RESEARCH PAPER

Tree Canopy Detection Using Artificial Intelligence Based - Deep Learning Approach

Rohit R. Vibhute[1*], **Dr. Seema S. Patil**[2*]

[1*]PG, Scholar, Department of Electronics and Telecommunication Engineering, Ashokrao Mane Group of Institutions, Vathar, Kolhapur, Maharashtra, India.
[2*]Associate Professor, Department of Electronics and Telecommunication Engineering, Ashokrao Mane Group of Institutions, Vathar, Kolhapur, Maharashtra, India.

[*]**Corresponding Author**: rohitvibhute040@gmail.com, seema.patil9@gmail.com

ABSTRACT

The structure of the plant canopy seems to play a very important role in shaping the microclimate around the crops. This is done by influencing the evaporation, light interception, and overall crop growth, which greatly depends on the canopy. It is also a very important indicator of biomass accumulation and directly related to yield potation. Due to this accurate information about the canopy presence and the size of the canopy is very much essential for environmental planning. The conventional approaches which are implemented for canopy detection such as manual field surveys and hemispherical photography are usually time-consuming and labour-intensive. This means they are often affected by the observer bias. Although various techniques such as remote sensing, which use satellite imagery or LiDAR exist and can cover large areas, they are typically limited due to spatial resolution and high operational costs. Over the recent years, we have seen a large technological progress in the field of deep learning and computer vision. This has opened infinite new possibilities for automated canopy analysis using ground-level images. In this paper, we discuss a lightweight and efficient framework for plant canopy detection and size estimation. A mobile net SSD-based convolutional neural network is fine-tuned using diverse plant images dataset to accurately detect canopy regions. The canopy size is then estimated for detection based on the pixel area of the detected regions. To support the practical deployment, the system consists of an interactive graphical interface. The experiment evaluation which is carried out shows reliable detection accuracy with low computational requirement, thus making the proposed approach suitable for real-time on-field applications. The important key observations which can be observed from this study include a comprehensive review of related work, end-to-end methodology, and covering model training with results and discussion along with the future directions.

Keywords: Tree Canopy Detection, Deep Learning, Convolutional Neural Network (CNN), Artificial Intelligence, MobileNet-SSD, Canopy Size Estimation

How to cite this article: Vibhute RR, Patil SS. Tree Canopy Detection Using Artificial Intelligence Based - Deep Learning Approach. Int J Drug Deliv Technol. 2026;16(11s): 223-231; DOI: 10.25258/ijddt.16.11s.21
.

1. INTRODUCTION

The interaction between vegetation and the surrounding environment has been studied for long time by researchers from agriculture, forestry, ecology, and urban planning fields. The plant canopy, which includes leaves, branches, and stems above the ground surface, plays an important role in regulating the local micro-climate. It controls the evapotranspiration process, also intercepts the rainfall, and influences soil moisture conditions. In addition to this, the shelter and the habitat is also provided by the canopy to many living species. The direct effect of the canopy n agricultural fields is seen to affect photosynthesis, weed growth, and nutrient absorption, which directly impacts the crop yield of cultivation. Similarly, on the other hand in the urban areas, these tree canopies help in reducing surface temperature and also improving air quality by providing shade and storing carbon. Hence, the measurement of canopy structure and its variation is very important for ecosystem monitoring and management decisions. However, accurate estimation of canopy parameters is still a challenging task. Hence, measuring the canopy structure and its changes is important for monitoring ecosystems and supporting management decisions. However, obtaining accurate canopy measurements is still difficult.

Currently, the traditional methods which are used for measurement of the canopy mainly depend on field-based techniques, which include point quadrant sampling, clinometer-based height management, and also photographic methods like hemispherical or digital cover photography. These approaches, though in-use, give detailed information at local level, but considering the real-world scenario, the data collection process is very much labor-intensive and time-consuming. The measurements which are obtained by such processes are also affected by the observer experience, which can lead to variation in the results. In addition to this, it is difficult to collect enough samples to represent large agricultural lands or forest regions, hence limiting the spatial coverage and monitoring frequency. Satellite imagery and airborne LiDAR systems have improved large-area canopy observation, but their spatial resolution is often not sufficient to capture small-scale canopy variations. These methods also require specialised skills and high operational cost, which makes them less suitable for small farmers and local governing bodies.
Recently the advanced methods such as the use of

---

*Author for Correspondence: rohitvibhute040@gmail.com machine learning and various computer vision techniques have created new possibilities for automatic vegetation analysis. Convolutional neural networks are widely used for image classification and object detection tasks, as features are automatically learned from raw image data. Lightweight models such as Mobile Net reduce computational complexity by using efficient convolution operations, thus allowing deployment on low-power devices. Transfer learning is commonly applied, where pre-trained models are fine-tuned for specific applications using limited dataset. In addition to deep learning, basic image processing techniques are also used to support the detection and estimation process. Hence, combining CNN-based detection with pixel-level analysis provides a practical solution for plant canopy detection and size estimation from ground-level images.

This paper proposes a deep learning framework to detect plant canopy regions and estimate their sizes using ground-level images. The objectives of this work are threefold: (i) to develop a dataset capturing diverse canopy structures under varying lighting and background conditions; (ii) to design and train a MobileNet-SSD model for accurate canopy detection with limited computational overhead; and (iii) to implement a user-friendly application that performs detection, size estimation, and visualisation. In addition, we present a detailed literature survey to contextualise our approach and discuss limitations and avenues for future research.

2. LITERATURE SURVEY

2.1 Traditional canopy measurement methods

Early studies on canopy assessment mainly depended on direct field-based measurements. Instruments such as the clinometer and spherical densitometer were commonly used to estimate canopy height and closure by measuring the angles towards canopy edges. The point quadrat sampling method was also used, where probes were inserted into the vegetation at fixed intervals and the number of foliage contacts were recorded. This approach helped in estimating the leaf area index (LAI). In addition to this, hemispherical photography and digital cover photography were applied by capturing upward-facing images using fisheye or wide-angle lenses. The canopy parameters were then obtained by separating sky and vegetation regions through image thresholding. Although these techniques provided detailed measurements at plot level, they required careful calibration and were highly dependent on observer judgement. Thus, the results often varied between users, and the methods were difficult to extend beyond small experimental areas.

With the advancement of remote sensing technologies, canopy monitoring at larger scales became possible. Multispectral and hyperspectral satellite images have been widely used to estimate canopy cover, LAI, and vegetation types. However, the spatial resolution of most satellite sensors is often larger than individual tree crowns, hence limiting the ability to capture fine canopy structures. LiDAR (Light Detection and Ranging) technology measures canopy height and structure by sending laser pulses and analysing their return signals. Ground-based terrestrial LiDAR systems can generate highly detailed three-dimensional point clouds, but the equipment cost is high and the measurements are affected by occlusion issues. Airborne LiDAR allows coverage of large areas, but the data collection process is expensive and the interpretation requires specialised expertise. Thus, despite their advantages, these methods are not always suitable for routine or low-cost canopy assessment. Zeng et al. [6] introduced a 3D LiDAR-based canopy imaging system mounted on an orchard utility vehicle. Their algorithm, implemented in MATLAB, effectively segmented trellis wires, poles, and trunks with accuracies above 82%. They also generated canopy density and depth maps, demonstrating the system's value in precision orchard operations such as pruning and variable-rate spraying.

Pawar and Patil [5] proposed a digital orchard tree canopy detection system aimed at reducing farmer workload. Their approach worked effectively across orchards containing mixed-age, variably spaced trees with diverse crown sizes and growth stages. The system demonstrated the practical benefits of automating canopy structure assessment compared to traditional manual inspection methods.

2.2 Remote sensing and deep learning for canopy detection

In recent years, machine learning techniques have been increasingly applied to remote sensing data for vegetation classification and canopy parameter estimation. Das and Dilkina [1] used a deep convolutional neural network to estimate urban tree canopy cover from high-resolution satellite images. The proposed model learned to differentiate between tree canopy and built-up surfaces, thus reducing the need for extensive manual annotation. Wang et al. [2] applied the YOLOv4 object detection model for identifying defects in aircraft glass canopies, which showed that modern CNN architectures are capable of handling complex object detection problems. Castelluccio et al. [6] used convolutional neural networks for land-use classification in remote sensing imagery, where the CNN-based approach performed better than conventional classifiers by learning spatial patterns automatically. Gao and Li [4] presented a research work in which they focused on detection and the measurement of the canopy height from ICESat-2 LiDAR data using improved filtering methods. In their work they highlighted how machine learning can be used effectively combined with signal processing to improve the vegetation analysis.

While satellite-based imagery gives an overall view of canopy coverage over large regions, ground-level images provide more detailed information about individual plants. Patil et al. [3] proposed a canopy detection and estimation system by combining deep learning with sensor fusion techniques in agricultural fields. In their method, image data was used along with additional sensors such as LiDAR or ultrasonic sensors to improve estimation accuracy. However, the use of multiple sensors increases system complexity, hardware cost, and maintenance effort. In addition to this, deployment becomes difficult for small-scale users. Hence, the present work focuses on canopy detection using only a single camera input, which makes the system simpler and more suitable for smallholder farmers and local communities.

Anagnostis et al. [7] introduced a U-Net–based semantic segmentation framework for orchard tree mapping using aerial imagery. Their dataset covered seven use-case scenarios across multiple walnut orchards, enabling the model to handle seasonal variations, weed presence, and differences in tree maturity. This work showed that deep learning–based semantic segmentation is highly effective for large-scale orchard canopy analysis.

Zhang et al. [10] addressed the difficulty of canopy edge segmentation in high-density orchards with overlapping branches and shadows. They proposed the MPAPR R-CNN framework—an enhanced Mask R-CNN integrated with PAFPN and PointRend—that achieved highly precise canopy boundary detection using UAV imagery. Their method successfully mitigated over- and under-sampling issues common in pixel-level segmentation.

Roslan et al. [8] attempted to detect tree crowns in dense tropical forests using a deep learning object detection model with optimized anchor ratios via K-Means clustering. Despite limited training data, their model achieved over 70% mAP, demonstrating strong potential for crown detection in complex natural ecosystems.

Partel et al. [9], Patil et al.[11] developed a smart agricultural sprayer using sensor fusion techniques, integrating LiDAR, machine vision, GPS, and AI-based object detection. Their system effectively classified trees, measured canopy height, and estimated canopy density for variable-rate spray applications. This approach highlighted the broader relevance of canopy detection in precision agriculture automation.

2.3 Convolutional neural networks and lightweight architectures

Deep convolutional neural networks have significantly changed the field of image recognition in recent years. Architectures such as AlexNet, VGGNet, and ResNet have shown very high accuracy when trained on large image datasets. Krizhevsky et al. [5] demonstrated the capability of CNNs for large-scale image classification by using multiple convolution layers followed by fully connected layers. However, these deep models contain a very large number of parameters and require high computational power. Thus, their deployment on low-resource or embedded devices becomes difficult.

To overcome this issue, lightweight CNN architectures were introduced to reduce model size and processing complexity. The MobileNet architecture achieves efficiency by using depthwise separable convolutions, where the standard convolution operation is divided into depthwise and pointwise convolutions. This reduces the total number of parameters and floating-point operations, while the accuracy is not reduced significantly. In addition to this, the MobileNet-SSD framework combines MobileNet with the Single Shot Multibox Detector, allowing object detection in a single forward pass. The use of depthwise separable convolutions reduces computational load, hence enabling faster inference while still learning spatial features effectively. Due to these advantages, MobileNet-SSD was selected in this work for plant canopy detection, as it is suitable for mobile and field-level applications. Recent studies such as Anagnostis et al. [7] and Roslan et al. [8] have demonstrated that even comparatively lightweight convolutional neural network (CNN) architectures can achieve strong performance in detecting canopy structures within highly complex orchard and forest environments. These works highlight that modern deep learning models do not always require heavy computational backbones; rather, well-designed lightweight networks can still capture essential spatial and textural cues necessary for vegetation segmentation and crown localization. Their findings also confirm that efficient models are more robust to variations in canopy shape, density, illumination, and background clutter—conditions frequently encountered in real-world agricultural fields. Given these insights, MobileNet-SSD emerges as an ideal choice for the present work. Its backbone, MobileNet, is specifically engineered using depthwise separable convolutions that drastically reduce parameter count and computational load without significantly sacrificing representational capability. When combined with the Single Shot Multibox Detector (SSD), the architecture delivers fast, single-stage object detection suitable for embedded and mobile platforms. This balance of speed, resource efficiency, and detection accuracy makes MobileNet-SSD particularly advantageous for field-deployable applications, where devices often operate with limited processing power and must provide real-time feedback.

By leveraging MobileNet-SSD, the proposed system ensures practical on-site canopy detection using only a standard camera and modest hardware resources. This enables broader accessibility for farmers, researchers, and local land-management personnel, ultimately supporting real-time monitoring and decision-making in resource-constrained agricultural settings.

2.4 Summary of previous research

The literature review clearly shows the gradual evolution of canopy assessment techniques over time. The traditional field-based methods provide accurate measurements at local level, but they are difficult to scale for large agricultural fields or forest areas. The remote sensing approaches have increased the spatial coverage, however the resolution is often not sufficient to capture fine canopy details. In addition to this, these methods require specialised knowledge and high operational cost, hence limiting their wider adoption. Machine learning techniques, especially deep learning models, have recently shown good potential for automating vegetation analysis by learning complex spatial patterns directly from image data. Lightweight convolutional neural network architectures make it possible to deploy such models on mobile or low-power devices, while transfer learning helps in adapting pre-trained models to new applications using limited training data. However, most of the existing research works mainly focus on aerial imagery, satellite data, specialised sensing devices, or multimodal sensor fusion techniques. Thus, the applicability of these systems becomes limited in low-resource and field-level environments. The present work addresses this research gap by proposing a deep learning-based canopy detection system that works on simple ground-level photographs. The system requires only a standard camera and moderate computational resources, hence making it more practical for small-scale farmers and local users.

3. Materials and Methodology

3.1 Dataset collection

The performance of any machine learning system mainly depends on the quality and variation present in the training dataset. To develop an effective canopy detection model, a diverse set of plant images was collected under natural field conditions. The images were captured from multiple sources, which included the agricultural farms, botanical gardens, and urban green areas, with the focus to include different plant species, different canopy structures, and also the surrounding as well as the background environments. The images were collected using consumer-grade digital cameras and also handheld mobile phones and smartphone devices at different times of the day to include the variations of the light. These variations in lighting conditions such as bright sunlight, partial shade, or cloudy weather were specifically included in the dataset. In addition to this, both broadleaf and conifer type plants were considered, with canopy size, density, and shape varying significantly. To make sure the model trains well on the collected dataset, the complexity of the dataset was increased by using clutter backgrounds which were intentionally put in the dataset. These backgrounds contained different unwanted objects such as buildings, vehicles, soil surfaces, and other vegetations, which made the task of detection of canopy more challenging. Hence, the model which was trained further to distinguish plant canopy regions from non-vegetation objects could generalize well. The final dataset which we collected consisted of nearly 1500 images with image resolutions which were ranging from 640x480 to 1920x180 pixels. Before we started the training, all images were resized to the fixed input resolution required by the MobileNet SSD architecture. This pre-processing step made sure that the data was uniform which could support efficient model training.

3.2 Data annotation

For the purpose of supervised learning, proper annotation and accurate labelling of training samples is required. In this paper, we used open-source annotation tool called Label Image, which was used to annotate the canopy regions in the collected images. The images were annotated manually with the objective to draw tight bounding boxes around the visible canopy portion of each and every plant image. If more than one plant appeared in the image, different bounding boxes were marked for each and every canopy in that image. Thus, multiple canopy instances were handled within a single image. The annotation files were saved in Pascal VOC XML format, which contained the image name, bounding box coordinates, and the class label defined as "canopy". In addition to this, cross-checking of the annotations was carried out to maintain consistency and reduce manual errors. Hence, the quality of labelled data was improved before training. The complete annotated dataset was then randomly split into training and testing sets, with almost 80% of the images in the dataset were used for training and the remaining 20% of the same reserved for testing. From the training set, around 10% of the data was further separated as a validation set for model tuning and early stopping during training.

3.3 Data preprocessing

Before the training process, image resizing was don't to resize them to 300×300 pixels, which matches the input size required by the MobileNet-SSD model used from the TensorFlow Object Detection API. The pixel intensity values were normalised by dividing each pixel value by 255. Thus, the input data was scaled uniformly before feeding into the network. In addition to this, data augmentation techniques such as random horizontal flipping, image scaling, and brightness variation were applied to increase the dataset diversity. The augmentation process was performed during training itself, so that new transformed samples were generated in each epoch. Hence, the model was exposed to different variations of the same image across training cycles. These transformations help in simulating real field conditions such as changes in lighting and viewpoint, and also reduce the chances of overfitting. As a result, the trained network shows better generalisation when tested on previously unseen images.

3.4 Model architecture

The selected model architecture in this work is the MobileNet-SSD, which combines the lightweight MobileNet network with the SSD detection framework. The MobileNet backbone uses depthwise separable convolutions to reduce the overall computational load. In this approach, each standard convolution operation is split into a depthwise convolution followed by a pointwise (1×1) convolution. Thus, the number of parameters and required computations are significantly reduced, while the ability to learn both spatial and channel-wise features is maintained. In addition to the backbone, the SSD detection head is attached with multiple convolution layers having decreasing spatial resolution. These layers allow the detection of objects at different scales within a single forward pass. Convolutional predictors are applied on these feature maps to estimate the bounding box locations and to classify the detected objects. Since the focus of this work is plant canopy detection, only one object class, labelled as "canopy", was considered along with the background class. Hence, the model structure remains simple and suitable for real-time field deployment.

3.5 Training details

The MobileNet-SSD model was initialised using weights pre-trained on the COCO dataset. The final classification and bounding box regression layers were replaced with newly initialised layers to match the single-class canopy detection task. Thus, the model was adapted specifically for the given problem. The training was carried out for a total of 50 epochs using SGD. Before training the Learning Rate (LR) of 0.001 was selected and was reduced gradually using exponential decay after every 10 epochs. The batch size was fixed to 24 images, mainly due to GPU memory limitations during training. The overall loss function consisted of two components: the localisation loss based on smooth L1 function for bounding box regression, and the confidence loss using softmax cross-entropy for classification. In addition to this, early stopping was applied based on the validation loss to avoid overfitting of the model. The training process was performed on NVIDIA RTX 2080 Ti GPU with the 32 GB RAM. Hence, the complete training process took nearly 4 hours to reach convergence.

3.6 Size estimation and metrics

After the canopy regions were detected by the model, the canopy size was estimated by calculating the area of the corresponding bounding boxes in pixel units. Since the captured images did not contain absolute scale information, the size estimation was expressed in relative terms. Thus, the canopy size was represented as a percentage of the total image area. In addition to this, if the distance between the camera and the plant is available, the pixel-based measurements can be converted into real-world dimensions using basic perspective projection relations. However, in the present experiments, the canopy size was reported as a proportion of the image frame to allow fair comparison across different images. The detection performance of the proposed system was evaluated using standard performance evaluation parameters for object detection metrics. Precision was calculated as the ratio of correctly detected canopy regions to the total detected regions, while recall represented the ratio of correctly detected canopies to the total ground-truth canopies present in the image. In addition to this, mean average precision (mAP) was used to evaluate overall detection accuracy. An intersection over union (IoU) threshold of 0.5 was considered to determine whether a detection was correct. Hence, these metrics provided a clear understanding of the detection and estimation performance of the system.

3.7 Implementation specifics

After the canopy regions were detected by the model, the canopy size was estimated by computing the area of the detected bounding boxes in pixel values. Since the captured images do not contain any absolute scale reference, the size estimation was carried out in relative form. Thus, the canopy size was expressed as a percentage of the total image area. In addition to this, if the distance between the camera and the plant is known, the pixel-based measurements can be converted into real-world dimensions using simple perspective projection relations. However, in the present experimental setup, the canopy size was reported only as a proportion of the image frame so that comparison across different images becomes easier. The performance of the proposed canopy detection system was evaluated using commonly used object detection metrics. The precision metric was calculated as the ratio of correctly detected canopy regions to the total number of detected regions, while the recall metric represented the ratio of correctly detected canopies to the total ground-truth canopies present in the image. In addition to this, the mean average precision (mAP) was used to assess the overall detection performance of the model. An intersection over union (IoU) threshold value of 0.5 was used to decide whether a detected bounding box was considered as a correct detection. Hence, these evaluation measures helped in understanding both the detection accuracy and estimation reliability of the system.

4. System Design

4.1 Module overview

The proposed system comprises five different main modules: (i) Image Acquisition, where images are captured using a camera or loaded from disk (ii) the Image Preprocessing, which resizes and normalizes images; (iii) the Canopy Detection, where the MobileNet-SSD model predicts bounding boxes around canopy regions; (iv) Size Estimation, which calculates the area of each detected canopy; and (v) Result Visualization, providing graphical overlays. The Figure 1 illustrates the data flow between these modules. We focused on the modular design, which made sure that maintenance is easy and allows the individual components to be replaced without affecting the overall pipeline of canopy detection.

4.2 The GUI App

The application which was developed in this research provided with a simple and user-friendly way to interact with the proposed system. The interface team developed in such a way that it included the basic controls for selecting the input image, running the canopy detection process by feeding the same to the trained deep learning model, and then viewing the output results on the same screen. After the detection, the identified canopy regions are shown by drawing the bounding box around them in the image where the box colour indicates the confidence level of the detection. Thus, the user who was using this GUI application could easily understand the detection output results visually.

In addition to the results which were shown on the image, the side panel was also provided in the GUI application which displayed the numerical information such as the total number of canopies detected in their image, their relative size values, and also the confidence level of each and every canopy which is detected. The GUI also provides options to save the annotated image containing the detection results and to export a text-based summary of the estimated canopy sizes. Hence, the system output can be easily stored fr further analysis or reporting.

The overall simplicity of the interface makes the tool usable even for users who do not have prior knowledge of machine learning or image processing.

4.3 Implementation pipeline

The detection pipeline starts with the acquisition of an input image, which can be obtained either from a live camera feed or from local system storage. The acquired image is first sent to the preprocessing module, where it is resized to 300×300 pixels and normalised before further processing. During the inference stage, data augmentation is not applied. Thus, the image is directly prepared for model input without additional transformations. The pre-processed image is then passed to the MobileNet-SSD model, which generates a set of bounding box predictions along with their corresponding confidence scores. In order to avoid multiple detections of the same canopy region, non-maximum suppression (NMS) is applied to remove overlapping bounding boxes having lower confidence values. Hence, only the most relevant detection for each canopy is retained. The final bounding boxes are then mapped back and drawn on the original image, and the area of each detected box is calculated for size estimation.

In addition to this, the annotated output image along with the estimated canopy size information is displayed to the user through the GUI. The average processing time for a single image was around 70 ms when tested on a workstation system.

4.4 Data flow and visual illustration

Figure 1 shows the flowchart of the system. The flowchart highlights each module and the direction of data flow, starting from image acquisition, through preprocessing and detection, to size estimation and visualization. The clear separation of modules allows for potential integration of additional features such as multi-class detection, temporal tracking, or fusion with depth sensors in future work.

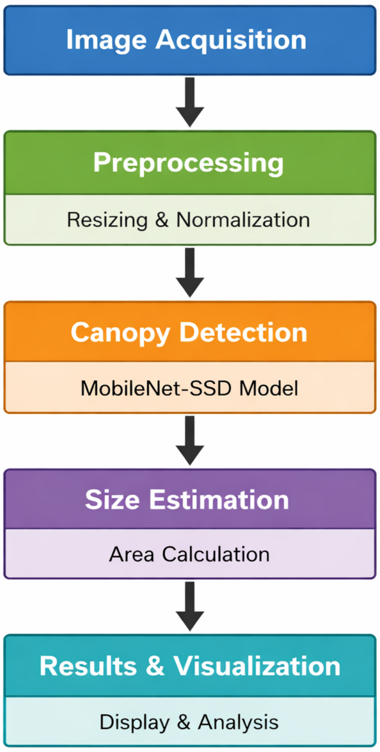

Figure 1: Flowchart of plant canopy detection and size estimation system

4.5 Conceptual illustration of canopy detection

To aid understanding, Figure 2 presents a conceptual illustration of canopy detection. The image depicts a plant with its canopy region highlighted to show what the detection algorithm aims to isolate. The illustration emphasises that the goal is to identify the green foliage mass while excluding the trunk, branches, and background elements. From the illustration seen above, we can observe that the goal is to identify the green area in the image while excluding the trunk, the branches, and the background elements. This conceptual diagram shown in the illustration are very much useful in the understanding of the exact flow of the system from the start to end.

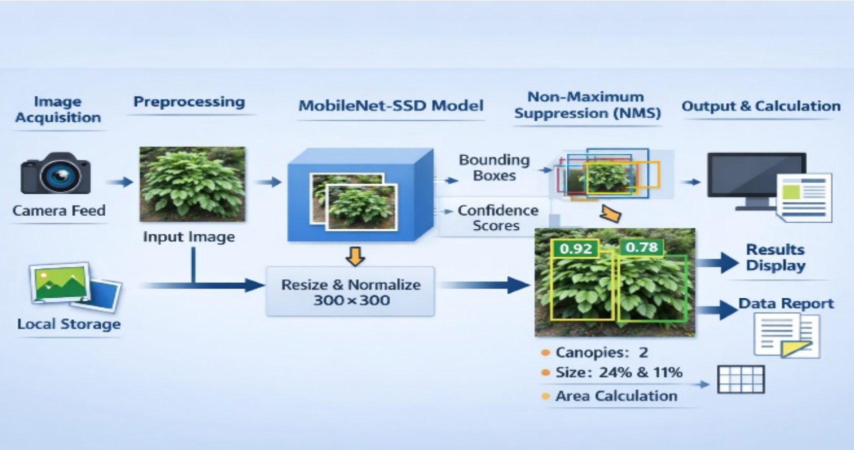

Figure 2: Conceptual illustration of plant canopy detection

5. Results and Discussion

6. 5.1 Detection accuracy and performance

Post-training it was observed that the model which was trained on MobileNet SSD achieved a mean average precision of 0.83 at an IOU threshold of 0.5 on the test dataset which was kept aside during training. From the precision and recall observed, we obtained the values of 0.85 and 0.80 respectively which indicated that there is a good balance between correctly identifying the canopy regions thereby minimizing the false detections. The relatively high precision value which is obtained shows that the model rarely misclassifies non-canopy objects while the recall can be used as an evaluation parameter that justifies that most of the true canopies were detected. Thus, from the results obtained, we can observe that the trained model was able to perform very much well across various lighting conditions and background, although few errors occurred when the canopy was blended with similar coloured backgrounds, i.e., when the intensity of green colour in the background was more than that of the expected. Inference speed is critical for practical deployment. On a desktop GPU, the model processed approximately 14 images per second, sufficient for near real-time applications. On a Raspberry Pi 4 equipped with a Coral Edge TPU, the model processed around 4 images per second, demonstrating feasibility for field-deployed devices.

5.2 Comparison with baseline models

To evaluate the benefits of using MobileNet-SSD, we compared its performance against two baseline detectors: a traditional image processing method using colour thresholding and morphological operations, and a heavier deep learning model (Faster R-CNN with a ResNet-50 backbone). Table 1 summarises the results. The colour-based method achieved poor precision (0.54) and recall (0.60) due to its sensitivity to lighting and background variations. Faster R-CNN achieved a slightly higher mAP (0.85) than MobileNet-SSD but required four times more computation and was unsuitable for low-power devices. Therefore, MobileNet-SSD offered the best	mAP@0.5	Precision	Recall	Inference time (ms/image)

trade-off between accuracy and efficiency. Model				
Colour thresholding	0.37	0.54	0.60	15
Faster R-CNN (ResNet-50)	0.85	0.88	0.82	280
MobileNet-SSD	**0.83**	**0.85**	**0.80**	**70**

5.3 Effect of data augmentation and training strategies

Data augmentation contributed significantly to model generalisation. Without augmentation, the model tended to overfit to specific lighting conditions and backgrounds, resulting in lower recall. Horizontal flipping helped the model generalise to canopy orientation, while brightness adjustments simulated varying illumination. Fine-tuning the learning rate schedule and using early stopping prevented overfitting and reduced training time. Transfer learning proved essential; training MobileNet-SSD from scratch on the limited dataset resulted in a mAP of 0.60, whereas fine-tuning the pre-trained network achieved a mAP of 0.83.

5.4 Discussion of results

The results indicate that a lightweight deep learning model can effectively detect plant canopies from ground-level images. The combination of transfer learning, data augmentation, and a carefully designed architecture allowed the model to generalise across diverse species and conditions. However, there are few limitations which are observed. From the results obtained, we can observe that the detection accuracy decreased little bit for very small or highly occluded canopies. The reason for this being, the model's receptive fields may not be able to capture the very fine details at such lower scales. In addition to this, the non-vegetative green objects, such as painted surfaces or artificial turf, was seen to occasionally trigger false positives, thereby decreasing the model's detection capabilities. By using the spectral information from multi-channel sensors or combining the RGB data from the cameras with near-infrared imagery, can be used in future to improve the discrimination which is existing due to similar green-coloured objects. Additionally, the size estimates can be expressed as relative pixel areas which are suitable for comparative analysis, but not for absolute measurements unless strict camera calibration is performed.

5.5 Limitations and future improvements

Several options and methods can be used for enhancing the performance of the system in the future. By incorporating depth estimation, by measuring the data from LiDAR sensors or stereo cameras can help us convert the pixel-level measurements into real-world dimensions. In addition to this, multi-class detection can be used, which could distinguish between the species or canopy health conditions, thereby including the biodiversity assessment and disease monitoring. Also, by using temporal tracking across image sequences could help us to allow the observation or tracking of the canopy growth over time. Finally, by including a greater number of images in the dataset and collection of the diverse dataset along with different environmental conditions and occluded backgrounds can help us further improve the detection capability of the model with very little false positives.

7. Conclusion

In this paper, we have presented a deep learning framework for canopy detection and size estimation from the image data. This proposed system uses a MobileNet SSD architecture which was fine-tuned on a diverse dataset of custom collected canopy images. The model performance observed post-training shows us that the model achieves a proper balance between the accuracy and the computational efficiency. It is suitable for deployment on mobile and edge devices which have resource-constrained computing capability. The GUI interfaced which was developed was user-friendly and helps in operations such as image loading, detection, as well as result visualization with ease. The results from the experimental results obtained in this research demonstrate that the system performs reliably well across varying lighting conditions and background. The comparison with the traditional methods and heavier model shows us the advantages of the proposed approach in terms of both precision as well as speed.

From the future point of view, the focus should be on extending the system to estimate the absolute canopy sizes using depth or spatial information, which focuses on multi-class detection to classify different plant species or health statuses. The solution should be a hardware-deployable, low-cost solution which can help in real-world agricultural and forestry applications. The use of additional sensors and temporal analysis can further enhance the capability of the system to monitor canopy dynamically and contribute to sustainable land management.

References

[1] N. Das and B. Dilkina, “A Deep Learning Approach to Assessing Urban Tree Canopy Using Satellite Imagery in the City of Atlanta,” Georgia Institute of Technology, 2017.

[2] J. Wang, S. Wei, K. Wang, and J. Li, “Surface Defect Detection of Aircraft Glass Canopy Based on Improved YOLOv4,” *Proceedings of the International Symposium on Communications and Information Technologies*, 2022.

[3] Patil, S. S., Y. M. Patil, and S. B. Patil. "Review

on automatic variable-rate spraying systems based on orchard canopy characterization." *Informatics and Automation* 22, no. 1 (2023): 57-86.

[4] S. Gao and Y. Li, "Retrieving Forest Canopy Height from ICESat-2 Data by an Improved DRAGANN Filtering Method," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Proceedings of the Neural Information Processing Systems Conference*, 2012, pp. 1097–1105.

[6] M. Castelluccio, G. Conte, A. F. Rossi, and D. Rohm, "Land Use Classification in Remote Sensing Images by Convolutional Neural Networks," *arXiv preprint arXiv:1508.00092*, 2015.

[7] Athanasios Anagnostis, Aristotelis C. Tagarakis, Dimitrios Kateris, Vasileios Moysiadis, Claus Grøn Sørensen, Simon Pearson and Dionysis Bochtis." Orchard Mapping with Deep Learning Semantic Segmentation", 2021 MDPI.

[8] Zhafri Roslan; Zalizah Awang; Mohd Nizam Husen; Roslan Ismail; Robiah Hamzah, "Deep Learning for Tree Crown Detection in Tropical.", 2020 14th International Conference on Ubiquitous Information Management and Communication (IMCOM).

[9] Victor Partel, Lucas Costa, Yiannis Ampatzidis,"Smart tree crop sprayer utilizing sensor fusion and artificial intelligence",2021, ELSEVIER.

[10] Weirong Zhang, Xuegeng Chen, Jiangtao Qi, Sisi Yang, "Automatic instance segmentation of the orchard canopy in unmanned aerial vehicle imagery using deep learning". December 2022, Sec. Technical Advances in Plant Science.

[11] Patil, S.S., Patil, Y.M., Patil, S.B. and Powar, R.V., 2024. A Comparative Analysis of Plant Canopy Detection Performance in a Variable-Rate Spraying System Using Deep Learning Models. Journal of The Institution of Engineers (India): Series A, *105*(3), pp.487-497.