying the tree species.

The test accuracy of the Custom CNN is the greatest (87) as compared to MobileNet V2 (81.28) and ResNet-50 (82.02). This is because the Custom CNN is a task-designed architecture designed in such a way that it is fine-tuned to detect fine-grained leaf features like venation structure, texture, and shape differences yielding critical tree species discriminatory information. Although the model has a moderate number of parameters of 2.12 million, they are in a compact architecture that fits the domain of the target instead of using a large generic representation featuring features.

MobileNetV2 has minimal parameters (0.01 million) that are supposed to be mobile and edge-friendly. Even though this enables it to be efficiently calculated, its lower representational capability limits its visual comparability of leaf species therefore reducing its classification precision. In this setup, ResNet-50 is incompatible with more time-constrained or resource-limited forestry applications, since it has a significantly higher computation cost (1.5686 s) and has a substantially smaller parameter space (0.02 million).

Generally, Custom CNN provides a reasonable trade-off among accuracy, complexity of model and inference efficiency. Being more than two times faster than the ResNet-50 (0.1513 s) and more accurate than both of the baseline models, it indicates that a well-constructed domain-specific CNN can be used effectively in the specialized forestry task compared to generic pre-trained models. Such a balance can be used to justify the fact that the proposed Custom CNN architecture was chosen to classify the tree species (Table 5).

**Table 5:** Dashboard indicating comparative results for Tree Species Classification

| Parameters | Values | | |
|---|---|---|---|
| Model | Custom CNN Model | MobileNet V2 | ResNet-50 |
| Test Accuracy | 0.87 | 0.81 | 0.82 |
| Params (M) | 2.12 | 0.01 | 0.02 |
| Inference Time (s) | 0.1513 | 0.2433 | 1.5686 |

## Discussion

The implementation of personal data obtained in Google Earth Engine led to a higher detection accuracy since high-resolution and uniform satellite data were available. This confirms the choice of Sentinel-2 imagery on Google Earth Engine to identify trees as the quality of input data is directly proportional to the ability to find objects more accurately and minimize false alarms. The correct classification of trees under the proposed framework once again proves that the methods relying on deep learning could be a helpful tool in conservation and habitat evaluation provided that they are trained on the datasets that are related to the specific domain.

The acquired performance measures can be considered to be effective as compared to the literature. Earlier studies had found accuracy values to be between 48.6 to 96.6 with the average accuracy of about 74.8% in varied experimental setups (Devidas et al., 2024). Conversely, the framework suggested demonstrated better robustness and generalization as it better and more consistently performed on tree enumeration and species classification. Professor M. Waleed et al. claimed a mean accuracy of 90 percent with multiple limb position analysis but their method had the disadvantage of lower reliability in classification because of the variations in poses (Waleed et al., 2020). In the current work, these constraints are overcome through the use of satellite images to enumerate and the use of leaf-based visual characteristics to classify, which means a higher degree of adaptability to diverse environmental factors.

The proposed approach is also justified by the existing studies. On satellite imagery, Suteri et al. (2024) achieved 85% of the accuracy with machine learning-based image segmentation but did not enter species-level identification (Suteri et al., 2024). Koneru et al. (2018) indicated that with the Cartosat-2 satellite images, the detection rate was 78% and only the large trees could be detected, but not classified (Koneru et al., 2018). Mahule et al. (2021) reached 88 percent accuracy with the UAV imagery and CNN-based analysis, but their method was restricted to structural evaluation and did not enumerate large-scale on an automated basis (Mahule et al., 2024). In contrast to these works, this framework is the only one to utilize both big-scale tree counting and species classification, which provides a more extensive and scalable option to forest monitoring.

The practical implications of the suggested system support the applicability of this system to the real-world implementation of forest management and policy-making. Quickly estimating the density of trees will help the government agencies and the planning authorities to estimate the level of forest cover prior to the approval of the land diversion projects thereby mitigating the risk to the environment. Also, species level classification aids in the evaluation of biodiversity and to determine ecologically sensitive species or endangered species within an area. Whereas this research is confined to the analysis of statistical images and does not take into consideration the optimization of embedded satellite software and real-time management of faults, the obtained results prove that AI-based surveying can deliver more accurate and feasible insights compared to the traditional manual survey. The dataset size (500 satellite images and 3,000 leaf images) was big enough to converge the model but would be insufficient to generalize the results on a global scale. All the above reasons substantiate the offered framework as an efficient instrument of sustainable forest management and decision-making in the environment.

## Conclusion

The project uses high-resolution satellite imagery for the purpose of enabling faster counting of trees on a wider scale and with greater precision in the context of development. The tree species identification is also used in order to enable more precise ecological analysis and enhanced biodiversity conservation practices. The tree species identification enables the system to supply information regarding ecosystem health and enable focused conservation practices. Satellite integration with the machine learning module presents real-time status during tree density time and is contributory towards effective decision-making to planning authorities, developers, as well as environmental authorities. The system is not impacted by human errors and offers prompt response during the implementation of conventional surveys of trees and thus optimizes forestry management to be efficient and effective. The suggested system can be deployed in real world owing to the small inference latency and moderate computation needs. The framework is practical because YOLOv8 can be used to perform analysis in near real-time, whereas the CNN model can be used on standard GPUs. In Future Scope we can use our framework to further achieve the same objective of improving its autonomy, scalability and real-world applicability by incorporating direct camera-based image acquisition in the web platform. The system can also be expanded in the future to enable the real-time image capture with device cameras, mobile phones, drones, or fixed monitoring cameras where manual uploading of images is eliminated and data is collected on site. This would enable the officials and field person in charge of forests to take canopy/leaf images in real time and get instant results on tree enumeration and species classification with the use of the deployed models. Also, the automated and periodic monitoring of the forest areas on the large scale can be assisted with the help of UAV-based imaging pipelines, which will help to detect the deforestation, illegal logging and seasonal changes of the vegetation. In order to enhance self-sufficiency, the trained CNN model can be deployed on edge computing devices in its lightweight version to conduct offline inferences in isolated forest areas with weak connectivity. To add geospatial tagging and time analysis, the system can be extended to provide predictions with GPS coordinates and time stamps which will enable long-term monitoring of the tree density and distribution of tree species. Future developments could involve incorporation of multispectral data, evaluation of tree health and carbon estimation to make the framework a complete autonomous decision-support system in the management of sustainable forests.

## Acknowledgment

Thank you to the publisher for their support in the publication of this research article. We are grateful for the resources and platform provided by the publisher, which have enabled us to share our findings with a wider audience. We appreciate the efforts of the editorial team in reviewing and editing our work, and we are thankful for the opportunity to contribute to the field of research through this publication.

We would like to sincerely thank our mentor, Dr. Nilesh Yadav, and his amazing guidance, knowledge, and constant availability, without whose enormous contribution, this study would not have been made possible. The mentorship of Dr. Nilesh Yadav was pivotal for helping in setting the path when it comes to methodology choice, interpreting results, and the help that came in making this overall project look better and highly influential to the outcome of the topic.

## Funding Information

The authors have not received any financial support or funding to report.

## Author’s Contributions

All the authors contributed to the preparation, writing and publication of this manuscript equally.

## Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

## References

Aparna, P., Hebbar, R., Harshita, M. P., Sounder, H., Nandkishore, K., & Vinod, P. V. (2018). *CNN-Based Technique for Automatic Tree Counting Using Very High Resolution Data*. 127–129. https://doi.org/10.1109/icdi3c.2018.00036

Birla, L., Bharadwaj, A., Jain, R., Deb, C. K., Sehgal, V. K., & Ramasubramanian, V. (2025). Deep learning model BiFPN-YOLOv8m for tree counting in mango orchards using satellite remote sensing data . *Scientific Reports*, *15*(1), 33791. https://doi.org/10.1038/s41598-025-97562-7

Devidas, D., Sumit, A., Swaroop, K., & Pushpak, N. (2024). Application of Image Analytics for Tree Enumeration for Diversion of Forest Land. *International Journal for Multidisciplinary Research*, *6*(1). https://doi.org/10.36948/ijfmr.2024.v06i01.11112

Djerriri, K., Ghabi, M., Karoui, M. S., & Adjoudj, R. (2018). Palm Trees Counting in Remote Sensing Imagery Using Regression Convolutional Neural Network. *Proceedings of the 2018 IEEE International Geoscience and Remote Sensing Symposium*, 2627–2630. https://doi.org/10.1109/igarss.2018.8519188

Freudenberg, M., Magdon, P., & Nölke, N. (2022). Individual tree crown delineation in high-resolution remote sensing images based on U-Net. *Neural Computing and Applications*, *34*(24), 22197–22207. https://doi.org/10.1007/s00521-022-07640-4

Hani, U., Munir, S., Younis, S., Saeed, T., & Younis, H. (2023). Automatic Tree Counting from Satellite Imagery Using YOLO V5, SSD and UNET Models: A case study of a campus in Islamabad, Pakistan. *Proceedings of the 2023 IEEE International Conference on Artificial Intelligence*, 88–94. https://doi.org/10.1109/icai58407.2023.10136679

Huang, Y., Ou, B., Meng, K., Yang, B., Carpenter, J., Jung, J., & Fei, S. (2024). Tree species classification from UAV canopy images with deep learning models. *Remote Sensing*, *16*(20), 3836. https://doi.org/10.3390/rs16203836

Jiang, T., Freudenberg, M., Kleinn, C., Lüddecke, T., Ecker, A., & Nölke, N. (2025). Detection transformer-based approach for mapping trees outside forests on high resolution satellite imagery. *Ecological Informatics*, *87*, 103114. https://doi.org/10.1016/j.ecoinf.2025.103114

Koneru, S. V., Arul Raj, M., Padmaja, M., Kollu, P. K., Bokinala, L., Ravi Raja, A., & Jitendra, A. (2018). Detection and Enumeration of Trees using Cartosat2 High Resolution Satellite Imagery. *Published in Proceedings of the 2018 International Conference on Advances in Remote Sensing*, 1–6. https://doi.org/10.1109/icares.2018.8547072

Ligade, D., Das, C., Parte, R., Kulkarni, M., Jadhav, S., & adeep, A. M. (2024). Satellite Image Analytics for Tree Enumeration for Diversion of Forest Land. *ICT Analysis and Applications*, 278–287. https://doi.org/10.1007/978-3-032-06694-7_26

Liu, C.-C., Chen, Y.-H., Wu, M.-H. M., Wei, C., & Ko, M.-H. (2019). Assessment of forest restoration with multitemporal remote sensing imagery. *Scientific Reports*, *9*(1), 7279. https://doi.org/10.1038/s41598-019-43544-5

Mahule, S., Pranathi, T., Ranjan, T. B., Kiran, T. A., & Siddhartha, A. (2024). Intelligent Forest Assessment: Advanced Tree Detection and Enumeration with AI. *International Journal of Research Publication and Reviews*, *5*(5), 10597–10603.

Mohana, S., Selvi, R. S., D.S., L. M., Felicia, D. P., & Kiruthika, I. S. (2023). Image Analytics for Tree Enumeration for Diversion of Forest Land. *International Research Journal of Modernization in Engineering Technology and Science*, *5*(9), 3260–3266.

Persson, H. J., Axelsson, C., Mukhopadhyay, R., Huo, L., & Holmgren, J. (2023). Comparison of Single Tree Species Classification Using Very Dense ALS Data or Dual-Wave ALS Data. *Published in Proceedings of the 2023 IEEE International Geoscience and Remote Sensing Symposium*, 3066–3069. https://doi.org/10.1109/igarss52108.2023.10281873

Safonova, A., Hamad, Y., Dmitriev, E., Georgiev, G., Trenkin, V., Georgieva, M., Dimitrov, S., & Iliev, M. (2021). Individual Tree Crown Delineation for the Species Classification and Assessment of Vital Status of Forest Stands from UAV Images. *Drones*, *5*(3), 77. https://doi.org/10.3390/drones5030077

Shukla, P. K., Veerasamy, B. D., Alduaiji, N., Addula, S. R., Pandey, A., & Shukla, P. K. (2025). Fraudulent account detection in social media using hybrid deep transformer model and hyperparameter optimization. *Scientific Reports*, *15*(1), 38447. https://doi.org/10.1038/s41598-025-24326-8

Suteri, H., Sharma, S., Choudhary, V., Goyal, N., & Kalonia, S. (2024). Forest Scan: An Image Analytics for Tree Enumeration for the Diversion of Forest Land. *2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE)*, 1–5. https://doi.org/10.1109/ic3se62002.2024.10593225

Ventura, J., Pawlak, C., Honsberger, M., Gonsalves, C., Rice, J., Love, N. L. R., Han, S., Nguyen, V., Sugano, K., Doremus, J., Fricker, G. A., Yost, J., & Ritter, M. (2024). Individual tree detection in large-scale urban environments using high-resolution multispectral imagery. *International Journal of Applied Earth Observation and Geoinformation*, *130*, 103848. https://doi.org/10.1016/j.jag.2024.103848

Waleed, M., Um, T.-W., Khan, A., & Ahmad, Z. (2020). An Automated Method for Detection and Enumeration of Olive Trees Through Remote Sensing. *IEEE Access*, *8*, 108592–108601. https://doi.org/10.1109/access.2020.2999078

Weinstein, B. G., Marconi, S., Aubry-Kientz, M., Vincent, G., Senyondo, H., & White, E. P. (2020). DeepForest: A Python package for RGB deep learning tree crown delineation. *Methods in Ecology and Evolution*, *11*(12), 1743–1751. https://doi.org/10.1111/2041-210x.13472

**DOI:** 10.3844/jcssp.2026.1100.1112

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

Zurqani, H. A. (2021). Using ML and Google Earth Engine for Forest Canopy Dataset. *IEEE Access*, *10*, 34000-34010.