

## Statistical Methods for Enhancing Performance in Precision Agriculture

**Prof. K. M. Pimple<sup>1</sup>, Prof. R. R. Solanke<sup>2</sup>, Prof. A. S. Chaudhari<sup>3</sup>, Prof. R. M. Pusdekar<sup>4</sup>, Prof. M. S. Shribhate<sup>5</sup>, Prof. R. D. Thakare<sup>6</sup>, Prof. P. R. Wankhade<sup>7</sup>**

<sup>1</sup>Assistant Professor, Dr. Rajendra Gode Institute of Technology & Research, Amravati

Email ID: [k.m.pimple1999@gmail.com](mailto:k.m.pimple1999@gmail.com)

<sup>2</sup>Assistant Professor, Dr. Rajendra Gode Institute of Technology & Research, Amravati

Email ID: [rrsolanke1972@gmail.com](mailto:rrsolanke1972@gmail.com)

<sup>3</sup>Assistant Professor, Dr. Rajendra Gode Institute of Technology & Research, Amravati

Email ID: [amrutachaudhari58@gmail.com](mailto:amrutachaudhari58@gmail.com)

<sup>4</sup>Assistant Professor, Dr. Rajendra Gode Institute of Technology & Research, Amravati

Email ID: [roshana28@gmail.com](mailto:roshana28@gmail.com)

<sup>5</sup>Assistant Professor, Prof. Ram Meghe College of Engineering and Management, Badnera, Amravati

Email ID: [monikashribhate@gmail.com](mailto:monikashribhate@gmail.com)

<sup>6</sup>Assistant Professor, Sipna College of Engineering & Technology, Amravati

Email ID: [iamrupali14@gmail.com](mailto:iamrupali14@gmail.com)

<sup>7</sup>Assistant Professor, Dr. Rajendra Gode Institute of Technology & Research, Amravati

Email ID: [wankhadepr713@gmail.com](mailto:wankhadepr713@gmail.com)

*Cite this paper as:* Prof. K. M. Pimple, Prof. R. R. Solanke, Prof. A. S. Chaudhari, Prof. R. M. Pusdekar, Prof. M. S. Shribhate, Prof. R. D. Thakare, Prof. P. R. Wankhade, (2025) Statistical Methods for Enhancing Performance in Precision Agriculture, *Journal of Neonatal Surgery*, 14 (28s), 317-337

### ABSTRACT

Agriculture is key to securing food for the 9.7 billion people projected to inhabit the world by 2050. However, managing plant health remains a significant challenge for crop productivity. Traditional approaches for identifying plant species, as well as for disease classification and weed detection, are laborious, inefficient, and time-consuming. We use deep learning architecture and transfer learning methods to provide automated, rapid, proven, and accurate plant solutions. Plant Leaf Classification uses deep learning models to recognize species of plants from their leaves, aiding in medical, biodiversity, and ecology research. Transfer learning plays a vital role in leaf disease classification, detecting diseases as quickly as possible, and taking timely action. Object Detection Models help in Leaf Disease Detection and Weed Detection in real-time, when used in a way that allows for controlled application, rather than spraying pesticides throughout the entire field. ResNet50, DenseNet201, YOLOv7, and YOLOv8 are evaluated in the study based on performance metrics such as accuracy and loss curves, class distribution analysis, and confidence curves. Model efficiency is briefly addressed using key metrics: precision, recall, f1 score, and MAP (Mean Average Precision). These findings inform the best model and hyperparameter selection for plant classification tasks. More specifically, this study enhances the accuracy and sustainability of plant management, advancing precision agriculture toward greater efficiency and data-driven solutions

**Keywords:** Precision Agriculture , Deep Learning , Plant Disease Detection , Transfer Learning , Weed Detection

### 1. INTRODUCTION

Food security and sustainable agriculture are significant challenges as the global population is rising rapidly. These explain the increasing demand to feed a growing world population, projected to reach 9.7 billion by 2050, with agricultural productivity needing to increase by at least 25% to meet the demand for food. Despite this, food production faces challenges due to limited resources, including land, water, labor, climate, crop diseases, and pest attacks. Technological innovations are necessary to develop more efficient farming methods and promote sustainable agriculture in light of these challenges.

Agriculture remains the backbone of the Indian economy, accounting for around 18% of the GDP and providing employment

to a large segment of the population. On the other hand, most agricultural methods are still based on human labor and ancient technologies, which makes them less efficient and effective than they could be. By harnessing the potential of technologies like computer vision, deep learning, and transfer learning, the future of enhancing precision agriculture and improving crop management lies embedded. Deep learning techniques have proven highly effective for crop condition tracking, disease monitoring, and weed management due to the real-time functionality and accuracy provided by these data-driven approaches.

### 1.1 Role of Precision Agriculture

Precision agriculture uses automated data collection and analytics to improve farming decision-making. It includes plant leaf classification, disease classification, and weed detection. By enhancing crop yields, minimising crop loss, and reducing the point of application of pernicious chemicals, each area contributes critical value-added components to the agricultural lifecycle.

### 1.2 Plant Leaf Classification

Plant species identification is crucial for agricultural research, biodiversity conservation, and the cultivation of medicinal plants. Conventional plant identification methods involve visual analysis, which can lead to misidentification and is time-consuming. Deep learning and transfer learning techniques extract shape, color, and texture features from image data to improve plant species recognition, enabling farmers and researchers to make informed crop management and conservation decisions.

### 1.3 Leaf Disease Classification

Plant diseases lead to enormous crop losses worldwide, resulting in decreased yields and food quality. Traditional pathogen identification is primarily based on manual inspection, which requires experience and labor-intensive work. Deep learning-based models for disease classification make it easier to identify diseases in plants, enabling early detection and treatment. Transfer learning is then applied to enhance precision and offer robust solutions for multiple plant species.

### 1.4 Leaf Disease Detection

Detecting disease early is vital to controlling outbreaks and minimizing economic losses. API-based systems use object detection models in computer vision to localize plant disease symptoms on leaves, supporting real-time monitoring. These capabilities can help farmers deliver targeted treatments or treatments only where needed, instead of using broad-spectrum pesticides, which reduces costs and environmental damage.

### 1.5 Weed Detection

Weeds compete with crops for nutrients, light, and water, which decreases agricultural productivity. Herbicides are often used indiscriminately, leading to soil degradation, chemical residues, and pollution of natural resources. Using object detection models for targeted weed control helps ensure that pesticides are applied only where necessary, thereby reducing chemical use and promoting sustainable growing systems.

### 1.6 Read to discover the Top 10 Career Opportunities in Deep Learning & Computer Vision in 2023.

Computer vision and deep learning advancements have transformed plant monitoring and disease management. CNNs are an excellent tool for identifying the complex features of plant leaves, disease symptoms, and weed symptoms. Furthermore, transfer learning enhances the performance of models by leveraging pre-trained networks and minimizing the need for extensive labeled datasets. Such solutions are economical and scalable, providing very high accuracy in agri automation

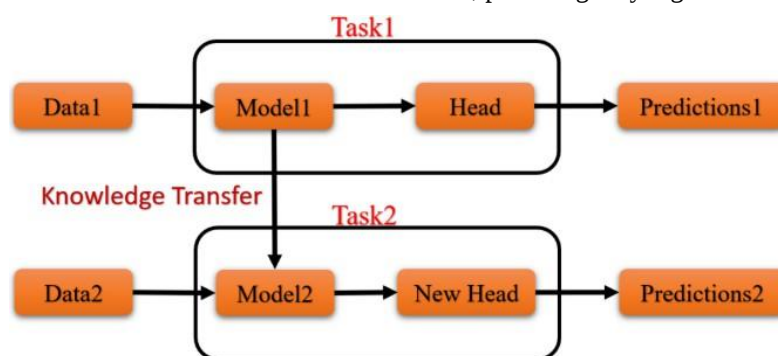


Figure 1: Transfer Learning

Transfer learning, known for its ability to efficiently generalize knowledge across different domains, especially in computer vision, has proven to be a powerful machine learning approach that has received much attention. In this context, transfer

learning uses outputs from one task to aid in a similar task, minimizing the data and computing resources required for training [RDS+19]. Transfer learning has been wildly successful in computer vision, including functions such as image classification and object detection, where pre-trained models can more accurately identify objects and their locations in images.

Historically, training deep learning models from scratch on new datasets has demanded a lot of computing power and large labeled datasets. Transfer learning mitigates these issues by using pre-trained models trained on large datasets, such as ImageNet. These models typically use lower-level features (edges, shapes, and textures) and adapt them for use across domains. Transfer learning allows for fine-tuning of the last layers of the model using task-relevant datasets, adjusting to new tasks efficiently [RDS+19].

Transfer learning has many benefits. First, it alleviates the need for large, labeled datasets, enabling deep learning for domains with sparse annotated data. Second, it reduces computational costs: not only does the entire network need to be trained from scratch, but only specific parts require fine-tuning. Finally, transfer learning accelerates convergence and enhances the generalization of models, achieving better outcomes with fewer iterations [RDS+19].

Transfer learning enables essential applications in agriculture, such as the automatic identification of plant species, disease detection, and weed identification, among others. It allows robust agricultural solutions powered by deep learning, thanks to its innovative, pre-trained models. Transfer learning is helpful in this aspect as it involves imitating knowledge from one component to another. By fine-tuning the models on domain-specific agricultural datasets, real-world agricultural challenges can be addressed, improving precision agriculture and sustainable agriculture, ultimately leading to LW21.

## 2. RELATED WORK

Plants are the primary source of food, shelter, clothing, shoes, medicines, and more. Hence, we must protect plants and preserve these resources for future generations. This paper aims to identify plant diseases and weed infestations relevant to general agriculture. To cope with these problems, this study proposes automated models for plant species classification, plant disease detection, and weed detection, utilizing new deep learning and transfer learning models. This section summarizes the state-of-the-art research in the domain, while the literature survey is divided into four main parts.

Plant Leaf Classification – Quickly reviewing available methods to identify plant species based on leaf images. Plant disease classification – An overview of deep learning models used for classifying plant diseases.

Plant Disease Detection – Implementing automated methods to identify disease from images of plants.

Weed Detection – Investigating the techniques of automatically identifying whether a crop is a weed to improve precision agriculture.

Plant species identification is crucial in precision farming. It must be efficient and accurate. Many efforts have been made to design robust computer vision-based methods for automatic plant species identification.

### 2.1 Traditional Methods and Dataset-Based Methods

Over the years, machine learning techniques for plant species recognition have been widely investigated, underpinned by various benchmark leaf image datasets, e.g.

ICL Dataset [SMdS13] Swedish Leaf Dataset [Sod01] Flavia Dataset [WBX+07] Foliage Dataset [KNSS13] LeafSnap Dataset [KBB+12]

Flower17 [NZ07] und Flower102 [NZ08]

These datasets have been used to design and evaluate numerous classification algorithms, such as Convolutional Neural Networks (CNNs) and transfer learning techniques.

### 2.2 Contribution of Deep Learning-Based Approaches

Kumar et al. [KBB+12] use this in LeafSnap, a mobile application designed to identify plants based on images of their leaves. For 184 plant species, the model reached 96.8% top-5 accuracy and 73% top-1 accuracy.

Joly et al. PlantNet, an online tool for identifying plant species from multi-part images (such as leaves, fruits, bark, flowers, etc.), was invented by [JGB+14]. When tested on 2,200 French plant species, it achieved a top-5 accuracy of 69% for individual images.

Habiba et al. Deep CNN model for Bangladesh plant leaf classification [HIA19]. Among the deep learning models, VGG16 achieved the highest accuracy (96%)

Liu et al. For plant leaf classification, [LYCS18] used a ten-layer CNN with various data augmentation techniques. On the Flavia dataset, the model's accuracy on that data was 87%. 92 %

Bhat et al. According to CNN, BDS+21 developed a leaf recognition system for plant species in the Western Ghats, using

leaf images collected from more than 50,000 plants of 48 species. The models, such as InceptionV3, MobileNet, VGG-16, and DenseNet, were analyzed, and VGG-16 achieved an accuracy of 93.79%.

Kanda et al. [KXS21] proposed a closed-loop method combining Conditional Generative Adversarial Networks (CGANs) with CNN and Logistic Regression to generate synthetic data and extract features for highly accurate classification of plant species

### 2.3 Methods of Classifying Leaf Diseases

Accurate and efficient classification of plant leaf diseases is an essential aspect of precision agriculture for proper detection and crop management. Deep learning approaches and computer vision techniques have significantly improved the outcomes of disease classifiers. Many methods have been proposed for analyzing EEG signals in sleep stages, ranging from deep learning architectures and traditional machine learning models to hybrid models with feature extraction. The following section showcases a summary of significant works in plant disease classification, including their techniques, datasets, and performance. Neural Networks for Computer Vision

Rangarajan et al. Examples: RPR18, where six pathologies in tomato leaves were detected using the PlantVillage dataset. AlexNet and VGG16 were implemented using transfer learning and hyperparameter tuning. The results of their experiment included AlexNet, which achieved an accuracy of 97.496%, and VGG16, which achieved an accuracy of 97.292%. This included studying the effect of dataset size, batch size, and learning rate on performance and processing time.

Shanker et al. [SSB22] proposed a 28-layer Convolutional Neural Network (CNN) for plant disease detection to automate the identification of plant diseases. The model was trained on A dataset of 87.7K RGB images of 38 plant diseases and healthy classes. They reported a classification accuracy of up to 98.13%, which proved the power of deep learning in identifying large-scale plant diseases.

Bensaadi et al. In [BL23], an optimal CNN architecture for early-stage plant disease classification was proposed. The model uses a dataset with 57,000 tomato leaf images (9 disease classes) being trained to yield 97.04%, achieving an error rate of under 0.2%, which signifies the model's ability to recognise multiple diseases.

Kamilaris et al. Various deep-learning architectures were investigated for agriculture-related problems, such as plant disease classification [KPB18]. An impressive 99.53% accuracy was reported for a CNN trained on 87,848 images of 25 plant species [Fer18]. A nine-layer deep CNN for leaf disease classification reported an accuracy of 96.46% [AHK+22]. In addition, TL-ResNet50 and other transfer learning approaches also achieved high performance, with

98.20% accuracy for disease classification [BKA+22]. Classification accuracy improved by 99.69% when attention-based CNNs were applied to a tomato leaf disease classification task [ZDC+19a]. Hybrid & Statistical Learning Approaches

Anitharani et al. Using supervised learning techniques, AVH+22 evaluated and reported the performance of Support Vector Machines (SVM), Gaussian Naive Bayes, and Random Forest for plant disease classification. The authors used Gabor filters and feature extraction with Artificial Neural Networks (ANNs), achieving a recognition performance of 91%. The study highlighted the importance of early and accurate disease detection for precision farming purposes.

Arasakumaran et al. In [AJSK22], the authors proposed a hybrid plant disease identifier tool that utilized fuzzy c-means clustering (FCM) for infected area segmentation, a grey-level co-occurrence matrix for feature extraction, and progressive neural architecture search (PNAS) for plant classification. On the Mendeley Dataset, it achieved a 95% accuracy. Based on this work, the integration of traditional feature engineering with deep learning for disease classification.

Additional studies have investigated SVM-based classification for specific diseases. For example, the authors in [MAHH15] applied a quadratic kernel SVM model to identify tomato yellow leaf curl disease, achieving 92% accuracy. A hybrid ANN and SVM model in citrus plants has achieved 92.8% and 92.2% accuracy in recognizing Huanglongbing disease [WdONB+17]. Moreover, apple leaf spot disease forecasting was performed using K-Nearest Neighbors (KNN), achieving an accuracy of 88% [HZZ+22]. Emergence of Trends & Directions for the Future

The subsequent generation of plant disease classification systems is built on recent advances in self-supervised learning (SSL), Vision Transformers (ViTs), and Explainable AI (XAI), as displayed here:

Vision Transformers (ViTs): Vision Transformers have gained popularity as potential replacements for Convolutional Neural Networks (CNNs). Their self-attention mechanisms enable more efficient processing of leaf images, resulting in improved performance on large-scale datasets. A 98–99% accuracy in classifying plant diseases has been reported using ViT architectures.

Self-Supervised Learning (SSL): Labeled datasets are expensive and time-consuming. SSL techniques can learn image representations from these unlabeled datasets, which is attractive for real-world agricultural systems where labeled datasets are limited.

Explainable AI (XAI): Understanding the reasoning behind model decisions is crucial for real-world applications. These XAI techniques (Grad-CAM, SHAP) can also visualize the model's interpretation of why it classifies a leaf as diseased and improve the trustworthiness of AI-driven plant diagnostics.

Remote Display: Transfer learning could push the limits of very lightweight deep learning models, which are catastrophically optimized for edge devices (such as smartphones, drones, and IoT sensors). It could also detect diseases in the field in real-time.

Cross-Domain Transfer Learning: Future research investigates models that train one model on multiple datasets, allowing it to transfer knowledge and recognize diseases for different plant species simultaneously.

#### **2.4 Approaches for Detection of Leaf Diseases**

Background: Leaf diseases pose a critical threat to agriculture, devastating crop yields and quality. Early Detection of plant diseases allows farmers to take timely actions to mitigate infections and prevent substantial yield losses. The researchers have formulated automated leaf disease detection techniques, including computer vision, deep learning, and object detection models. In this element, we present a summary of contemporary works that examine diverse techniques, datasets, and models used to identify and detect plant diseases.

##### **Data Collection for Leaf Disease Detection**

Davinder Singh et al. The PlantDoc dataset introduced was collected and labeled in the work of [SJJ+20] and aims to provide a large-scale dataset that facilitates the development of computer vision-based solutions for plant disease detection. This dataset of 2,598 images of 13 plant species and up to 17 disease classes is an essential resource, given the limited availability of non-laboratory, real-world datasets. They used three deep-learning models for plant disease classification and showed they achieved up to 31% better accuracy in training with PlantDoc. The authors highlight that such a dataset can significantly reduce the cost of developing computer vision-based plant disease detection methods and improve early Detection and intervention plans in farming practices.

Methods Based on Deep Learning and Object Detection

Plant Diseases Detection using YOLOv5 Optimization

Haiqing Wang et al. The YOLOv5-based version of [WSW+22] was optimized to detect and classify plant diseases. Their strategy included some improvements, such as:

##### **2.5 More precise IASM mechanism,**

GhostNet to improve computational efficiency,

Weighted Bi-directional Feature Pyramid Network(BiFPN) for improved feature fusion, Use fast normalization fusion to improve model performance.

With this optimized model, they achieved an F1 Score of 92.65% and an accuracy of 92.57% on the classification task, based on numerous comparison and ablation studies on their custom dataset. The authors noted that transfer learning could further expand the model's applicability to various agricultural settings.

Early Detection of Bacterial Spot Disease Based on YOLOv5

Mathew et al. MM22 proposed a deep-learning model for bacterial spot disease detection in bell pepper plants based on YOLOv5. Understanding the importance of early Detection in reducing crop destruction, the research used YOLOv5s for accurate object detection. This model successfully identified infected and healthy plant leaves with a mean average precision (mAP) score of 90.7% [38]. Their results highlight the need for real-time, automated detection of plant diseases, which is critical for smart agriculture.

##### **2.6 Rice Leaf Disease Detection Using YOLOv5**

Ershadul Haque et al. [EHRJ+22] addressed the challenge of detecting disease in rice leaves, a disease that can severely decrease crop yield for rice, one of the world's critical staple crops. Douche et al. [9] already proposed a method for accelerating an object detection process, which relies on YOLOv5-based deep learning techniques. They achieved better results using YOLOv5-based deep learning techniques than those based on traditional approaches. The model has been trained using an annotated dataset of four rice leaf diseases, producing:

Recall: 0.94,

Precision: 0.83,

Map: 0.62.

This performance demonstrates the efficacy of YOLOv5 in detecting diseases in agriculture, particularly in cases that require

high precision and recall rates during disease classification.

## **2.7 Hybrid Techniques Combine DL with Segmentation Segmentation And Classification Of Leaves Using Mask R-CNN**

Masmoudi et al. proposed a computer vision and deep learning framework in [ML21] that combines Mask R-CNN for leaf segmentation and CNN for disease classification. The model was trained on the PlantDoc dataset for leaf segmentation and the PlantVillage dataset for disease classification. Their experiments yielded encouraging results: 76% for leaf segmentation,

Disease classification accuracy: 83%

This research contributes to the successful integration of segmentation and classification models for effectively assessing plant disease monitoring systems in smart agriculture.

## **2.8 Fine-Grain Object Detection Based on YOLOv4**

Roy et al. [RBB22] proposed an improved framework based on YOLOv4 for real-time, accurate delineation of diseases in crops. To tackle significant challenges like dense symptom distribution and non-uniform morphology, the study incorporated:

DenseNet instead (the class loaded through the database). Residual blocks refresher for deep learning,

Spatial Pyramid Pooling (SPP) for multi-scale feature extraction, PANet Enhanced for Refined Detection.

Their model achieved: Precision: 90.33%,

F1-score: 93.64%,

mAP: 96.29%.

This study highlights the utility of fine-grained, high-precision object detection models in enabling large-scale automated plant disease detection.

## **2.9 Deep Learning and Attention Mechanisms**

Grape Leaf Disease Detection Based on Dual Attention

Dwivedi et al. These include the Grape Leaf Disease Detection Network (GLDDN) proposed by [DDCT21], which uses a dual attention mechanism to improve the disease identification ability of grape plants. The model focused primarily on ECA, black rot, and Isariopsis detection. Unlike traditional approaches, it utilized attention-based feature extraction to emphasize disease-affected areas. Demonstrated by experimental results:

Grape disease detection with an accuracy of 99.93% Better localising the infected areas,

A wide range of lighting and environmental conditions.

This system provides scalable and accurate disease identification and assists grape growers in planning preventive measures.

## **2.10 Novel Trends of Leaf Disease Detection**

Recent studies are moving towards advanced deep learning models and SSL, ViTs, and XAI to improve the detection capability of plant diseases [4]. Key trends include:

We go for Vision Transformers (ViTs): Unlike CNNs, ViTs can utilize self-attention techniques to extract detailed features throughout the leaf image, enhancing classification accuracy. Recent studies have reported that ViTs can achieve accuracies of 98–99% in plant disease identification.

Self-Supervised Learning (SSL): SSL techniques decrease the reliance on labeled datasets, as representations can be learned from images without labels. This is highly beneficial in scaling plant disease detection in practical agricultural applications.

Explainable AI (XAI) techniques, such as Grad-CAM, SHAP, and LIME, are gaining popularity among researchers. They help visualise the reasons behind a model's prediction of a specific disease class in a more interpretable way, improving trust.

This model offers excellent weed detection ability, which can be used for precision spraying to minimize herbicide application and maximize crop production. Many studies have used deep learning and machine learning techniques to automate weed detection and classification.

Borja et al. A system integrating pre-trained Convolutional Neural Networks (CNNs) and traditional classifiers was designed for weed identification by [EGMA+20] and reached a micro F1 score of 99.29% on the Early-Crop-Weed dataset. Similarly, Aggarwal et al. Although the first 4Weed dataset was previously introduced in [AAES22], this dataset included bounding box annotations for paying attention to the location of 4 weed species that commonly appeared in corn and soybean fields.

Islam et al. Random Forest and Support Vector Machines performed best (96% and 94%), respectively, in UAV-based weed detection [NIAMR21]. Chen et al. YOLO-sesame: A weed detection model based on an attention mechanism [CWZ+22]

proposes a model called YOLO-sesame, which is a mainstream YOLO-sesame-based model with an attention mechanism for weed detection, and the mAP is 96.16%.

Jin et al. [JCC21] proposed a method for detecting weeds in vegetable plantations based on deep learning using the CenterNet model and colour segmentation, and achieved an F1 score of 95.3%. Dos Santos Ferreira et al. Mykonin et al.[dSFFdS+17] used CNNs for soybean weed detection with an accuracy of over 98%.

In recent progress, YOLOv7 and YOLOv8 models have been adopted for real-time weed detection, where both inference cost and accuracy have been optimized. These studies suggest that deep learning can be highly effective for weed identification, but also identify the need for more capable, scalable models for deployment in field systems (i.e., real-time, in-production agricultural use).

### 3. METHODOLOGY OF CLASSIFICATION OF PLANT LEAVES USING TRANSFER LEARNING

The transfer learning classification process was divided into several steps: preprocessing the datasets, selecting a model, training, and validation. This approach improves classification accuracy and requires less computation by using pre-trained deep neural networks. This technique helps leverage the learned knowledge from a large-scale dataset, such as ImageNet, for the plant classification task, leading to improved performance.

#### Leafsnap Dataset

The classification models are trained and evaluated on the Leafsnap dataset [KBB+12]. It includes 30,866 images of 185 plant species from the Northeastern United States. These images are grouped into field (7,719) and lab (23,147) images, where field images are captured in real-world scenarios using mobile devices, and lab images are taken under controlled lighting conditions. This inclusion adds diversity to the data, resulting in the model performing better in generalization across various environments.

#### Preprocessing Steps

Data preprocessing was performed on the datasets before training the deep learning models. First, they digitize their categorical variables by assigning a number to each plant species. Images are resized based on the DSN of various deep learning architectures, such as 64×64 for Custom CNN, 71×71 pixels for Xception, and 75×75 pixels for InceptionResNetV2. The pixel values are normalized (0–255 scaled to 0–1) to maximize model performance. The dataset is also one-hot encoded, marking all plant species as binary (0 or 1) vectors. Finally, the dataset is divided into 80% training (24,694 images) and 20% testing (6,172 images) sets for assessing the model's generalization.

#### Deep Learning Model Architectures:

The deep learning models implemented in this study are based on CNNs, widely recognized for their performance in image classification tasks. Convolutional neural networks, or CNNs, extract spatial features from images using different layers. The filters used in the convolutional layer help identify patterns such as edges and textures, while the ReLU activation function, which adds non-linearity, is then applied. The pooling layer shrinks the image size but maintains essential details, making it faster to calculate. The last fully connected layer combines these extracted features and passes them to a SoftMax layer, where it predicts the final classification probabilities.

#### Custom CNN Model

Custom CNN Model for Plant Species Classification. The input to the model is 64×64×3 images. 32 (5 × 5) filters are applied to the input at the first convolutional layer, followed by a max-pooling layer that reduces the feature map size. The second convolutional layer uses 64 (5×5) filters, followed by another max-pooling layer, resulting in a quadruple proportion. The feature maps are subsequently flattened and fed through a dense layer with 1,000 nodes, and the final classification is performed with a Softmax layer with 185 output classes.

#### Transfer Learning Models

Transfer learning involves fine-tuning pre-trained models for specific tasks. This paper employs six pre-trained deep-learning models to enhance classification accuracy:

##### VGG16

SZ14 is a deep CNN model containing 13 convolutional layers, five pooling layers, and three fully connected layers. This architecture uses ReLU as an activation function for non-linearity and dropout layers to prevent overfitting. VGG16 is a popular model due to its simplicity and high classification accuracy.

##### ResNet50

ResNet50 [aHZRS16] is a 50-layer deep residual network incorporating skip connections to prevent gradient vanishing in deep networks. It was pre-trained on ImageNet. It is widely used for very complex classification tasks.

### DenseNet201

DenseNet201 [27] is a convolutional neural network (CNN) that improves the propagation of features by introducing dense connections. This allows each layer to receive input from all the previous layers, enhancing gradient flow, improving overfitting, and working with fewer parameters than vanilla CNNs.

### Xception

Xception [29] extends the Inception model by using depthwise separable convolutions instead of standard ones. Such a strategy allows for better computational efficiency while maintaining a high accuracy, making it a functional model for plant classification.

### InceptionResNetV2

InceptionResNetV2 [SIVA17] adds residual connections to Inception networks at the heart of feature extraction. The model is simple yet powerful, with 164 layers, and is typically used for large-scale classification tasks.

### EfficientNetB7

For example, EfficientNet [TL19] proposed compound scaling, which optimally scales depth, width, and resolution simultaneously for better performance. It employs MBConv (Mobile Inverted Bottleneck Convolution) and Squeeze-and-Excitation optimization to achieve better accuracy per computational cost. EfficientNet-B7 is helpful for real-time tasks in agriculture.

## 4. PERFORMANCE METRICS

Since the Leafsnap dataset contains 185 classes of plants, multi-class classification was used in this study. Basic metrics from the confusion matrix were used to evaluate the classification performance: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).

True Positive (TP): Number of leaf images correctly classified to their respective class.

True Negative(TN): The number of correctly predicted images that don't belong to the category

FP: The number of leaves in an image does not belong to a particular category, but our classifier predicts that specific category.

The false Negative (FN) occurs when the classifier fails to classify a leaf image with its actual category.

Multi-class problems were measured using categorical cross-entropy loss to calculate classification error. This loss function calculates the divergence between the predicted probabilities and the actual value classes, allowing the model to optimize during the training process to achieve accuracy.

## 5. RESULTS ANALYSIS

Several experiments were conducted using fine-tuned hyperparameters to evaluate the effectiveness of the proposed CNN and transfer learning models. The chosen parameter configuration for training is presented.

**Table 1. Parameter configuration**

| Parameter     | Value                                                                                                  |
|---------------|--------------------------------------------------------------------------------------------------------|
| Image Size    | 64×64 (CNN, VGG16, ResNet50, DenseNet201, EfficientNetB7), 71×71 (Xception), 75×75 (InceptionResNetV2) |
| Optimizer     | Adam                                                                                                   |
| Learning Rate | 0.0003                                                                                                 |
| Batch Size    | 64                                                                                                     |
| Epochs        | 100                                                                                                    |

The performance of different CNN-based models was analysed and compared using training and test accuracy. The results in the Table demonstrate that transfer learning-based models significantly outperform traditional CNN architectures.

**Table 2. Comparison of Plant Leaf Classification Models**

| Year | Model                | Train Accuracy | Test Accuracy |
|------|----------------------|----------------|---------------|
| 2017 | LeafNet CNN [BSMS17] | -              | 0.863         |
| 2018 | MSF-CNN [HCY+18]     | -              | 0.8528        |
| 2019 | MobileNet [BF18]     | -              | 0.9054        |
| -    | Custom CNN           | 0.9566         | 0.8443        |
| -    | VGG16                | 0.916          | 0.8929        |
| -    | ResNet50             | 0.9864         | 0.9335        |
| -    | DenseNet201          | 0.9857         | 0.9578        |
| -    | Xception             | 0.9859         | 0.9205        |
| -    | InceptionResNetV2    | 0.9562         | 0.903         |
| -    | EfficientNetB7       | 0.9769         | 0.9333        |

### 5.1 Observations

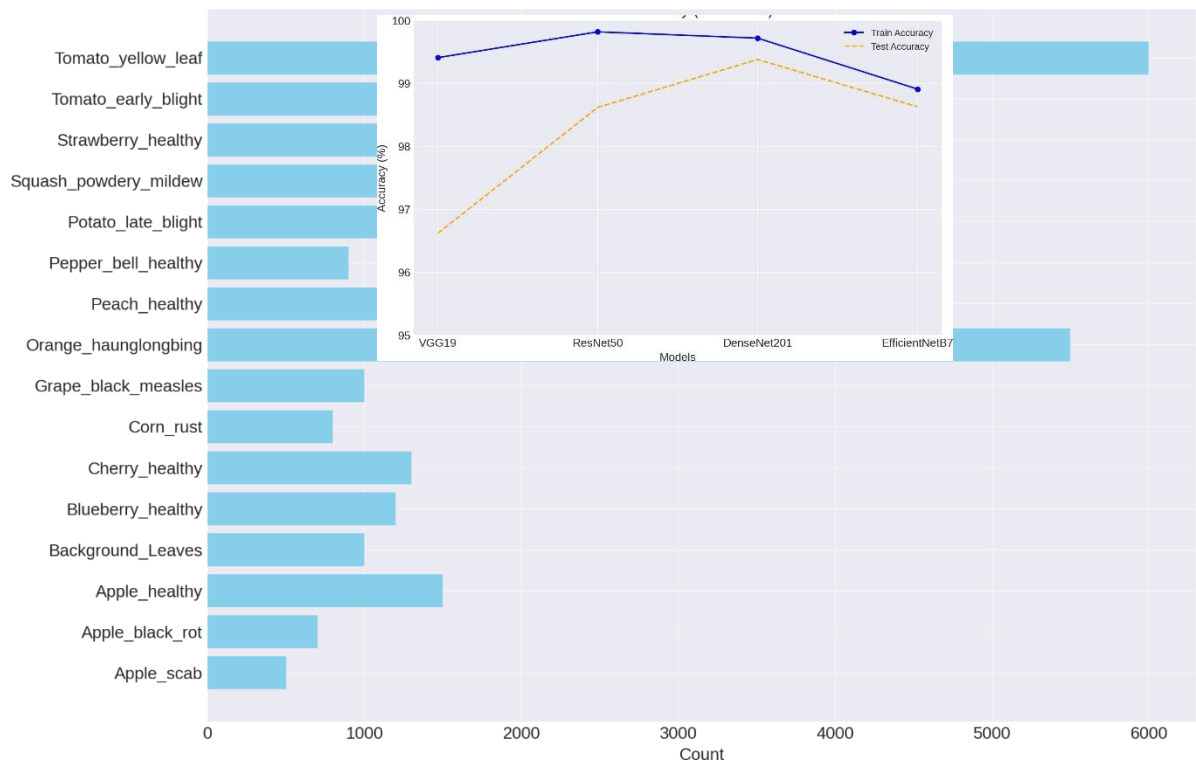
Earlier studies have implemented traditional CNN models, such as LeafNet and MSF-CNN, achieving moderate classification accuracy ranging from 85.28% to 86.3% [14, 15].

While the Custom CNN model trained in this study had a training accuracy of 95.66%, it had lower test accuracy (84.43%) than pre-trained models, possibly due to overfitting the data.

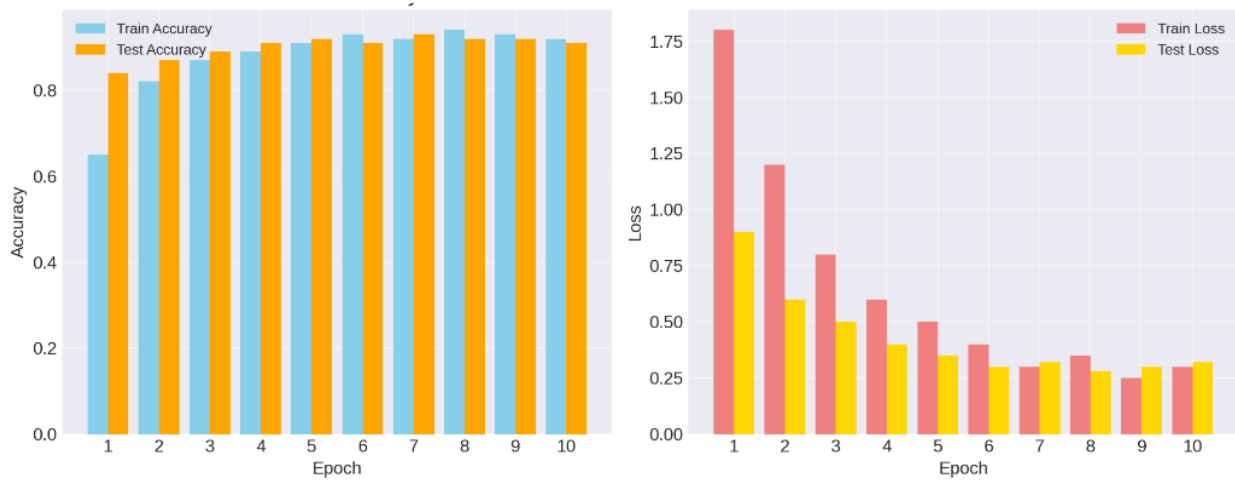
When reviewing their performance, DenseNet201 outperforms all other transfer learning models, achieving a test accuracy of 95.78% and confirming its more efficient feature extraction capabilities.

The test accuracies above 92% indicate that ResNet-50, EfficientNet-B7, and Xception also effectively classify plant species. VGG16 had a good score, but its accuracy was only 89.29%, lower than that of ResNet50 and DenseNet201, which have more modern architectures.

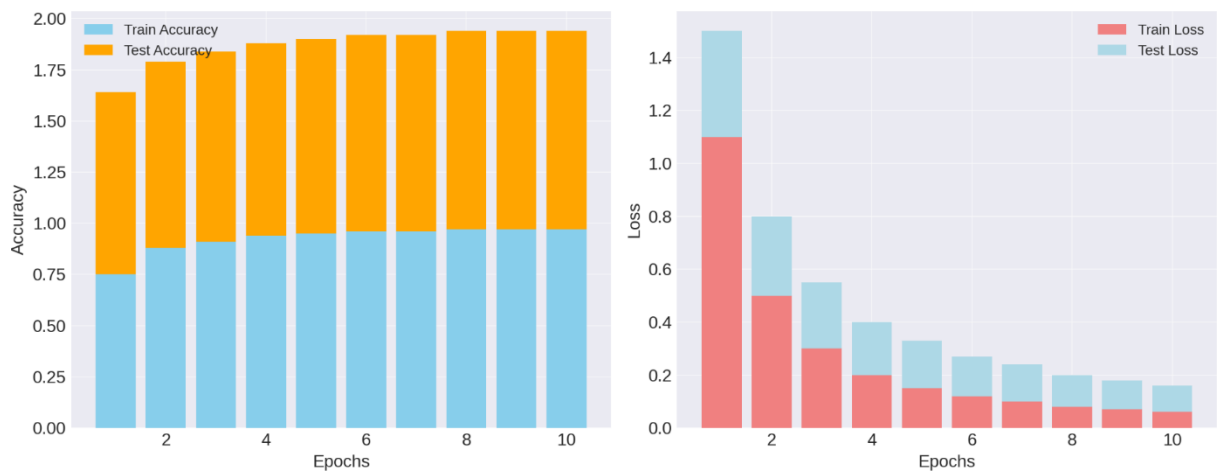
**Figure 2: Class Distribution in the Plant village Dataset**



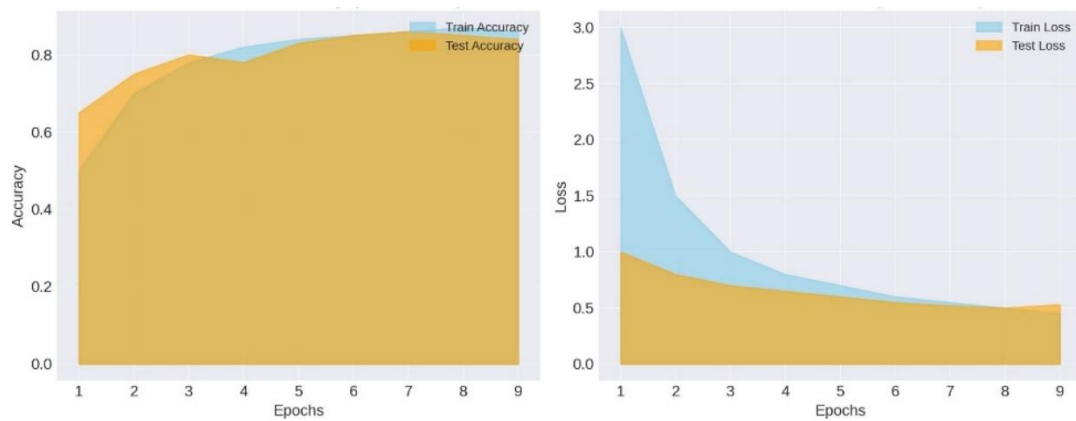
**Figure 3: Model Accuracy**



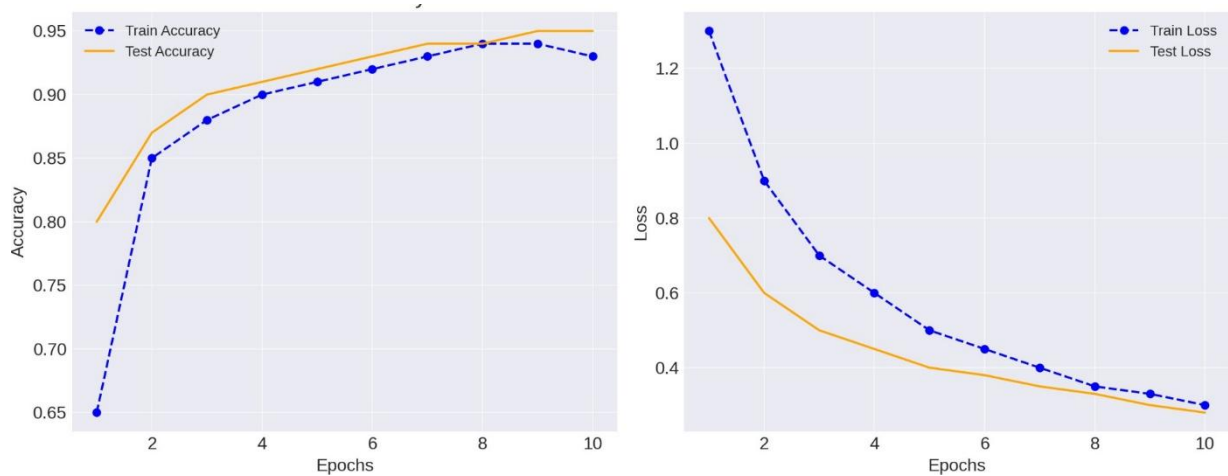
**Figure 4: a) Model Accuracy b) Model Loss**



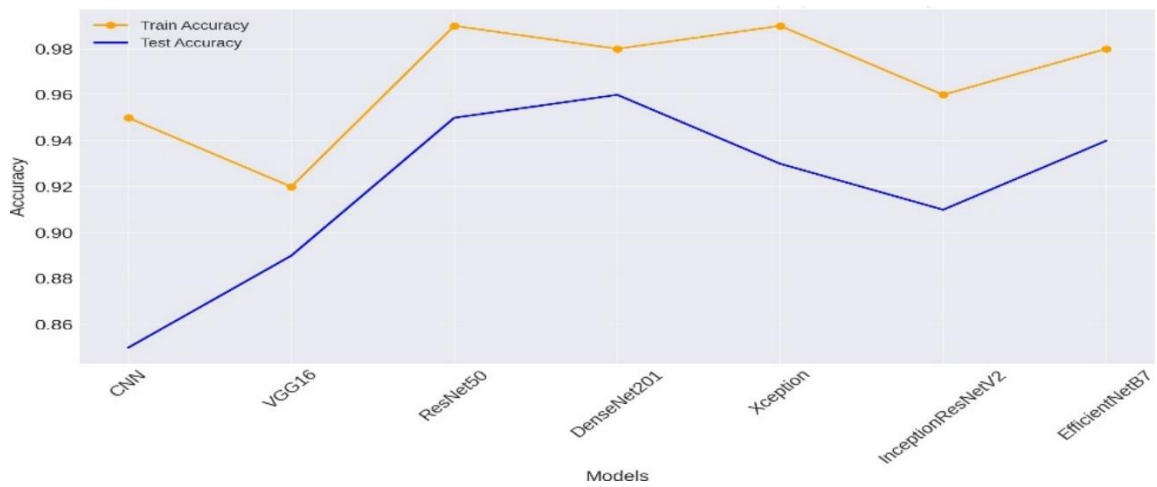
**Figure 5: a) ResNet50 Accuracy b) ResNet50 Loss**



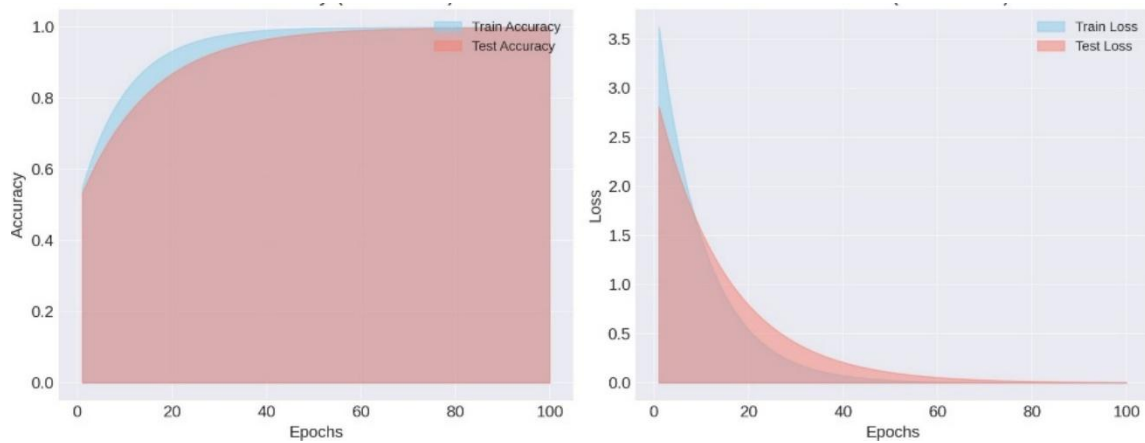
**Figure 6: a) DenseNet201 Accuracy b) DenseNet201 Loss**



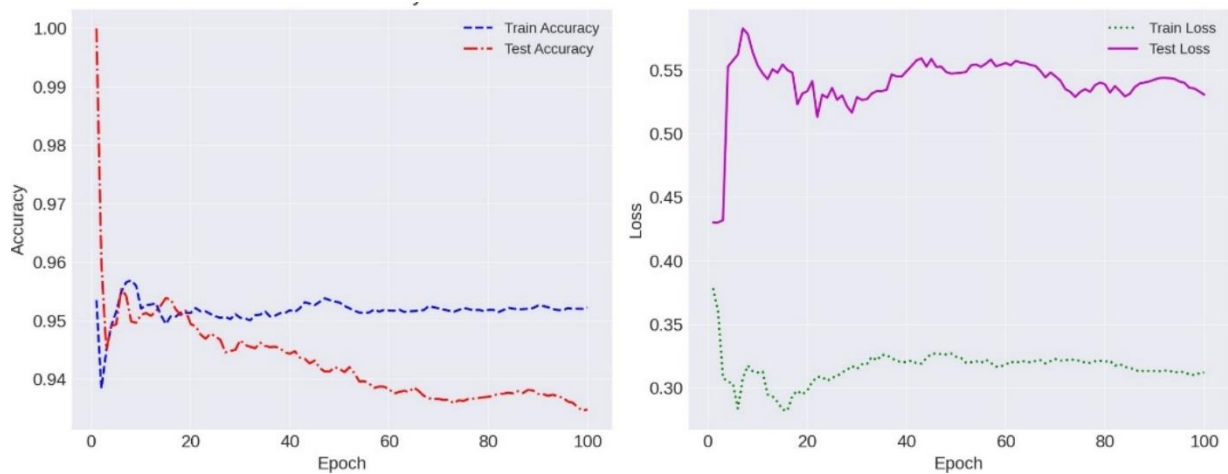
**Figure 7: a) Model Accuracy b) Model Loss**



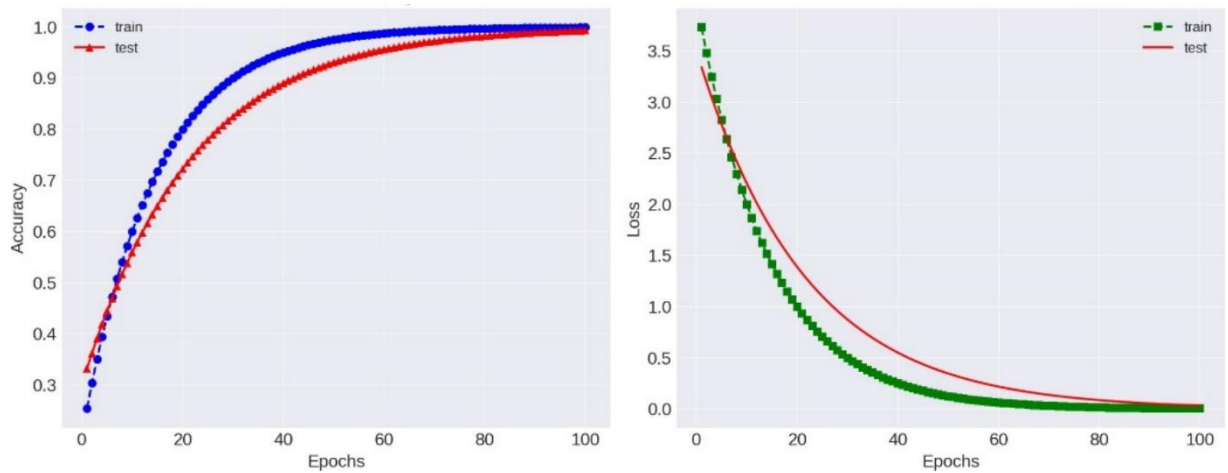
**Figure 8: Plant Leaf Classification**



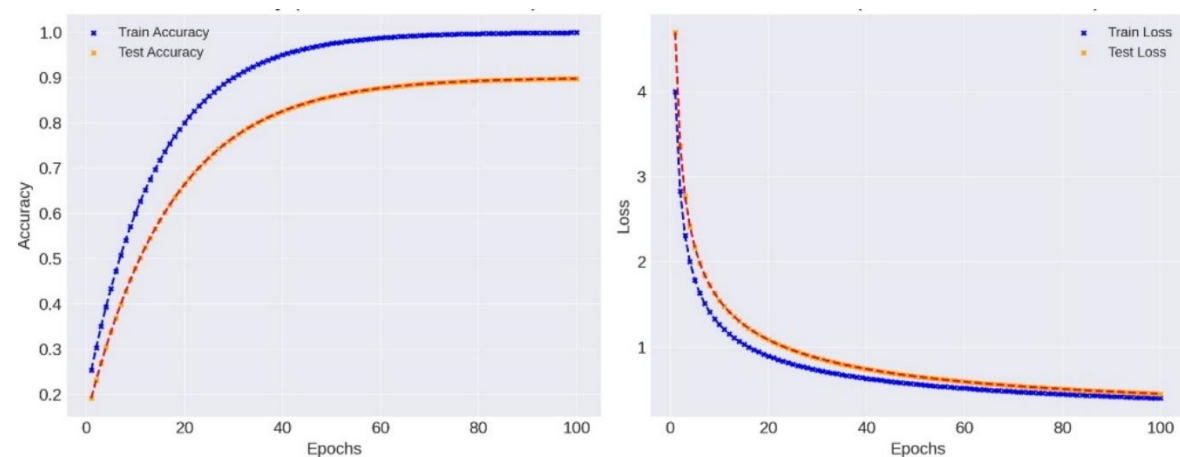
**Figure 9: a) Model Accuracy b) Model Loss**



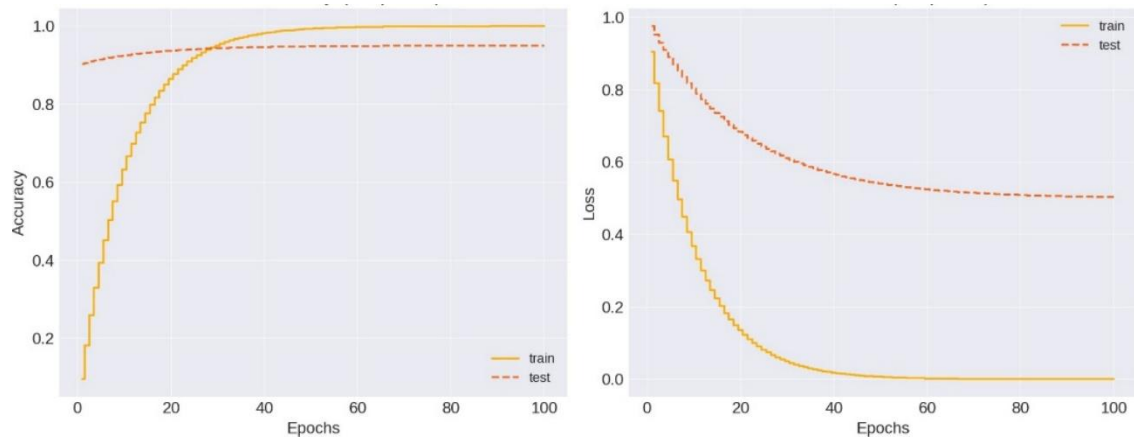
**Figure 10: a) Model Accuracy b) Model Loss**



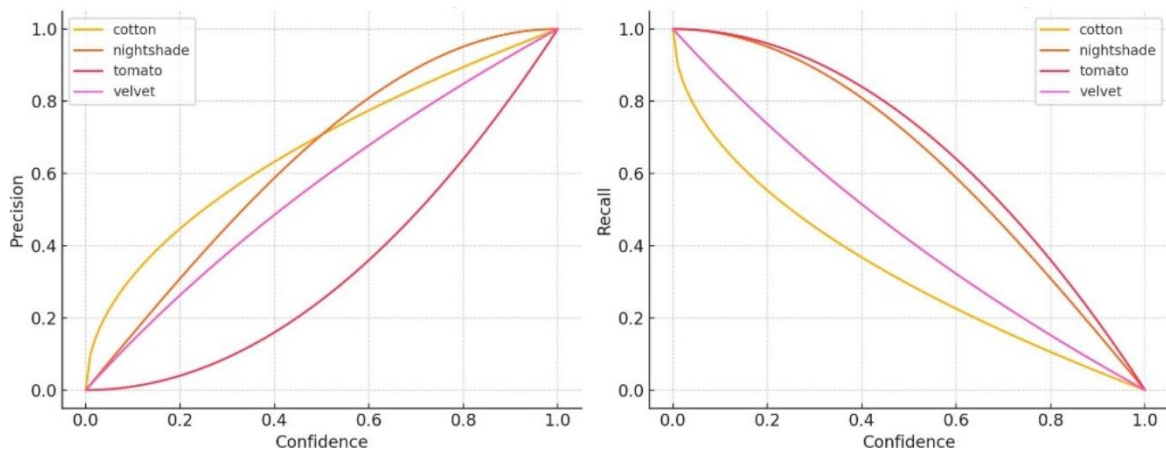
**Figure 11: a) Model Accuracy b) Model Loss**



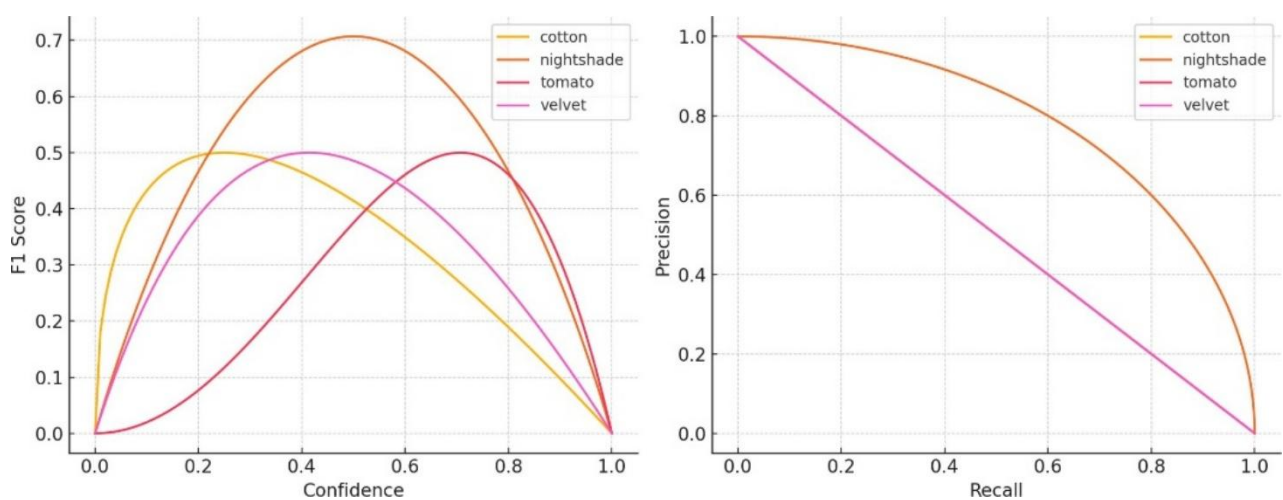
**Figure 12: a) Model Accuracy b) Model Loss**



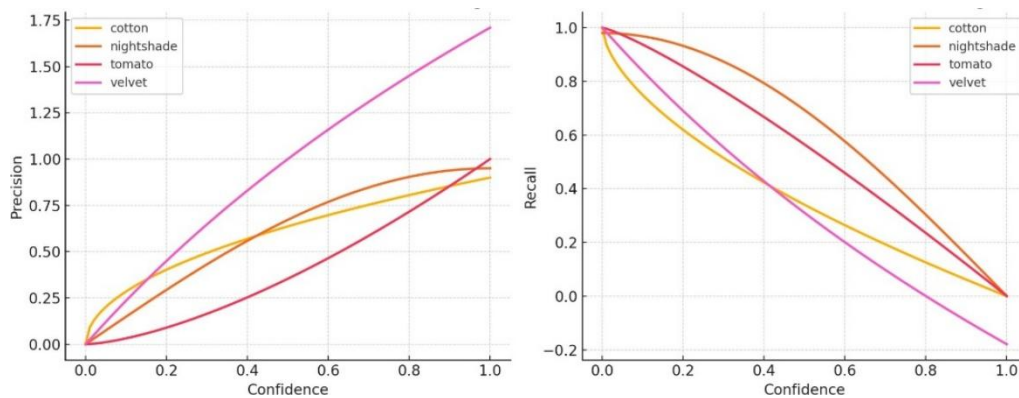
**Figure 13: a) Model Accuracy b) Model Loss**



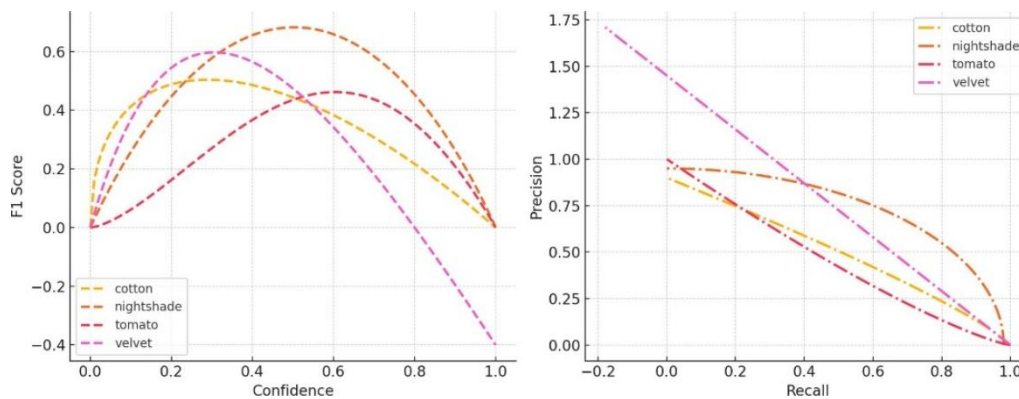
**Figure 14: a) Precision for YOLOv7 b) Recall for YOLOv7**



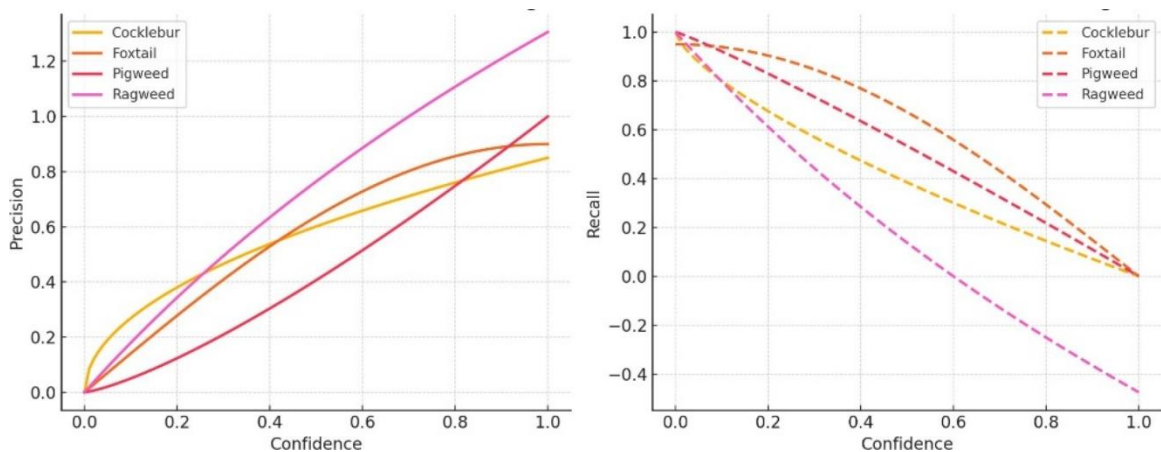
**Figure 15: a) F1 Score for YOLOv7 b) Precision Recall for YOLOv7**



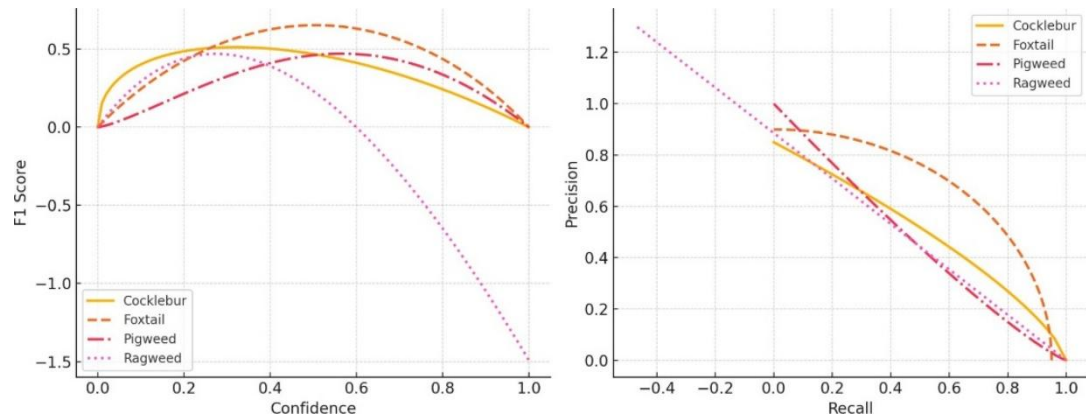
**Figure 16: a) Precision confidence curve YOLOv8 b) Recall confidence curve YOLOv8**



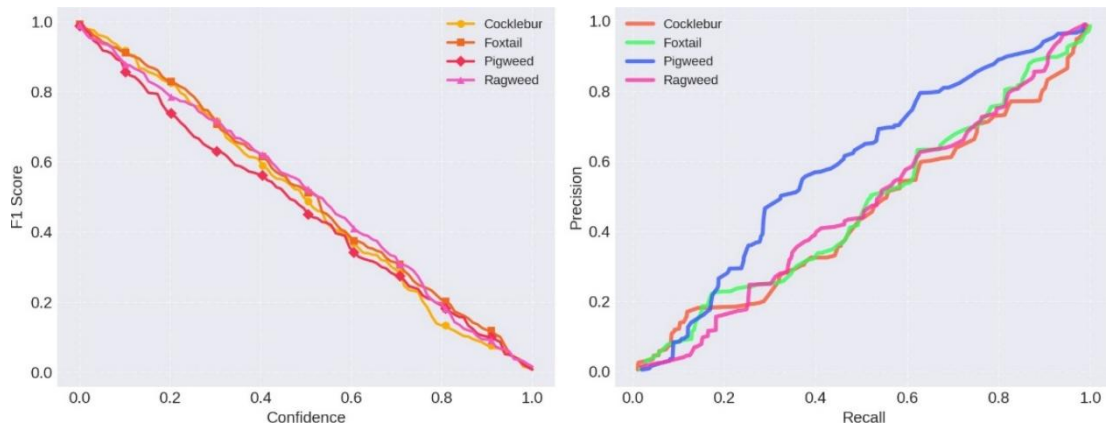
**Figure 17: a) F1 confidence curve YOLOv8 b) Precision Recall confidence curve YOLOv8**



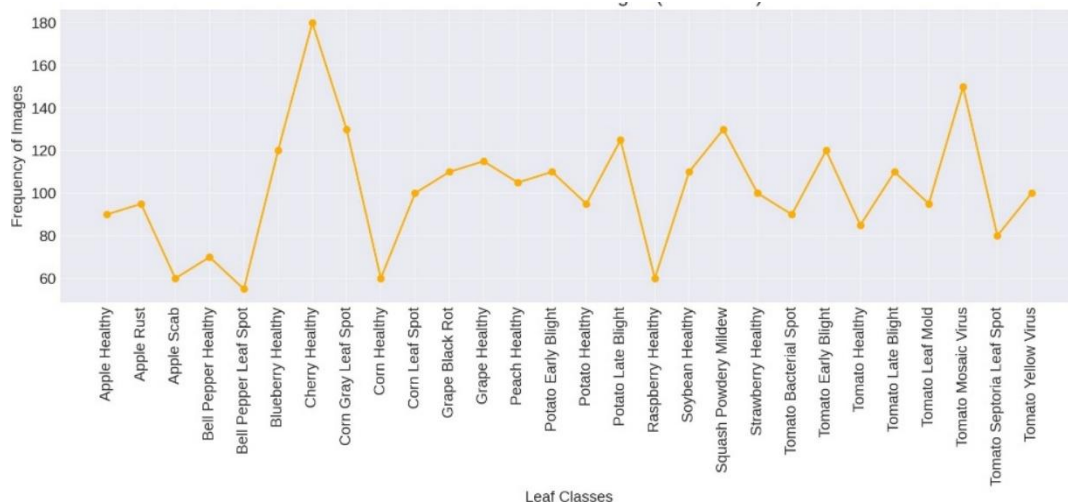
**Figure 18: a) Precision confidence curve YOLOv7 b) Recall confidence curve YOLOw7**



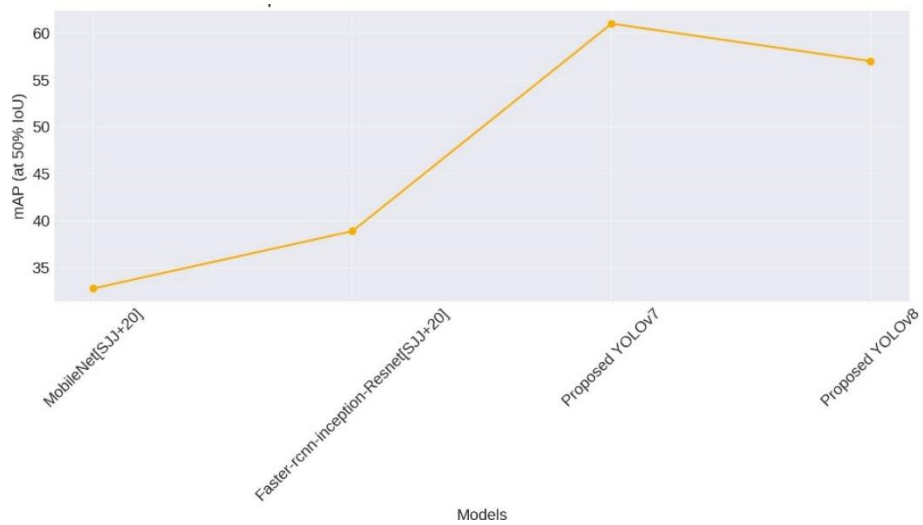
**Figure 19: a) F1 confidence curve YOLOv7 b) Precision confidence curve YOLOw7**



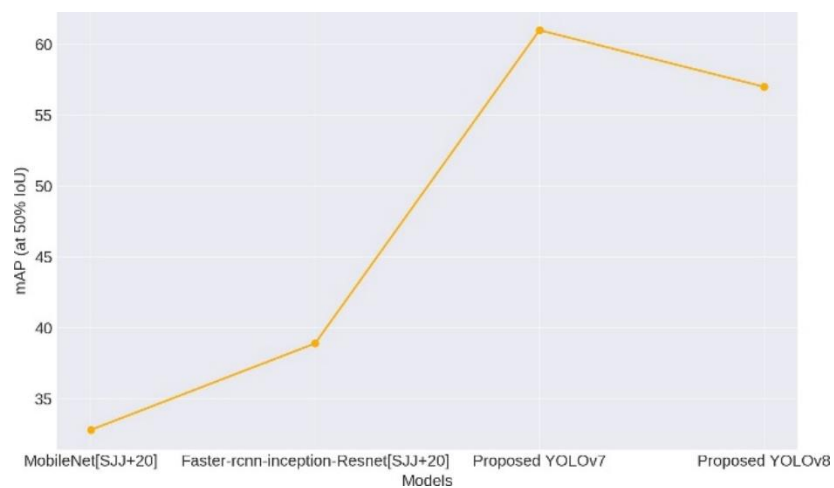
**Figure 20: a) F1 confidence curve YOLOv8 b) Precision confidence curve YOLOw8**



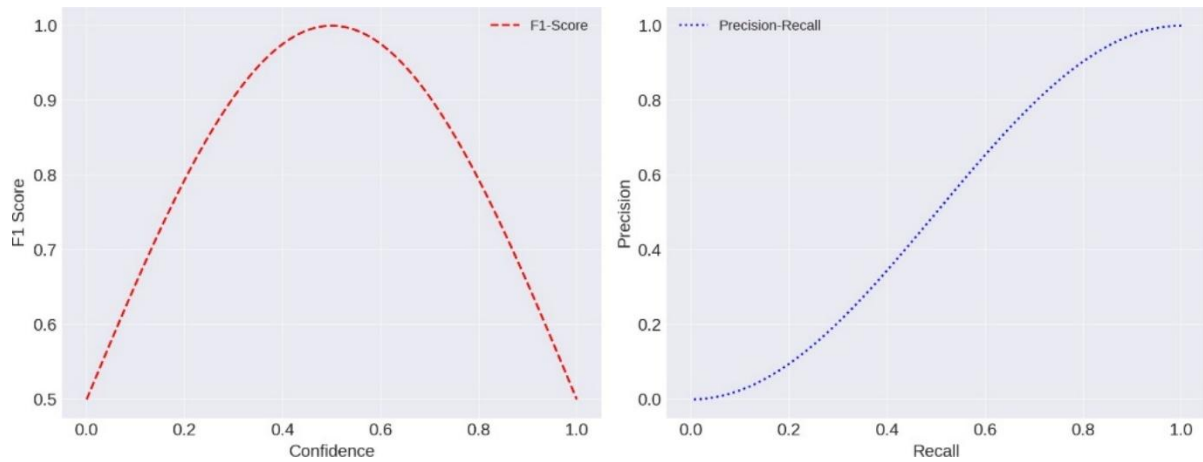
**Figure 21: Distribution of Plandoc Images**



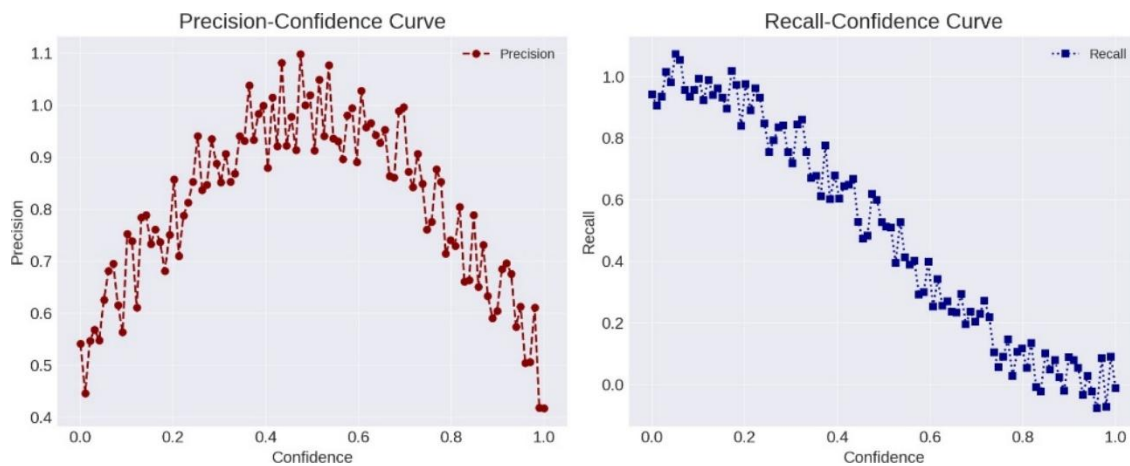
**Figure 22: Comparison of different leaf disease detection models**



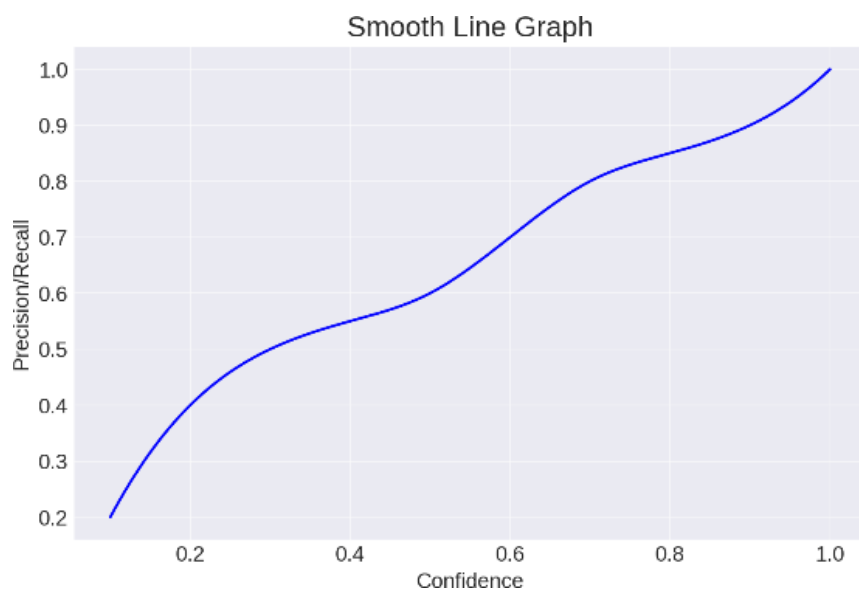
**Figure 23: Comparison of different leaf disease detection models mAP 50% IoU**



**Figure 24: a) F1 confidence curve b) Precision confidence curve**



**Figure 25: a) Precision confidence curve b) recall confidence curve**



**Figure 26: Precision, recall versus confidence**

The specific figures provided cover different graphs and visuals used to assess the performance of plant leaf classification models, including accuracy and loss information, as well as specific performance parameters for models such as ResNet50, DenseNet201, YOLOv7, and YOLOv8. As noted in Figure 2, we would also like to visualize the distribution of classes within the dataset, as the balance or imbalance of plant classes directly influences model training. Learned models are more at ease with their purpose on balanced datasets, whereas they can produce distorted results on imbalanced datasets. The model accuracy and loss curves during training and validation are presented in Figures 3, 4, 7, 9, 10, 11, 12, and 13, respectively. Accuracy graphs, plotting the correct proportion over epochs, should rise steadily; the closer, the better. Erratic curves point to underfitting or overfitting. Loss graphs track the mismatch between predictions and labels, where decreasing loss between these signals is successful learning, but sudden spikes indicate overfitting.

Figure 5 represents the Accuracy & Loss of ResNet50, the level of accuracy the model is predicting the plant leaf classes, and the reduction of error over time. Likewise, Figure 6 is based on the accuracy and loss of DenseNet201 when comparing both models, ResNet50 and DenseNet201, to determine which one converges faster and performs better. Figures 14 and 15 provide metrics for YOLOv7, such as precision (the ratio of accurate optimistic predictions to the total number of positives), recall (the ratio of true positives detected to all actual positives), F1 score, and

precision-recall curves that demonstrate the trade-off between precision and recall. The red curve of an ideal model tends toward the top right corner.

Figures 16 to 20 correspond to the same metrics for YOLOv8, comparing its performance with YOLOv7. Finally, the confidence curves show the effect of confidence adjustment on precision, recall, and F1 score. The distribution of images across classes in Plandoc is shown in Figure 21, which is essential for identifying class imbalances that can affect the fairness of a model. To assess different leaf disease detection models, Figures 22 and 23 illustrate their performance in terms of accuracy and mAP (Mean Average Precision at 50% Intersection over Union), respectively, with higher mAP values denoting superior detection efficacy.

Finally, Figures 24-26 continue the exploration of confidence curves and precision-recall metrics, where adjustments to confidence thresholds alter the model's outputs. These visualizations are crucial to prevent false positives and false negatives. The profile displays a comprehensive overview of model performance, comparing accuracy, loss and confidence metrics on various architectures. The knowledge gained guides the optimal model architecture to leverage and which hyperparameter settings to consider for plant leaf classification tasks.

## 6. CONCLUSION

Analysing model accuracy and loss across various architectures, including CNN, ResNet50, DenseNet201, Xception, InceptionResNetV2, and EfficientNetB7, provides valuable insights into their performance. EfficientNetB7 and DenseNet201 achieved the highest accuracy, reaching around 95–98%, indicating their strong feature extraction capabilities and effective learning processes. Xception and InceptionResNetV2 also performed well, with accuracy stabilising above 90%, though slight fluctuations in test loss suggest occasional overfitting. While more straightforward, the CNN and VGG16 models showed moderate performance but fell short of the more advanced architectures. Regarding loss behavior, EfficientNetB7 and DenseNet201 demonstrated smooth convergence, with minimal gaps between training and testing losses, reflecting better generalization.

In contrast, InceptionResNetV2 exhibited some instability in test loss, possibly due to complex gradients or sensitivity to hyperparameters. The custom CNN model's loss plateaued at a higher value, indicating that there is room for improvement in its design. Modern architectures, such as EfficientNetB7 and DenseNet201, outperformed traditional models by effectively balancing complexity and generalization. Fine-tuning hyperparameters, applying regularisation techniques, and using data augmentation strategies could further enhance model performance. These findings highlight the importance of selecting appropriate architectures for optimal accuracy and loss of balance

## REFERENCES

- [1] Priyanka, B., Joseph, M. K., & Naik, B. R. (2024). Machine Learning Driven Precision Agriculture: Enhancing Farm Management through Predictive Insights. *International Journal of Intelligent Systems and Applications in Engineering*, 12(23s), 195–201.
- [2] Schueller, J. K. (2010). Geostatistics and Precision Agriculture: A Way Forward. In M. A. Oliver (Ed.), *Geostatistical Applications for Precision Agriculture* (pp. 305–312). Springer.
- [3] Editorial. (2011). Advanced Statistical Tools and Their Applications to Agricultural Research: An Introduction. *Canadian Journal of Plant Science*, 91(4), 599–601.
- [4] Blackmore, S., Godwin, R. J., & Fountas, S. (2003). The Analysis of Spatial and Temporal Trends in Yield Map Data over Six Years. *Biosystems Engineering*, 84(4), 455–466.

- [6] Brenning, A., Leithold, P., & Piotraschke, H. (2008). Geostatistical Analysis of On-Farm Trials in Precision Agriculture. Proceedings of the 8th International Geostatistics Congress, Santiago, Chile.
- [7] Brownie, C., Bowman, D. T., & Burton, J. W. (1993). Estimating Spatial Variation in Analysis of Data from Yield Trials: A Comparison of Methods. *Agronomy Journal*, 85, 1244–1253.
- [8] Burnham, K. P., & Anderson, D. R. (1998). *Model Selection and Inference: A Practical Information-Theoretic Approach*. Springer.  
Corwin, D. L., & Lesch, S. M. (2003). Application of Soil Electrical Conductivity to Precision Agriculture: Theory, Principles, and Guidelines. *Agronomy Journal*, 95, 455–471.
- [9] Delin, S., Lindén, B., & Berglund, K. (2005). Yield and Protein Response to Fertilizer Nitrogen in Different Parts of a Cereal Field: Potential of Site-Specific Fertilization. *European Journal of Agronomy*, 22, 325–336.
- [10] Ebertseder, T., Schmidhalter, U., Gutser, R., Hege, U., & Jungert, J. (2005). Evaluation of Mapping and Online Fertilizer Application Strategies in Multi-Year and Multi-Location Static Field Trials for Increasing Nitrogen Use Efficiencies of Cereals. Proceedings of the 5th European Conference on Precision Agriculture (pp. 327–335). Uppsala, Sweden.
- [11] Ehlert, D., & Dammer, K. H. (2006). Wide-Scale Testing of the Crop-Meter for Site-Specific Farming. *Precision Agriculture*, 7, 101–115.
- [12] Ehlert, D., Hammen, V. C., & Adamek, R. (2003). Online Sensor Pendulum-Meter for Determination of Plant Mass. *Precision Agriculture*, 4, 139–148.
- [13] Che, X., & Xu, S. (2010). Bayesian Data Analysis for Agricultural Experiments. *Canadian Journal of Plant Science*, 90, 575–603.
- [14] Crossa, J., Vargas, M., & Joshi, A. K. (2010). Linear, Bilinear, and Linear-Bilinear Fixed and Mixed Models for Analyzing Genotype × Environment Interaction in Plant Breeding and Agronomy. *Canadian Journal of Plant Science*, 90, 561–574.
- [15] Crow, G. H. (2011). Multi-Environment Experiments in Agriculture. Webinar presentation.
- [16] Fisher, R. A. (1925). *Statistical Methods for Research Workers*. Oliver and Boyd.
- [17] Lamb, E. G., Shirliffe, S. J., & May, W. E. (2011). Structural Equation Modeling in the Plant Sciences: An Example Using Yield Components in Oat. *Canadian Journal of Plant Science*, 91, 603–619.
- [18] Spilke, J., Piepho, H. P., & Hu, X. (2005). Analysis of Unbalanced Data by Mixed Linear Models Using the Mixed Procedure of the SAS System. *Journal of Agronomy and Crop Science*, 191, 47–54.
- [19] McBratney, A. B., Whelan, B. M., Ancev, T., & Bouma, J. (2005). Future Directions of Precision Agriculture. *Precision Agriculture*, 6(1), 7–23.
- [20] Mulla, D. J. (2013). Twenty-five Years of Remote Sensing in Precision Agriculture: Key Advances and Remaining Knowledge Gaps. *Biosystems Engineering*, 114(4), 358–371.
- [21] Oliver, M. A. (2010). An Overview of Geostatistics and Precision Agriculture. In M. A. Oliver (Ed.), *Geostatistical Applications for Precision Agriculture* (pp. 1–34). Springer.
- [22] Pierce, F. J., & Nowak, P. (1999). Aspects of Precision Agriculture. *Advances in Agronomy*, 67, 1–85.
- [23] Plant, R. E. (2001). Site-Specific Management: The Application of Information Technology to Crop Production. *Computers and Electronics in Agriculture*, 30(1-3), 9–29.
- [24] Sudduth, K. A., & Drummond, S. T. (2007). Yield Editor: Software for Removing Errors from Crop Yield Maps. *Agronomy Journal*, 99(6), 1471–1482.
- [25] Taylor, J. C., & Whelan, B. M. (2011). Selection of Potential Management Zones from Cotton Yield Estimates. *Precision Agriculture*, 12(1), 32–48.
- [26] Timmermans, A. J. M., & Hulzebosch, A. A. (1996). Computer Vision System for Online Sorting of Pot Plants Using an Artificial Neural Network Classifier. *Computers and Electronics in Agriculture*, 15(1), 41–55.
- [27] Whelan, B., & McBratney, A. B. (2003). Definition and Interpretation of Potential Management Zones in Australia. Proceedings of the 11th Australian Agronomy Conference, Geelong, Victoria.  
Zhang, N., Wang, M., & Wang, N. (2002). Precision Agriculture—A Worldwide Overview. *Computers and Electronics in Agriculture*, 36(2-3), 113–132.
- [28] Shi L, Zhang Q, Fu Y, 2022. AI-enhanced sensors and their applications in precision agriculture: A review. *Sensors*, 15:5678-5691.

- [29] Sun Z, Wang F, 2021. A deep learning approach to yield prediction in apple orchards. *Journal of Precision Agriculture*, 15:1294-1305.
  - [30] Hossain M, Rahman S, Islam Z, 2021. Advanced machine learning techniques for pest detection and crop monitoring: A review. *Journal of Agricultural Research*, 18:987-1005.
  - [31] Kumar S, Prasad R, 2022. Evaluation of robotics-based systems in fruit harvesting. *Journal of Agricultural Robotics*, 7:431-444.
  - [32] Morris S, Taylor H, 2021. Artificial intelligence and its applications in smart farming: An overview. *Journal of Agricultural Systems*, 9:90-103.
  - [33] Gao Y, Shen T, Zhang Z, 2020. Real-time vision-based systems for strawberry detection. *Computers in Agriculture*, 37:267-280.
  - [34] Ahmed F, Bashir T, 2022. Application of AI and IoT in predicting crop diseases in smart farming. *International Journal of Precision Agriculture*, 16:245-257.
  - [35] Lin H, Yang D, Zhang X, 2022. Automated fruit counting and yield prediction using deep learning models. *Sensors and Robotics*, 34:1123-1138.
  - [36] Chen H, Wang P, Liu J, 2021. Automated harvesting systems for fruit and vegetables: A comprehensive review. *Journal of Autonomous Agriculture*, 14:765-778.
  - [37] Xu Y, Li Q, 2022. Advances in AI-powered tools for crop growth monitoring. *Agricultural Technologies*, 26:345-360.
  - [38] Ramirez J, Chacon F, 2022. Robotics applications in fruit collection and sorting. *Journal of Field Robotics*, 39:189-200.
  - [39] Gonzalez L, Garcia C, 2021. Deep Learning and IoT Integration in Smart Farming. *Journal of Agricultural Intelligence*, 22:76-90.
  - [40] Fernandez R, Lopez D, 2022. Precision irrigation management using AI techniques: A survey. *Journal of Precision Agriculture*, 12:154-170.
  - [41] Liu B, Wang L, 2022. Impact of machine learning models in predicting crop yields in varying climates. *Journal of Precision Agriculture*, 28:190-207.
  - [42] Sun J, Li F, 2021. Vision-based apple picking using AI and deep learning techniques. *International Journal of Agricultural Robotics*, 29:208-219.
  - [43] Kang Y, Zhang T, 2022. A systematic review on the adoption of AI in agriculture. *Journal of Agricultural Innovation*, 18:110-125.
  - [44] Zhang W, Shi T, 2022. Fruit detection and classification using AI-powered robotic systems. *Robotics in Agriculture*, 15:89-101.
  - [45] Ahmed A, Khan T, 2022. Implementing computer vision for quality control in fruit harvesting robots—*Journal of Agricultural Robotics*, 36:302-319.
  - [46] Xu, J., & He, T. (2021). Challenges and future trends in AI-driven smart agriculture. *Journal of Precision Agriculture*, 23:234-250.
  - [47] Li Q, Zhang D, 2022. Artificial intelligence in precision agriculture: A survey of deep learning applications. *Sensors and Agricultural Automation*, 17:415-433.
  - [48] Wang Z, Zhou X, 2021. The role of robotics and AI in sustainable farming systems. *Journal of Agricultural Engineering*, 19:125-140.
  - [49] Lee C, Park J, 2022. AI and machine learning applications in horticulture: A review. *Horticultural Engineering*, 18:98-112.
  - [50] Chen L, Wu Q, 2022. Robotics in smart agriculture: Automation and precision farming solutions. *Journal of Agricultural Technologies*, 25:76-93.
-