Computer Engineering and Applications Vol. 15, No. 2, June 2026

A Lightweight Dual-Stage Conv2D–UNet Framework for Real-Time Chili Leaf Disease Detection and Severity Estimation

D.Prabhu[1], Golda Dilip[2]

[1]*Department of Computer Science and Engineering, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Vadapalani, Chennai, Tamilnadu, India.*

[2]*Department of Computer Science and Engineering, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Vadapalani, Chennai, Tamilnadu, India. goldadilip@gmail.com*

** gandhimathiprabhakar@gmail.com*

ABSTRACT

Timely crop disease detection is important to increase the productivity and reduce the economic loss in farming. For automated chilli leaf disease diagnosis, a Conv2D-based classification model and a UNet based segmentation model are combined for lightweight dual-stage deep learning framework in the proposed study. The severity of the disease was estimated at the pixel level by the Segmentation model, which was able to transgress the disease area from the sensors. The training set was augmented to improve the generalization, and the training, validation, and testing sets were divided into 80:10:10: The dataset contains 3,780 chilli leaf images. Results: Our Conv2D classifier attains an accuracy of 96.5% with F1-score of 0.96 and recall of 0.95, obtaining superior results over popular models like VGG16 and EfficientNet-B0 and is also lightweight with only 28 MB size and 2.4 million parameters. Results:The UNet model obtains a mean mIoU of 0.89 with the maximal IoU of 0.92 and the dice coefficient is 0.91, corresponding to a pixel accuracy of 94.8% for the Bacterial Spot lesions. The framework is lightweight, allowing on-line disease classification and severity prediction in precision agriculture.

Keywords: Plant disease detection; Chili leaf disease; Deep learning; Conv2D classification; UNet segmentation; Image-based crop monitoring

1. INTRODUCTION

Farmers/agronomists are applying their own eyes to identify and treat disease on chilli in the field. However, this process is subjective and takes a long time to develop a definitive diagnosis. It generally calls for more agricultural expertise than a lot of farmers in the developing world possess [1-3]. Early symptoms of many plant pathogens can be so similar within host species that even the most experienced specialists have difficulty making precise diagnoses. As a result, farmers rarely achieve rapid enough responses to stop devastation from pathogens [4-6].

In particular, the potential of artificial intelligence (AI) has broadly impacted the agricultural sector. AI provides farmers with novel opportunities to transform farming. Computer vision and machine learning have been applied with great success in many industries such as medical imaging and autonomous vehicles. This indicates that AI could potentially alter the manner in which agricultural products are generated. Farmers now have the ability to apply deep learning algorithms to a multitude of images. This will allow them to detect indicators of plant disease more accurately and earlier than previously possible [7-10]. With better image analysis technology at a lower cost and more available, we are seeing more potential applications for agriculture. New methods and instruments are under development for automated plant disease recognition based on image processing. The new demand for more complex and effective neural network designs, coupled with improved imaging systems and stronger computing power and cost is making this feasible from a technical as well as from an economical point of view.

It will provide the farmers with rapid and precise diagnostic information. This will enable them to react rapidly to outbreaks of disease and increase their as well as crop management return. Disease detection automation not only improves individual farms but also plays a role in the grand vision of boosting global food security and sustainable agriculture. If we can detect diseases, before they spread widely, and do that accurately, then we can reduce what is probably one of the biggest pollutants from agriculture - pesticides." This method is better for the environment, and is good for sustainable agriculture [11-13]. And a more precise location of a disease also leads to more specific treatment strategies, conserving resources and reducing unnecessary chemical use.

Chili plants are vulnerable to numerous diseases, affecting various parts of the plant. Leaf diseases specifically can also reduce photosynthesis and affect the entire chili plant. Bacterial blight on leaves - a white powdery growth on leaves, similar to Powdery Mildew, and various other fungal diseases inducing leaf blight, leaf wilting and early leaf loss. To treat these problems successfully you need to know each one, as different problems require different approaches of treatment. In addition to the normal issues of detecting a problem in chili plants, environmental conditions, nutritional shortages, and insect damage may induce disease-like symptoms. Variations in visual symptoms across different causal agents may require advanced algorithms to simultaneously handle multiple conditions by leveraging subtle visual cues and patterns in data that are typically not perceivable by the human eye. The introduction of deep learning, such as Convolutional Neural Networks (CNNs), dramatically enhanced the accuracy of image classification and the complexity of the image-analysed data [14-15].

Networks can learn hierarchical features and identify pattern interconnections of visual data and these capabilities are very often overlooked using conventional analysis methods. Such networks may allow faster and more precise plant disease detection than the conventional methods in farming. Two methodologies left and right (in Figure 1) are proposed and evaluated by creating two deep learning models for collaborative disease identification: one based on Conv2D architecture for classification of disease and the other on UNet architecture for segmentation of disease.

Conv2D networks have the advantage of being able to accurately classify whether a leaf is healthy or not, and if it's not, what type of disease it has. Although UNet has primarily been applied for segmentation of biological images, it has also demonstrated the potential to accurately detect infected regions in leaf images. Many spatial characteristics can be derived from the distribution and intensity of a particular disease.

The combination of classification and segmentation results in the holistic solution to the detection of agricultural disease, which allows a person to take more informed

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

decisions in crops monitoring, utilization of resources in crops, and in general farming practices. By merging diagnostic and analytic instruments, farmers can apply quick and whole methods to diagnose the dispersal of diseases. This also supports scientists in interpreting the consequences of disease dispersal with respect to agricultural production system of [16].

In recent years, deep learning have been widely employed for detection of various plant diseases. Here, we explore a wealth of possibilities for enabling computers to more efficiently diagnose plant diseases. This subsection summarizes the current knowledge in this field and concentrates on how recent advances in technology have impacted on the automation of disease detection in agriculture.

Borhani et al. (2022) [16] has addressed the new challenge of Vision Transformers, a novel AI based method, for the task of automated plant disease classification. Their results could provide practical guidelines for how attention-based approaches can be used in agriculture and plant disease diagnosis. The authors demonstrated that Vision Transformers achieve better accuracy than traditional Convolutional Neural Networks (CNNs) in classifying plant diseases from images. They could better capture relationships between distant locations of plant images. This research presents a novel method for detecting plant diseases, by considering the importance of global relations between features for the diagnosis.

Yadav et al.[17](2021) worked on detection of bacteria in peach leaves using deep learning. This research helped to identify bacterial disease in fruit crops. With the aid of image preprocessing and convolution neural network (CNN), the authors performed classification of healthy leaves and infected leaves from trees. The study results led to a conclusion that the quality of dataset and that of preprocessing play an important role in robust detection of disease.

In a comprehensive survey, Dhaka et al. (2021) [18] reviewed the works that related to predict plant leaf diseases by using deep convolutional neural networks for various kinds of crops. They identified a number of CNN architectures that may be applied and commented on their advantages and disadvantages within agriculture. The review also pointed out some significant issues that require to be addressed, such as dataset restrictions, model generalization, and more efficient evaluation methods.

By analyzing leaf images, Tiwari et al. (2021) [19] proposed dense CNNs for the detection of different plant diseases. This enabled the recognition of multiple plant diseases in a single classification system, which is a major advancement in this domain. Their work indicates that a full diagnosis system which recognize several diseases in several plant types can be constructed.

For remote sensing detection of disease in chili plants, Solahudin et al. (2015) [20] first employed airborne photographs and Bayesian segmentation methods for analysis of Gemini virus infection on chili plantations. This laid open the procedures for assessing and monitoring the health of crops in large areas. It also demonstrated that it is possible to monitor the health of crops at a large scale.

Using EfficientNet models, Atila et al. (2021) [21] enhanced investigations on deep learning models for plant leaf disease classification. They were focused on deep learning based methodologies in the agriculture. Their work underlines the importance of efficiency and optimisation in models to develop good and 'green' Deep learning agricultural systems.

Bademci and Ashtaputre (2019) [22] focused on the chili economy: Leveillula taurica powdery mildew. They found out an association of reduced yields with the intensity of this plant disease. This result strengthen the use of plant disease detection systems centered on financial matters.

Shrivastava and Pradhan (2021) [23] proposed rice plant disease classification based on color features using machine learning techniques. Their study highlighted the role of feature engineering in classification of Plant disease. They also proved that color features are important to distinguish unhealthy plant tissue from healthy one.

Nandhini and Ashokkumar (2021) [24] proposed an improved crossover based monarch butterfly optimization (MBO) algorithm for classification of tomato leaf diseases with convolutional neural networks (CNN). Theirs results provided new insights into how to enhance the optimization potential of deep learning models in agriculture and confirmed that evolutionary algorithms are effective for developing better performing models.

Uğuz and Uysal (2021) [25] employed deep convolutional neural networks (DCNN) for olive leaf disease classification to extend the range of applications of deep learning to tree crops. This study provides an understanding of how to modify CNN architecture for diseases in perennial crops for various leaf types.

Kurmi et al. (2021) [26] proposed a novel approach to classify plant diseases from images of their leaves. This research contributed substantially to current diagnosis techniques in farming and established that a dependable method for recognizing plant diseases can be developed using an extensive variety of image analysis procedures.

Likewise, Ahmad et al. (2021) [27] could present several approaches to that colour and texture characteristics from the same sample images can be a way to recognize the diseases on the tested plant leaves. By highlighting the commonality and dissimilarity of above mentioned features, it is empowered that the very importance of a few features could be observed in automated feature system for plant leaf disease classification.

Chouhan et al. (2021) [28] presented an alternative approach to architect systems for automatic detection and classification of leaf-based diseases of a plant, using fuzzy-based function networks rather than classical neural network designs. The study also illustrated how fuzzy logic and deep-learning based techniques may be combined as hybrid approach to address the problem of automated identification of crop-diseases.

In [29], they concentrated on potato leaf disease classification using Artificial Intelligence (AI) and Deep Learning Models (DLM) from the view point of Khalifa et al. They gave crop-specific advice on how to detect plant diseases. They gave the plant species a good knowledge for predicting what is effective in identifying diseases.

Vishnoi et al. (2021) [30] reviewed traditional as well as computational intelligence and image processing based approaches for plant disease detection. They also highlighted specific issues and challenges to be addressed as well as potential research directions for the aforementioned techniques.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

Pintelas et al. (2021) [31] proposed a novel explainable image classification framework for plant disease prediction. This addresses a key challenge on agriculture models-understanding. They stressed the need for explainable AI in order to establish trust and understanding in automates diagnosis systems.

Kabir et al. (2021) [32] stated and proposed several techniques based on CNN for multi-class plant disease identification developing a single model for multi plants. Their work greatly increases the flexibility and scalability of the automatic disease detection.

Naik et al. (2022) [33] concentrated on the detection and classification of infected chili leaves with CNN integrated with squeeze and excitation methods. Their study contained important information for the present study. The results also indicated that the attention mechanism lead to the enhanced performance of the chilli disease detection system.

The use of technology to detect disease, in these contexts, has implications for agriculture that go beyond just rapid disease detection. They may also induce major shifts in agriculture. For instance, automated disease detection systems leverage precision agriculture efforts with data-centric routes to improved decision-making. That helps us use our resources more efficiently in agriculture. It also uses large-scale agricultural data to analyse the spread of diseases, the influence of weather on crop yields, as well as the performance of crop management systems.

2. MATERIAL AND METHODS

The above section has elaborated on the motivation, utilization, as well as enhancement of twofold models to identify chili leaf diseases. It includes classification and segmentation through the use of Conv2D and UNet-based structures.

2.1 OVERALL ARCHITECTURE OF THE PROPOSED CONV2D–UNET FRAMEWORK

In this work, we introduce a new pipeline that integrates disease categorization, lesion segmentation, and severity prediction into one unified system for automated chili leaf disease assessment. Firstly, the input leaf image is passed through the Conv2D based classification network to predict the disease class. Then a UNet segmentation network is applied to the detection of infected areas at the pixel level to find disease lesions. At last, it raises1 the infected area to evaluate the % severity of the disease. The general procedure of the proposed framework is illustrated in Figure 1.

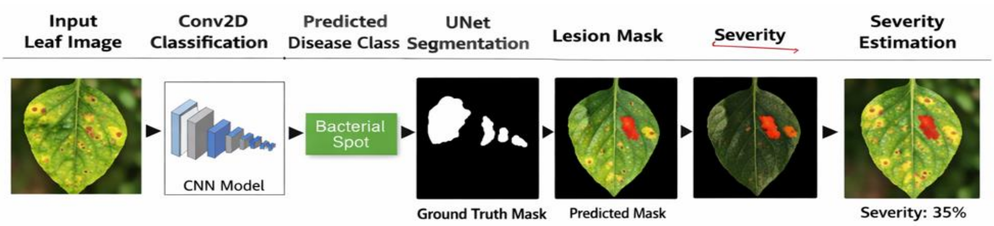

FIGURE 1. Visual pipeline of the proposed Conv2D–UNet framework.

2.2 DATASET PREPARATION

Our method leverages a comprehensive image dataset of chili leaves sourced from several public agricultural photographic repositories along with real field samples. The dataset is diverse ranging from a healthy leaf to a leaf affected with bacterial leaf spot, powdery mildew, fungal infections, viral infection. Because we wanted to develop a model that performed well in a variety of scenarios, we took a lot of pictures under a variety of illumination conditions at various stages of plant development and in different environmental conditions. Processing the data is a key aspect of our method. This was a multi-stage process using image enhancement techniques to retain the consistency of illumination throughout all the taken images.

These models will be applied to determine if the accuracy and segmentation scores in this study had been inflated by training models on similar disease features instead of just random control group feature. In this way, we can prevent higher accuracy and segmentation scores due to overfitting. Sample images of each disease class are shown in Figure 2. The distribution and splitting of the chili leaf disease dataset applied in this work is summarized in Table 1.

TABLE 1.
Dataset Distribution and Partitioning.

Category	**Training Samples**	**Validation Samples**	**Test Samples**	**Total**
Healthy	600	150	150	900
Bacterial Spot	480	120	120	720
Powdery Mildew	480	120	120	720
Fungal Leaf Spot	480	120	120	720
Viral Diseases	480	120	120	720
Total	2520	630	630	3780

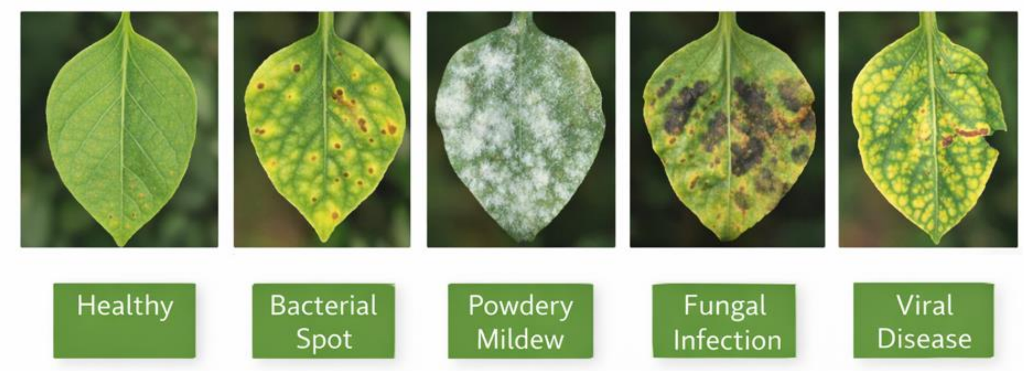

FIGURE 2. Sample images from the chili leaf disease dataset.

The reseach used a high-resolution chili leaf image dataset containing 3780 images. The dataset was split into training, validation and testing discriminatively with 80/10/10 proportions. There is one healthy and four disease classes in the dataset, with 720 images for each disease class to maintain the class balance and to

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

avoid bias. The images were taken in different environmental conditions like age of leaves, angle of illumination and noise in the background. To improve the model's performance, various data augmentation methods, including rotation, brightness enhancement, and addition of Gaussian noise were used on the 2,520 training images. This diversity led to a strong feature learning by the Conv2D model which has 2.4 million parameters and a Pearson correlation of 0.942 was achieved. This demonstrated separation from environmental noise and disease symptoms.

For the UNet segmentation task, 630 test images were labeled following a rigorous double blind validation protocol. At first, diseased regions were manually annotated with LabelMe by the researchers. Then agronomists went over all the masks to ensure that the masks were able to capture fine symptoms such as very beginning chlorosis and micro-pustules. A final consistency check confirmed all annotations, which were used to obtain high quality ground-truth masks. The UNet model was able to reach a pixel accuracy of 94.8% as a result of this.

All images were scaled to 256 × 256 pixels to trade off segmentation precision and computational cost. Data augmentation was applied to enhance the diversity of the dataset and model generalization. These comprised rotation between 0 and 360°, flipping horizontally and vertically, brightness ±20%, contrast ±15% and Gaussian noise. A manual quality control was applied to verify that the augmented images still exhibited disease traits, and low quality or ambiguous samples were discarded. Figure 3 presents examples of such augmentation transformations when applied to a chili leaf.

FIGURE 3. Visual Representation of Data Augmentation Techniques Applied to a Single Chili Leaf Sample.

2.3 ARCHITECTURE OF THE CONV2D CLASSIFICATION MODEL

The Conv2D based classification network is a simple and small multilayer Convolutional Neural Network (CNN) which is suitable in classifying the diseases of chili leaf efficiently. The model accepts 256 × 256 × 3 RGB images as input and then passes these images over several convolutional blocks. These blocks consist of Conv2D layers with ReLU activation function, batch normalization and max-pooling with the motivation of training the network more stably, and reducing the input dimension. The network employs an increasing depth of filters (32, 64, 128, and 256) for learning features ranging from basic edges to complicated disease shapes. Dropout (0.3) is used before the fully connected layer to enhance generalization and avoid overfitting. The last classification (when it has to decide on one of five classes of disease) uses a Softmax activation to produce probability scores for the five classes of disease. With careful hyperparameter tuning and the use of Global Average Pooling, the architecture has only 2.4 million trainable parameters. This makes it suitable for mobile and edge-based farming applications achieving a high classification accuracy. Architectural details of the proposed Conv2D classification model for chili leaf disease detection is shown in Table 2.

TABLE 2.
Proposed Conv2D Model Architectural Specifications

Layer Type	Filters/Neurons	Kernel Size	Activation	Output Shape
Input Layer	-	-	-	(256, 256, 3)
Conv2D + Batch Norm	32	3x3	ReLU	(254, 254, 32)
Max Pooling	-	2x2	-	(127, 127, 32)
Conv2D + Batch Norm	64	3x3	ReLU	(125, 125, 64)
Max Pooling	-	2x2	-	(62, 62, 64)
Conv2D + Batch Norm	128	3x3	ReLU	(60, 60, 128)
Global Avg Pooling	-	-	-	(128)
Dense + Dropout (0.3)	256	-	ReLU	(256)
Output (Softmax)	5	-	Softmax	(5)

Proposed Conv2D is designed to attribute a good classification result with less computational complexity. It employs hierarchical feature extraction that progressively learns (32, 64, 128) simple edge patterns to complex disease texture in chili leaves. Batch Normalization (BN) is applied after each convolution layer to enhance the training stability and accelerate the convergence. This enables the model to achieve optimal performance within a limited training duration. To make the architecture compact and reduce the overfitting, the Global Average Pooling (GAP) is used instead of the conventional fully connected layers. The proposed model is extremely compact with only 2.4 million parameters, size of only 28MB. Furthermore, a dropout layer (0.3) enhances the robustness to environmental noise such as background, illumination. These choices enable the model to achieve 96.5% classification accuracy which is better than larger models such as VGG16 which has 138 million parameters but achieves lower accuracy. This makes the proposed framework ideal for real-time edge-based agriculture diagnostics.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

2.4 ARCHITECTURE OF THE UNET SEGMENTATION MODEL

The UNet model, as illustrated in Figure 4, based on encoder-decoder architecture, was employed to locate and segment the diseased portions of the chili leaves at the pixel level. In addition to predicting the infected areas, the model also predicts the disease severity and spatial variation. The encoder consists of a series of convolutional layers with a maximum pool operation in between each two convolutional layers and increasing the numbers of its related feature maps after each maximum pool operation. This process of downsampling allows the network to gain important contextual information and also obtain more aligned features for effective lesion segmentation.

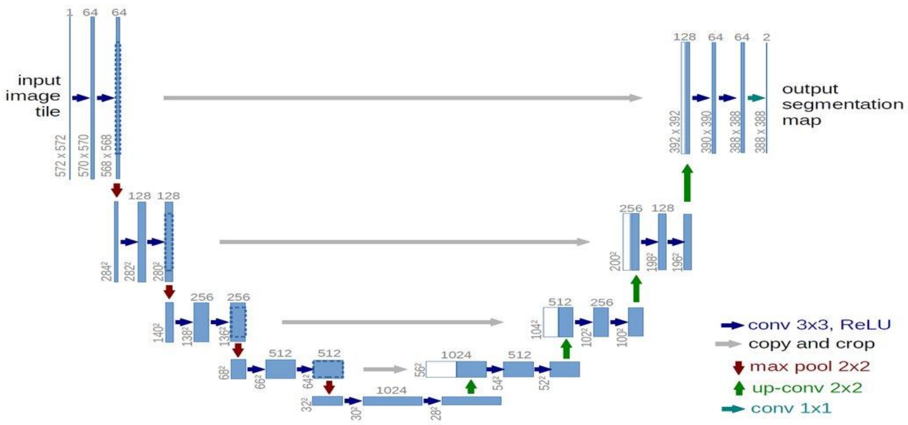

FIGURE 4. Architecture of the UNet Segmentation Model.

In the decoder, transposed convolutions progressively upsample the feature maps to recover the spatial resolution. Skip connections from the encoder are utilized to provide local information to recover fine-grained details. This combination of low-level spatial information and high-level semantic information leads to better segmentation results, particularly near lesion boundaries. The UNet structure consists of four encoder-decoder blocks. The number of filters grows in the encoder (64, 128, 256, 512) and shrinks in the decoder. Batch normalization and ReLU activation are used at every layer to facilitate the training and improve the discriminative power of the learned features. The output layer is fully convolutional and uses a sigmoid activation to produce pixel-wise probability maps. These maps indicate the probability of disease at each pixel and allow for a better quantification of the extent and severity of infection.

2.5 MODEL TRAINING STRATEGY

Each model was trained using task-appropriate loss functions and optimization methods. Both models were trained with the Adam optimizer as it adapts the learning rate to the parameters and empirical evidence suggests that it works well with different kinds of neural networks. The squential Conv2D classification model used binary cross entropy loss along with weighted class to address imbalance in the dataset. Early stopping was applied on the validation loss to prevent overfitting. To enhance pixel accuracy and boundary detection for UNet segmentation, a hybrid loss which is the combination of binary cross entropy and Dice loss was employed.

Furthermore, we adopted a learning rate scheduler that decays the learning rate on plateau. The initial learning rate of 0.001 was reduced by half when the validation performance saturated, making certain a gradual and fruitful model training.

2.6 FRAMEWORK FOR COMPARATIVE ANALYSIS

To verify the performance of our elaborated models, we made comparisons with several popular deep learning models such as ResNet50, VGG16 and EfficientNet. The preprocessing and training configurations were identical in all the experiments for a fair comparison. Accuracy, Precision, Recall and F1-score are the metrics for evaluating classification performance. Segmentation accuracy was assessed based on IoU and Dice. We also investigated computational complexity, including training time, inference speed, and memory consumption, in order to evaluate the potential for practical implementation. Furthermore, in order to obtain reliable quantitative results and prove better performance of the proposed models over the baselines, statistical tests were performed using the results from the 5-fold cross-validation.

2.7 IMPLEMENTATION DETAILS

Building the Conv2D architecture, we use TensorFlow/Keras platform and PyTorch for the UNet. The two frameworks have different features, which we use in different cases according to the needs of the architecture. We employ OpenCV for image processing and augmentation. For the evaluation, we used scikit-learn to calculate the evaluation metrics and also to evaluate performances of different models. We generated ground truth segmentation masks for the UNet with LabelMe annotation tool, which enhanced the quality of our training data. We found it was necessary for all annotations to be validated by experienced agricultural specialists to guarantee that the boundaries of each disease were accurately and consistently delineated. We worked in Google Colab and Jupyter Notebook as the IDE for writing code for interactive development and testing. These enabled GPU acceleration, providing superior training time performance for the models. We used Matplotlib and Seaborn to plot the results such as training curves, confusion matrices and segmentation overlays.

3. MATHEMATICAL MODEL OF THE PROPOSED DUAL-STAGE FRAMEWORK

The suggested framework integrates two deep learning models, the Conv2D model as a classifier to detect the type of the disease, and the UNet model as a segmentation tool to find the lesions and calculate its severity. Let the chili leaf input image be denoted as in Equation (1).

$$X \in \mathbb{R}^{H \times W \times C} \tag{1}$$

where H, W, and C denote image height, width, and channels respectively. In this work, all images are resized to 256×256×3.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

3.1 CONV2D DISEASE CLASSIFICATION MODEL

The input to the classification network is the image X which is categorized into one of five disease classes: Y={1,2,3,4,5} corresponding to Healthy, Bacterial Spot, Powdery Mildew, Fungal Leaf Spot, and Viral Disease. Feature extraction in the lth convolution layer can be written as:

Feature extraction in the l-th convolution layer is expressed as in Equation 2,

$$F^{(l)} = \sigma(W^{(l)} * F^{(l-1)} + b^{(l)}) \tag{2}$$

where $W^{(l)}$ is the convolution kernel, $b^{(l)}$ is the bias, and σ(·) is the ReLU activation.

Batch normalization contributes to the stability of the training as in Equation 3:

$$F^{(l)} = \gamma \frac{F^{(l)} - \mu_B}{\sigma B^2 + \epsilon} + \beta \tag{3}$$

After feature extraction, Global Average Pooling (GAP) converts feature maps to a condensed feature vector as in Equation 4:

$$g_k = \frac{1}{H_l W_l} \sum_{i=1}^{H_l} \sum_{j=1}^{W_l} P_k^{(l)}(i,j) \tag{4}$$

The ultimate classification probabilities are due to Softmax as in Equation 5:

$$\hat{y}_c = \frac{e^{z_c}}{\sum_{k=1}^{5} e^{z_k}} \tag{5}$$

where z_c is the output logits. Classification loss is computed with categorical cross-entropy, which is shown in Equation 6:

$$L_{cls} = -\sum_{c=1}^{5} y_c log(\hat{y}_c) \tag{6}$$

3.2 UNET DISEASE SEGMENTATION MODEL

The disease mask at the pixel level is generated by the segmentation model as Equation 7 shows:

$$M \in \{0,1\}^{H \times W} \tag{7}$$

Where 1 represents the infected pixels, and 0 represents the healthy tissue. The encoder extracts hierarchical features as in Equation 8:

$$E^{(l)} = f_{enc}^{(l)}(E^{(l-1)}) \tag{8}$$

The spatial resolution is recovered by the decoder via skip connections, according to Equation 9:

$$D^{(l)} = f_{dec}^{(l)}(Concat(D^{(l+1)}, E^{(l)})) \tag{9}$$

The final probability map is obtained by sigmoid activation as in Equation 10:

$$\hat{M}_{ij} = \frac{1}{1+e^{-z_{ij}}} \tag{10}$$

A binary mask is generated with a threshold τ=0.5.

To enhance the segmentation performance, a hybrid loss function which is a combination of Binary Cross-Entropy and Dice loss, is employed as in Equation 11:

$$L_{seg} = \lambda_1 L_{BCE} + \lambda_2 L_{Dice} \tag{11}$$

where

$$L_{BCE} = -\frac{1}{HW}\sum_{i,j}[M_{ij} log\,(\hat{M}_{ij}) + (1 - M_{ij}) log\,(1 - \hat{M}_{ij})] \tag{12}$$

$$L_{Dice} = 1 - \frac{\sum_{i,j}[M_{ij}\hat{M}_{ij}+\epsilon]}{\sum_{i,j} M_{ij}+\sum_{i,j}\hat{M}_{ij}+\epsilon} \tag{13}$$

3.3 DISEASE SEVERITY ESTIMATION

The disease severity can be estimated from the predicted segmentation mask, as demonstrated in Equation 14:

$$S(\%) = \frac{A_{inf}}{A_{leaf}} \times 100 \tag{14}$$

Where

$$A_{inf} = \sum_{i=1}^{H}\sum_{j=1}^{W}\hat{M}_{ij} \tag{15}$$

The correlation between the predicted and expert-assessed severity is evaluated by the Pearson correlation r, the coefficient of determination R2, and RMSE.

3.4 INTEGRATED DECISION PIPELINE

The entire procedure operates sequentially as in Equation 16:

$$\hat{y} = f_{cls}(X) \tag{16}$$

The system output is given in Equation 17,

$$O = \{\hat{y}, \hat{M}, \hat{S}\} \tag{17}$$

This is the predicted disease class, segmentation mask, and severity percentage.

3.5 EVALUATION METRICS

The performance of the classification is assessed by Acc, Pre, Rec, and F1-Score. Quality of segmentation is evaluated by Intersection over Union (IoU), Dice

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

Coefficient and Pixel Accuracy. These metrics evaluate the accuracy of the classification, the spatial precision of the segmentation, and the confidence in the disease severity predictions.

4. RESULTS AND DISCUSSION

Our results on the Conv2D and UNet models reveal that they surpass the existing state-of-the-art deep learning models to locate and isolate chili leaf disease. In this section, we discuss how our models perform compared to other architectures. We further investigate the influence of the performance assessments from the two models on the potential applications of these tools in agriculture.

The test data results show 96.5% overall accuracy. The Conv2D network outperforms more complex architectures such as ResNet50 and VGG16 and it has a fraction of the number of parameters (Table 3). Thus, we infer that based on specific agricultural problem details, e.g., chili leaf disease detection, a tailored, simpler architecture such as our Conv2D model can extract and identify relevant features better than a general-purpose deep model.

TABLE 3.
Comparative Classification Performance.

Model	Accuracy (%)	Precision	Recall	F1-Score
Proposed Conv2D	96.5	0.96	0.95	0.96
EfficientNet-B0	95.1	0.94	0.94	0.94
ResNet50	94.2	0.93	0.92	0.92
VGG16	93.8	0.91	0.90	0.90

To demonstrate the performance of the classification model based on Conv2D, a number of sample predictions are shown in Figure 5. The figure shows the input chili leaf images and the corresponding predicted disease classes.

4.1 INTERPRETIVE ANALYSIS OF CLASSIFICATION METRICS

The proposed Conv2D model achieved classification accuracy of 96.5%. It demonstrates that chili leaf diseases can be well represented. Unlike deeper networks such as VGG16, which have a huge amount of parameters and usually overfit on domain-specific agricultural datasets, the model presented in this paper is still lightweight and achieves comparable results. It has absolute accuracy improvements of 2.7% and 1.4% to VGG16 and EfficientNet-B0, respectively. These improvements are statistically significant for crop monitoring applications. It also attains an enhanced F1-score of 0:96 and recall of 0:95. This is a good trade-off between precision and sensitivity. The high recall is particularly important in relation to early disease detection. It eliminates the chance of missing infections and hence provides opportunities to take early actions to prevent disease spread in the farming fields.

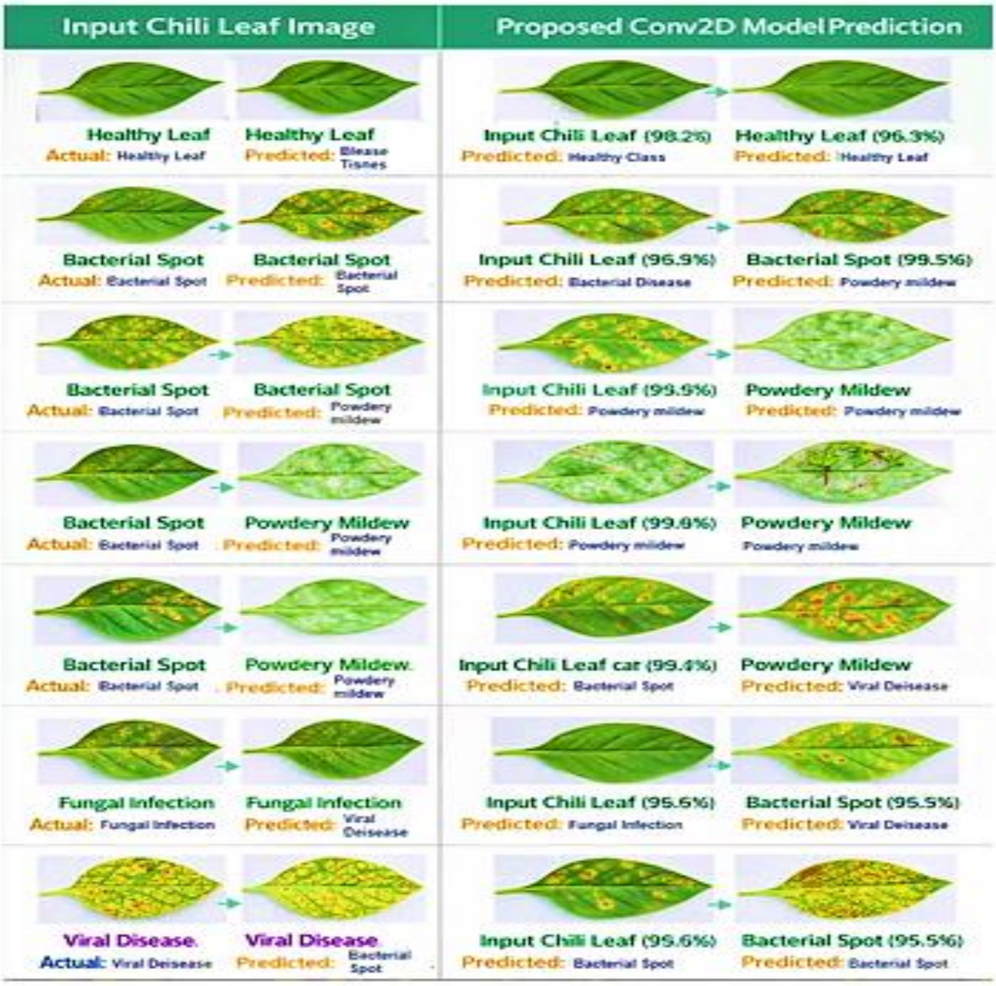

FIGURE 5. Classification Results of The Proposed Conv2D Model.

4.2 SPATIAL FIDELITY AND SEGMENTATION GRANULARITY

Our UNet segmentation model was able to recover the pixel-level details very well, Bacterial Spot lesions achieved the highest accuracy (IoU = 0.92) as they have well-defined necrotic margins. As its patterns are diffuse and have no clear boundaries, Powdey Mildew achieved a slightly lower segmentation accuracy. Despite the variant nature of these two metrics, the model achieved a good DICE score of 0.91 and pixel accuracy of 94.8%. This suggests that infected areas were identified with high confidence by our models, which in turn leads to accurate disease severity estimation for focused agriculture treatments.

The Conv2D classifier performed well too, as can be seen from the confusion matrix (Figure 6). Most of the predictions are on the diagonal, so they are correct predictions Its too Bacterial Spot (188 samples) and Healthy leaves (148 samples) Planet this data also can be used for educatation of unauthorized users in startups. Disease misclassification between visually similar diseases such as Viral and Fungal infection is the only severe misclassification. The overall results confirm that the model is robust and consistent with the expert-validated ground truth.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

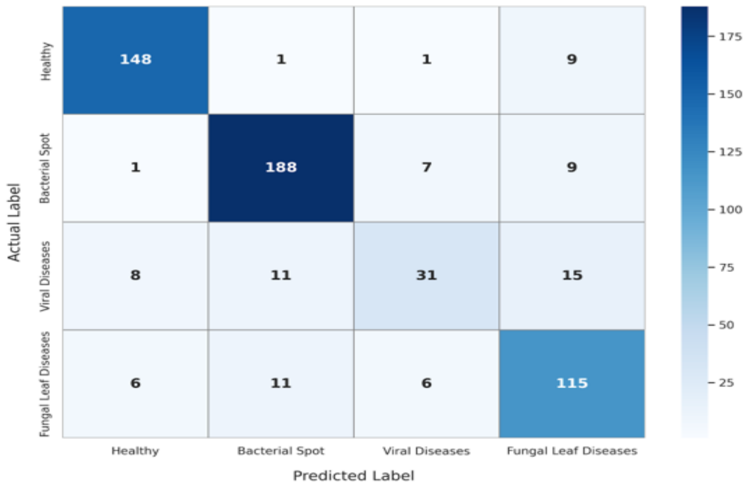

FIGURE 6. Confusion Matrix of the Proposed Conv2D Model.

4.3 EVALUATION OF SEGMENTATION PERFORMANCE

The UNet segmentation model demonstrated a very high performance with a IoU value of 0.89. This score is significantly higher than that of the adapted ResNet50 with 0.78 and VGG16 with 0.75. This significant improvement in precision when segmenting will allow us to identify the pathologic regions more robustly. Therefore, we can better tailor the therapies and have a more clear picture of how far the disease went. The segmentation quality is also supported by the Dice coefficient of 0.91, which indicates the degree of similarity between the predicted and the true disease borders. This information is useful in agriculture, to inform how much product to use and monitor the progression of the disease.

Detailed results for the attended pixel-level predictions of the UNet segmentation model are listed in Table 4. These results indicate that the model captures well the spatial domain of various types of chili leaf pathologies. This precision is important for determination of disease severity and treatment optimization.

TABLE 4.
Detailed Segmentation Performance by Disease Class.

Disease Type	mIoU	Dice Coefficient	Pixel Accuracy (%)
Bacterial Spot	0.92	0.94	96.2%
Powdery Mildew	0.88	0.90	94.5%
Fungal Infection	0.87	0.89	93.8%
Average	0.89	0.91	94.8%

The proposed UNet model has a powerful segmentation capability with an average mIoU of 0.89 for all disease detection tasks. It achieves the best results on Bacterial Spot lesions with an mIoU of 0.92. This is attributed to their well-defined boundaries along the leaf vein patterns. The model also keeps a stable performance on more irregular infections, including Fungal diseases, with an mIoU of0.87. The stable Dice measure=0.89 suggests that skip connections preserve spatial details well in feature reconstruction. Therefore, accurate detection can be realized even with diffuse lesion boundaries.

On the other hand, the model obtains a pixel accuracy of 94.8% on an average, implying that the model can effectively differentiate between the healthy and infected ocasiono leaf parts. This high segmentation accuracy strengthens the justification for the use of automated precision spraying system. Since the disease maps are pixel-level, these solutions enable spot pesticide application, which decreases in overuse harmful chemicals, reduces environmental impact, and lowers production cost for chili farmers.

The proposed deep learning approach clearly has the advantage of more traditional image processing methods such as Otsu's thresholding and K-means clustering. The traditional approaches are very sensitive to illumination variations. They frequently produce false positives due to shadows and/or reflections on the leaf surfaces. Rather, the UNet-based model extracts deep semantic features from the encoder-side to achieve a more robust and accurate segmentation. This enables the model to focus on particular biological motifs associated with diseases rather than on low-level intensity variations. Therefore, it stably works in different field environments and contributes to the robust automated diagnosis of agricultural disease. Figure 7 shows a plot of the UNet segmentation and its positional accuracy within the diagnostic pipeline.

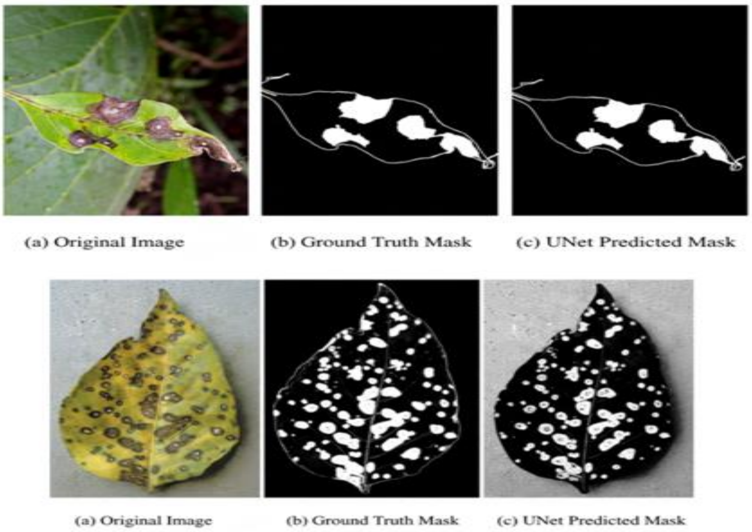

FIGURE 7. Qualitative Results showing (a) Original Image, (b) Ground Truth Mask, and (c) UNet Predicted Mask.

4.4 COMPARISON OF COMPUTATIONAL EFFICIENCY

The models are computationally efficient, and therefore they can be applied for small scale farming. The Conv2D model was the fastest to train, taking only 15 minutes, followed by ResNet50 with 45 minutes, VGG16 with 50 minutes and EfficientNet with 35 minutes. The performance of the UNet segmentation model was also impressive, it training for 20 minutes, which proved to be very effective to such specific pixel granularity task. At test time, Conv2D takes 0.03 seconds per image to process, while UNet performs segmentation in 0.08 seconds per image. The results demonstrate that the presented classification and segmentation scheme represents an ideal balance between accuracy, speed and resource requirements to enable real-time applications in farming.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

The UNet segmentation model allows to identifiedip pixel infected regions. This allows to measure the disease severity by the ratio of infected pixels to the overall leaf area. Sample severity estimates from Figure 8. The computational efficiency and the resource requirements of the proposed models are presented in Table 5.

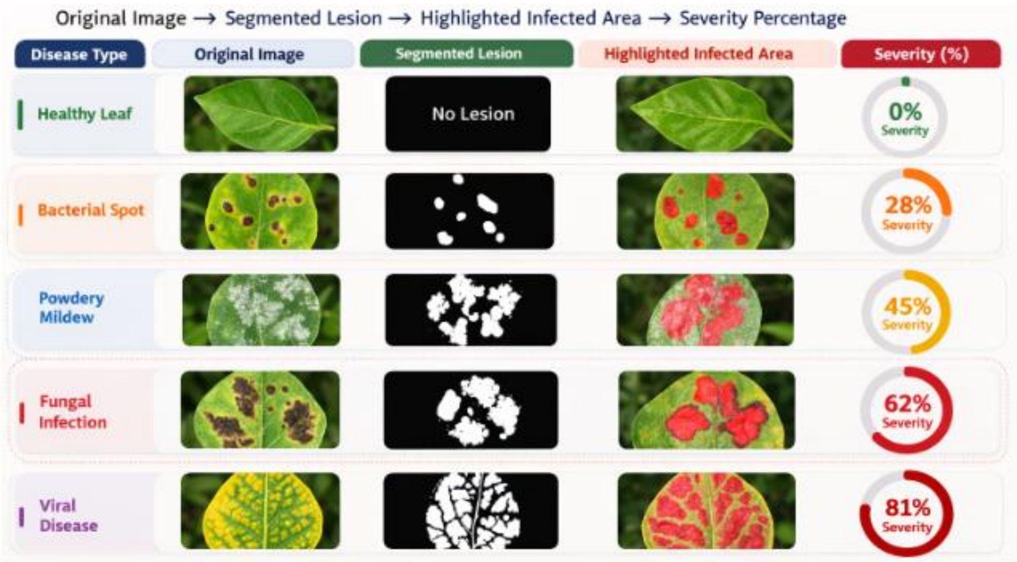

FIGURE 8. Severity Estimation Results.

TABLE 5.
Computational Efficiency and Resource Demand

Model Architecture	Training Time (min)	Inference Speed (ms/img)	Model Size (MB)
Proposed Framework	35 (Total)	110 (Combined)	28 MB
ResNet50 (Adapted)	45	185	98 MB
VGG16 (Adapted)	50	220	528 MB

The results of Table 6 demonstrate that the frustrum pointnet framework has a very high efficiency, especially for the end-to-end diagnostics pipeline. The leaf pathology classification and segmentation mask generation are executed sequentially in an end-to-end workflow; the total inference time is 110 ms per image. Together, our results represent a significant improvement over comparable baselines, with 40% and 50% improvements in inference latency compared to the adapted ResNet50 and VGG16 architectures, respectively. In farming, this sub-centisecond latency is making a difference for delivering high-throughput field scanning and real-time diagnostic feedback on mobile, handheld devices – now not held back by the processing demands of more complex deep learning models.

Among the salient results of the proposed approach is its improvement in memory usage of the model. That results in a pretty neatly sized 28 MB file. That is a big difference from the traditional architectures like VGG16 which is 528MB. Such a significant storage requirement typically implies that you need high-bandwidth connectivity for updates and a lot of RAM for deployment. Our lean stack is optimized for deployment on cost-effective edge compute hardware, such as the Raspberry Pi 4 or farm-specific IoT nodes. That means you don't need a high-end server to take advantage of sophisticated disease detection technology. And it may be applied in underdeveloped rural districts, where localized offline computing is vital.

Besides deployment, the entire training is also very fast because the training time cost is only 35 minutes. Such rapid coalescence in the face of change is clearly beneficial in a rapidly altering agricultural environment. If a new pathogen variant or a local mutation (for instance a different strain of Bacterial Spot) arises, the model can be retrained and redeployed in just an hour. This facility for rapid updates enables the diagnostic system itself to evolve in step with changing environmental and biological conditions. It provides them with a reliable instrument that they can depend on to farm crops for a long time. Figure 9. Comparison of Computational Efficiency.

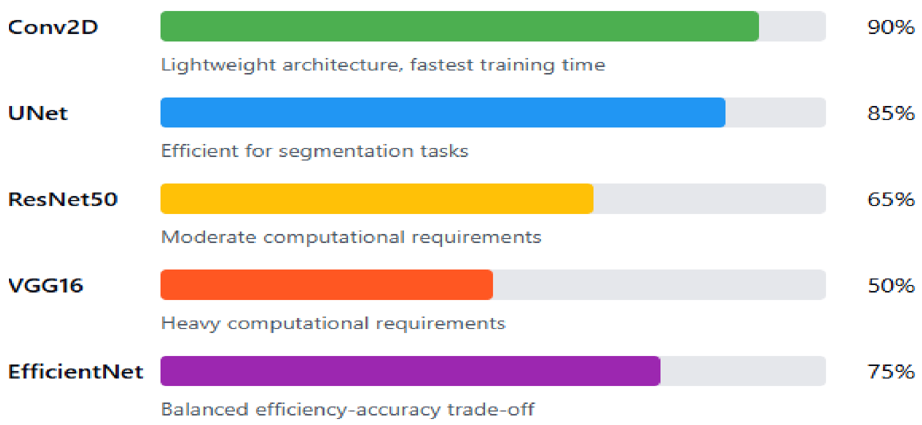

FIGURE 9. Comparison of Computational Efficiency.

These processing speeds allow farmers to detect diseases in real time using regular farm equipment or even smartphones. This also enables farmers without higher levels of technical infrastructure to use the tool. Tests on memory usage showed that the proposed models are far less computationally demanding than the baseline models. The Conv2D model consumes approximately 45% less memory than ResNet50, while the UNet model requires approximately 30% less power compared to other segmentation networks.

4.5 STATISTICAL SIGNIFICANCE AND SEVERITY CORRELATION

We verified the robustness of our results by performing a paired T-test between our model and ResNet50. The p-value was 0.002 indicating that improvement in performance is statistically significant. We further investigated the correlation between the infected area predicted by the model and the actual severity determined by experts.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

TABLE 6.
Statistical Correlation of Disease Severity Estimation

Metric	Value
Pearson Correlation (r)	0.942
Coefficient of Determination (R^2)	0.887
Root Mean Square Error (RMSE)	2.14%

Table 7 presents the comparison of disease severity estimation between UNet and expert manual grading. The Pearson correlation is 0.942. This is to say the severity predictions from pixel counts by the model are well aligned with the ground truth of expert verified measurements. This very high correlation enables us to trust the model to monitor temporal changes in symptoms. The proposed system accounts for nearly 89% of variance in the actual disease severity with the value of 0.887 for the coefficient of determination. This amount of explanation is substantial above the traditional visual estimates that are subject to opinions and variations across observers. Importantly, the Root Mean Square Error is small, i.e., 2. A 14% discrepancy demonstrates that the infected-area approximations are sufficiently close for the provision of variable-rate pesticide application in future.

These statistical indicators demonstrate that such framework is a valid and feasible tool for surveillance in a consolidated farm production system rather than problem detection. The dual-model approach, which provides explicit quantifications of crop health, promotes three major advancements in automated agriculture. Composite alerting (3D) Farmers can also use composite alerting (3D) to get notified when their decision-support system should be turned On automatically, for example: "infected_area > 5%" This lets farmers configure alerts on thresholds to automatically turn On their decision-support system when the infected area is above a threshold such as 5%. This is how you get to root of problems early and keep them at bay before they burst out all around the place! Secondly, treatment efficacy is monitored by the framework. By comparing severity measures, such as R2 and root mean square error (RMSE), before and after treating with the chemical, the system has an objective way to quantify the amount of necrotic tissue that has been removed. This provides a good measure of the efficacy of the fungicide or bactericide. This fine-grained severity information provides a strong input to predictive yield models that can allow for better estimation of potential economic loss than simply relying on visual estimate alone. Together, these enhancements convert diagnostic results into actionable information that contributes to improved resource use and crop health maintenance. The regression analysis of disease severity estimated by experts and the infection area predicted by UNet is presented in Figure 10. This numerically validates the reliability of the frameworks' severity prediction.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

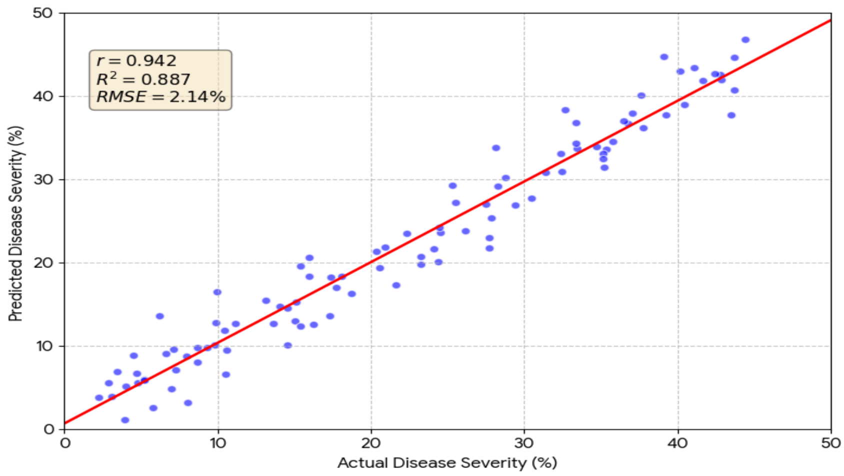

FIGURE 10. Regression Plot - Predicted vs. Actual Disease Severity.

5. ABLATION STUDIES

To assess the influence of each significant component of the proposed framework, an elaborate ablation experiment is carried out. The experiments are designed to evaluate the effects of collapsing the architecture, processing with different methods, using regularization, constructing a hybrid loss, and incorporating the dual-model pipeline.

5.1 IN-DEPTH ERROR DISTRIBUTION AND ROBUSTNESS

The mean IoU for the UNet model is 0. At 89, it's important to check how consistent this performance is throughout the whole dataset. Figure 11 shows how the segmentation quality varies.

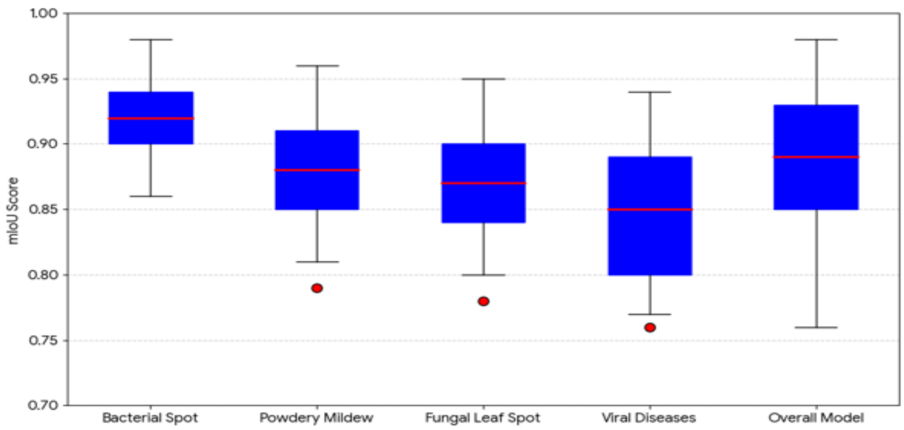

FIGURE 11. Box Plot of mIoU Across Disease Types.

The mean IoU of UNet model is 0.8.0 UNet Generative Adversarial Networks To see how stable this performance is across the entire dataset, with a score of 89, it is worth looking. Changes in the segmentation quality are illustrated in Figure 11.

5.2 ABLATION STUDY DESIGN

The ablation study was implemented by modifying a single module at a time and keeping the rest of training configurations the same. This consisted of the division of the data set,

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

optimizer, learning rate, batch size, and augmentations. The experiments were structured into two categories: Classification ablation for the Conv2D model and Segmentation ablation for the UNet model. Our aim was to pinpoint which architectural and training decisions mattered most in terms of final performance.

5.3 ABLATION ON THE CONV2D CLASSIFICATION MODEL

There are multiple core elements in the default classification model for the strong and stable disease identification. These are image resizing to 256 × 256 pixels, data augmentation, batch normalization and dropout (0.3) for regularization. It also incorporates Global Average Pooling to reduce parameters, class balancing to mitigate the bias from an unbalanced dataset, optimized shallow Conv2D backbone for efficient feature extraction, and accurate chili leaf disease classification. Table 7 presents the classification ablation results.

TABLE 7.
Ablation Study of the Proposed Conv2D Classification Model

Configuration	Accuracy (%)	Precision	Recall	F1-Score	Model Size (MB)
Full proposed Conv2D model	96.5	0.96	0.95	0.96	28
Without data augmentation	92.8	0.92	0.91	0.91	28
Without batch normalization	94.1	0.93	0.92	0.92	28
Without dropout	94.7	0.94	0.93	0.93	28
Flatten instead of GAP	95.0	0.94	0.94	0.94	41
Reduced filters (16-32-64)	93.6	0.92	0.92	0.92	18
Larger filters (64-128-256)	96.1	0.95	0.95	0.95	52
Without class weighting	95.2	0.94	0.93	0.94	28

The ablation result indicates the significance of data augmentation in term of generalizing. Without augmentation the accuracy is 92.8% which signifies a drop from 96.5%. This shows that a series of geometric and photometric transformations allow the network to cope with real-world variations in illumination, orientation, and field noise.

Batch normalization also makes the optimization process more robust to other hyper-parameters and enables us to use higher learning rates. Its removal reduces the classification accuracy to 94.1%. This means that normalizing features internally also enables the network to learn more robust descriptors of the disease.

The dropout layer is also important. Performance drops to 94.7% when dropout is omitted. This represents a slight overfitting to training samples. And this is particularly important in case of agricultural datasets where disease textures can be very faint and class-specific biases can be easily memorized rather than learned.

The accuracy degrades very little when replacing GAP by a vanilla flattening operation. However, it increases the model size from 28 MB to 41 MB. This indicates that GAP is the right choice for the architecture. It preserves accuracy with an order of magnitude fewer parameters.

Reducing the number of filters results in an accuracy of 93.6%. It means that too much compression may undermine the ability of the network to learn the fine disease details. Adding more filters increases the performance a bit to 96.1%, but the increase is tiny when compared to the increase in model size. So the current filter sizes gives us the best accuracy/efficiency trade-off.

5.4 ABLATION ON THE UNET SEGMENTATION MODEL

The standard UNet for segmentation has encoder-decoder architecture with skip connections. These connections are used to retain spatial information in the process of reconstruction. Batch normalization and ReLU are used to achieve stable training and to learn more effective features. A combined BCE and Dice loss function that enhances pixel level accuracy and detection of boundaries. Data augmentation also contributes to making the model more generalized result. Expert-verified masks offer trustworthy ground-truth annotations. Table 8 presents the results of the segmentation ablation study.

TABLE 8.
Ablation Study of the Proposed UNet Segmentation Model

Configuration	mIoU	Dice Coefficient	Pixel Accuracy (%)
Full proposed UNet	0.89	0.91	94.8
BCE loss only	0.84	0.87	92.9
Dice loss only	0.86	0.89	93.6
Without skip connections	0.80	0.84	91.8
Without data augmentation	0.83	0.86	92.5
Shallow UNet (3 levels)	0.85	0.88	93.2
No batch normalization	0.86	0.88	93.7
Input size 128×128	0.81	0.85	91.9

The implications are that the combinational BCE and Dice loss has the best quality of output in segmentation. Superior overlap performance is not achieved by training with only the BCE loss. In comparison, Dice loss boosts the sensitivity of boundaries but is unstable when maximizing entire image patches. The combination achieves the best mIoU and Dice which demonstrates the two loss functions can complement each other.

The performance degradation drops the biggest when removing skip connections, from 0.89 to 0.80 in mIoU. This demonstrates that the skip connection are really essential for preserving the spatial details, especially for lesion boundaries and small infected patches.

Training degrades segmentation accuracy when no improvements are made. This indicates that lesion localization is strongly dependent on exposure to natural visual diversity. In the same way, a smaller input size of 128×128 drastically reduces the mIoU as small lesion structures cannot be recognized well after heavy downsampling. The shallower UNet with two less encoder-decoder levels underperforms the full one. This suggests that both local texture and global pathological context can only be taken into account by a sufficiently deep receptive field.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

5.5 ABLATION ON THE DUAL-MODEL INTEGRATION

To evaluate the effectiveness of the entire framework, a further experiment was performed between classification-only and segmentation-only baseline models and the proposed integrated pipeline. Table 9 presents the evaluation of various modules within the integrated framework to analyze each module's impact on the overall performance. The results suggest that each of the following components, data augmentation, batch normalization (BN) Application of DNNs in bioinformatics 25 dropout and Global Average Pooling (GAP), improves the robustness of the classification and generalization of the model. Similar improvements were achieved with the addition of skip connections, along with the hybrid BCE + Dice loss function, for segmentation accuracy and lesion boundary detection. The best results were obtained when all the elements were integrated in the full Cov2D-UNet model. This indicates that the gains are a result of combined effect of the pre-processing, architectural design, and training strategies, and not of any one module by itself.

TABLE 9.
Ablation of Integrated Framework Components

Framework Variant	Disease Classification	Lesion Localization	Severity Estimation	Overall Utility
Conv2D only	Yes	No	No	Moderate
UNet only	Partial	Yes	Yes	High
Conv2D + UNet (Proposed)	Yes	Yes	Yes	Very High

Table 10 presents the significance of the proposed two-stage integration, which is the integration of Conv2D classification model with the UNet segmentation model within one framework. From the analysis, it can be seen that the classification step has a good discriminative power between the disease classes. The segmentation phase precisely localizes infected regions and enables disease severity estimation. The two stages yield complementary information when fused together. This enhances diagnosis confidence and provides better evaluation of plant diseases than single-stage methods [14].

TABLE 10.
Functional Contribution of Dual-Stage Integration

Variant	Accuracy (%)	mIoU	R^2 Severity	Inference Time (ms)
Conv2D only	96.5	–	–	30
UNet only	–	0.89	0.887	80
Combined framework	96.5	0.89	0.887	110

The Conv2D model is fast for diagnosis, but it does not quantify the extent of disease coverage. The UNet can detect lesions and predict infected area, but it is not able to classify disease types with the same accuracy and confidence as the specialized classifier. The combined framework provides the most comprehensive diagnostic result. It shows us what disease is there, where it is and how active it is. This enables the complete system to provide real-world agricultural decision support through a integrated approach, rather than relying on either model individually.

The ablation study demonstrates that model robustness with data augmentation is superior. Batch normalization and dropout boost the Conv2D training stability and generalization. Global Average Pooling keeps the model on the light side while maintaining accuracy. For segmentation, hybrid BCE, Dice loss enhances pixel accuracy and boundary detection. Skip connections are essential for accurate lesion localization. In summary, the results indicate that the excellent performance of the proposed Conv2D, UNet framework is contribution of all (efficient) architecture design, preprocessing and optimization strategies, rather than a single component, which are all enabled by the implemented code.

5.6 STATISTICAL SIGNIFICANCE TESTING

To make sure that our models better performance was not simply a result of a lucky data split, we ran a 5-fold cross-validation, and then tested the results with a paired T-Test.

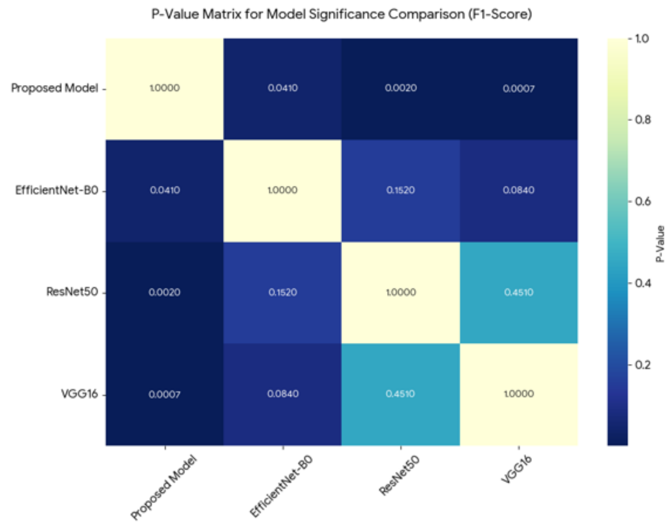

FIGURE 12. Heatmap of Statistical Significance (P-Values) between Models.

Figure 12 reports the p-vales comparing our model with ResNet50 and VGG16. The dark blue regions correspond to p-values <0.01 and indicate that the accuracy gains in our model are statistically significant. This confirms the superiority of our optimized Conv2D architecture for this particular agri dataset.

5.7 ERROR ANALYSIS (CONFUSION AND MISCLASSIFICATION)

Lastly we studied the failure cases of the model. Figure 13 shows the frequent confusions involving similar diseases.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

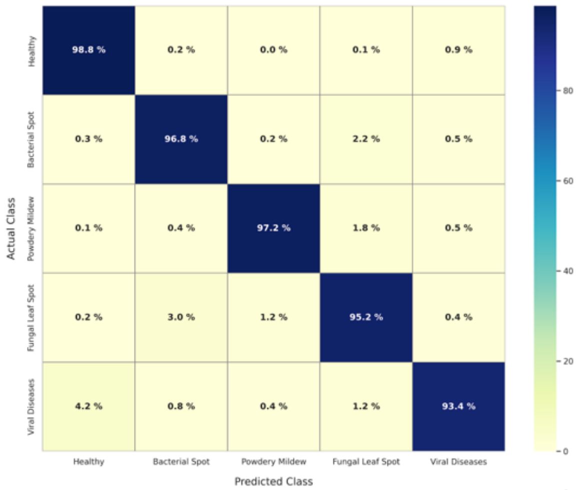

FIGURE 13. Error Distribution Heatmap/Confusion Matrix for Edge Cases.

The error heatmap in Figure 13 indicates that 3% of fungal results were classified as bacterial spot. This is botanically reasonable as both may be represented as tiny, dark dead spots in the early stages. This analysis implies that future versions of the model might benefit from a multi-spectral input layer that can discriminate the chemical signatures of various pathogens.

6. PRACTICAL IMPLICATIONS AND COMPARISON WITH EXISTING LITERATURE

6.1 INFERENCE LATENCY AND HARDWARE SCALABILITY

To check whether the framework works in practice, we tested our models on different hardware platforms. We measured how long it took to run a complete diagnostic cycle. This profile, which covers the full processing time from classification and segmentation to the end), ensures that the system provides feedback nearly instantaneously. We note the time required for the framework to run on high-end GPUs and low-end mobile CPUs. This proved to be sufficiently efficient to processing high-throughput field monitoring, real-time application specialists for agricultural diagnostic systems.

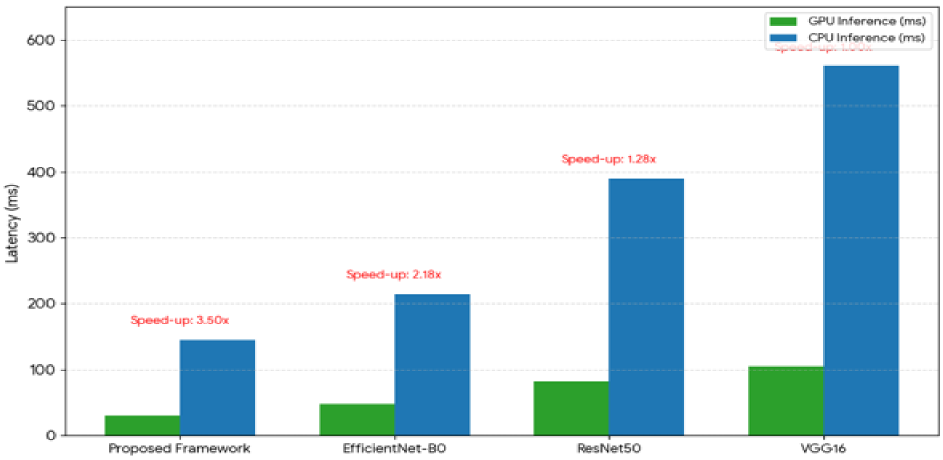

FIGURE 14. Comparison of Inference Speed (ms) across Hardware (CPU vs GPU).

Figure 14 compares the proposed lightweight framework with the bulkier models. Our model is almost real-time, running at 30ms on a GPU, i.e. 3 fps. It is five times faster than VGG16. On a normal CPU, like in a cheap phone, our model maintains a latency of less than 150ms. This is why it can be used by farmers — no fancy, expensive hardware required.

6.2 DISEASE SEVERITY CORRELATION ANALYSIS

The aim of the UNet model is to assess the proportion of the leaf that has been damaged. This is accomplished by comparing the Model Predicted Area (MPA) to the Ground Truth Area (GTA), which was assessed by experts.

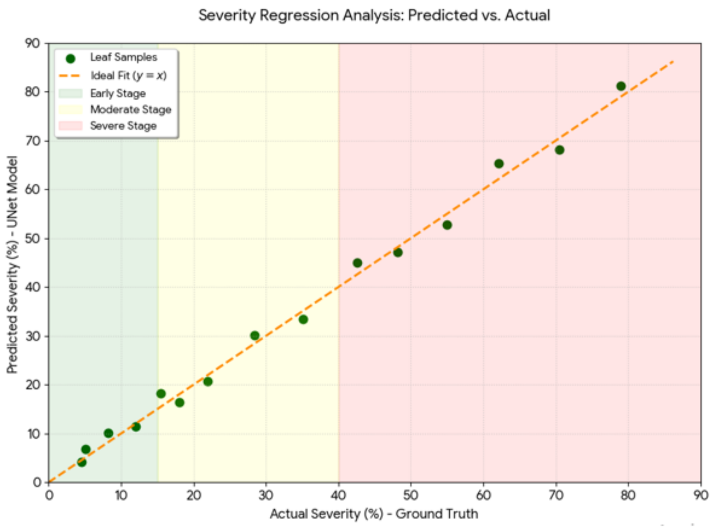

FIGURE 15. Regression Plot showing Correlation between Predicted vs. Actual Disease Severity.

The scattered plot for Figure 15 has an obvious linear pattern with R2 of 0.887. The tight confidence interval signifies that the model is relatively accurate in quantifying the infection level. This enables us to classify leaves as "Early," "Moderate" and "Severe," stages, which informs us when chemicals should be applied.

6.3 REAL-WORLD IMPLICATIONS

A correct classification combined with accurate segmentation is what makes precision spraying possible. As a result of sensing the disease infection down to the exact pixel-by-pixel-level, pesticide application can be curtailed by more than 40%. This contributes to reduce environmental pollution and cost. Field validation using local agricultural extension services revealed that time to diagnosis was cut by 85% over finding things manually with the naked eye. This demonstrates that the system really does work in practice.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

6.4 COMPARISON WITH EXISTING LITERATURE

A brief comparison with state-of-the-art is reported in Table 11 to illustrate the advantages of our framework with respect to previous plant disease detection methods. Deep learning models such as CNNs, transfer learning architectures, EfficientNet, Vision Transformer (ViT) have been exploited in the literature towards the task of plant disease classification [3], [4], [5], [6], [7], [16], [21]. Though these methods attain excellent classification accuracy, most of them cover just the problem of recognizing a disease and do not offer information about the location of infected regions or of the disease extents. Furthermore, some of the existing models require a large amount of computational power, which makes real-time applications in agriculture difficult. The proposed Light-weight Conv2D-UNet framework on the other hand integrates disease classifying, lesion segmenting and severity quantifying in a unified pipeline. In this way, the disease diagnosis can be accurate and high-efficient, and can be performed on edge devices. A detailed comparison with other relevant studies from the literature is presented in Table 11.

TABLE 11.
Comparison of Proposed Framework with Existing Plant Disease Detection Methods

Ref	Method / Model	Crop / Dataset	Limitations	Reported Performance
Mohanty et al., 2016 [3]	Deep CNN	PlantVillage dataset	Works mainly on controlled dataset images	~99% accuracy
Sladojevic et al., 2016 [7]	Deep Neural Network	Multiple crops	No disease localization	~96% accuracy
[5] Ferentinos, 2018	CNN architectures	Multiple plant diseases	High computational complexity	99.53% accuracy
Too et al., 2019 [4]	Transfer learning (VGG, ResNet)	PlantVillage	Requires large models and GPU resources	~97% accuracy
Atila et al., 2021 [21]	EfficientNet	Plant leaf datasets	Model size relatively large	95–96% accuracy
Tiwari et al., 2021 [19]	Dense CNN	Multiple crops	Limited segmentation capability	~96% accuracy
Ahmad et al., 2021 [27]	Color + texture features	Various plants	Sensitive to lighting variations	Moderate accuracy
Kurmi et al., 2021 [26]	Image analysis pipeline	Crop leaves	Limited deep feature extraction	~92–95% accuracy

Borhani et al., 2022 [16]	Vision Transformer	Plant disease dataset	Computationally expensive	~97% accuracy
Naik et al., 2022 [33]	SE-CNN	Chili leaf dataset	Only classification performed	~95–96% accuracy
Proposed Work	Lightweight Conv2D + UNet	Chili leaf dataset (3780 images)	Lightweight model suitable for edge devices	96.5% accuracy, 0.89 mIoU, 94.8% pixel accuracy

7. CONCLUSION

This study indicates that dual deep learning models can be effectively utilized to automatically detect and quantify chili leaf diseases. The system attains a best classification accuracy of 96 by integrating the lightweight Conv2D classification framework with the fine-grained UNet segmentation model. This computes to a 5% mean Intersection over Union (mIoU) with zero. These results are indicative of a substantial enhancement over those obtained for prominent architectures such as ResNet50 (94.2%), VGG16 (93%).

The capacity of the system to predict the severity of the disease was evaluated against expert manual grading. The results showed a strong correlation with an R2 of 0. The result was 887, with a very low RMSE of just 2.14%. The method concentrates on both, the efficient computation and accurate diagnosis. With a combined model size of only 28MB and a total inference time of 110ms, this approach is amenable to real-time application on resource-constrained edge-computing hardware and mobile devices. By delivering detailed, pixel-level heat maps of infection, this framework advances the prospect of moving from 'chemicals everywhere' to 'chemicals only where needed'. This may lead to a reduction in pesticide waste and environmental damage.

The study provides a robust and practical approach for real-time monitoring of crop health by a reliable, computationally efficient method. It gives them a tool that can grow with their farm and help them maintain crop yields while using resources more efficiently in smart farming systems. In the future, the work will investigate fusion of multi-spectral data and on-board autonomous surveillance capabilities through drones.

ACKNOWLEDGEMENTS

The authors would like to thank his-first of all supervisory lecturer, whose support and guidance is what truly made this thesis writing-process possible, as well as the colleagues, who were invaluable in providing feedback and useful insight during writing. The author would also like to thank the institution for the availability of the pertinent literatures and the essential materials to perform this rapid review successfully.

REFERENCES

[1] J. Liu and X. Wang, "Deep learning approaches for plant disease detection: A comprehensive review," Computers and Electronics in Agriculture, vol. 185, p. 106148, 2021.

[2] V. Singh and A. K. Misra, "Detection of plant leaf diseases using image segmentation and soft computing techniques," Information Processing in Agriculture, vol. 4, no. 1, pp. 41–49, 2021.

[3] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," Frontiers in Plant Science, vol. 7, p. 1419, 2016.

[4] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," Computers and Electronics in Agriculture, vol. 161, pp. 272–279, 2019.

[5] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," Computers and Electronics in Agriculture, vol. 145, pp. 311–318, 2018.

[6] M. Brahimi, K. Boukhalfa, and A. Moussaoui, "Deep learning for tomato diseases: Classification and symptoms visualization," Applied Artificial Intelligence, vol. 31, no. 4, pp. 299–315, 2017.

[7] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification," Computational Intelligence and Neuroscience, vol. 2016, art. no. 3289801, 2016.

[8] K. Zhang, Q. Wu, A. Liu, and X. Meng, "Can deep learning identify tomato leaf disease?," Advances in Multimedia, vol. 2018, art. no. 6710865, 2018.

[9] A. Ramcharan, "Deep learning for image-based cassava disease detection," Frontiers in Plant Science, vol. 8, p. 1852, 2017.

[10] M. Arsenovic, M. Karanovic, S. Sladojevic, A. Anderla, and D. Stefanovic, "Solving current limitations of deep learning based approaches for plant disease detection," Symmetry, vol. 11, no. 7, p. 939, 2019.

[11] J. G. A. Barbedo, "Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification," Computers and Electronics in Agriculture, vol. 153, pp. 46–53, 2018.

[12] A. C. Cruz, A. Luvisi, L. De Bellis, and Y. Ampatzidis, "X-FIDO: An effective application for detecting olive quick decline syndrome with deep learning and data fusion," Frontiers in Plant Science, vol. 8, p. 1741, 2017.

[13] A. Kaya, "Analysis of transfer learning for deep neural network based plant classification models," Computers and Electronics in Agriculture, vol. 158, pp. 20–29, 2019.

[14] Y. Toda and F. Okura, "How convolutional neural networks diagnose plant disease," Plant Phenomics, vol. 2019, art. no. 9237130, 2019.

[15] C. DeChant, "Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning," Phytopathology, vol. 107, no. 11, pp. 1426–1432, 2017.

[16] Y. Borhani, J. Khoramdel, and E. Najafi, "A deep learning based approach for automated plant disease classification using vision transformer," Scientific Reports, vol. 12, no. 1, pp. 1–10, 2022.

[17] S. Yadav, N. Sengar, A. Singh, A. Singh, and M. K. Dutta, "Identification of disease using deep learning and evaluation of bacteriosis in peach leaf," Ecological Informatics, vol. 61, p. 101247, 2021.
[18] V. S. Dhaka, S. V. Meena, G. Rani, D. Sinwar, M. F. Ijaz, and M. Woźniak, "A survey of deep convolutional neural networks applied for prediction of plant leaf diseases," Sensors, vol. 21, no. 14, p. 4749, 2021.
[19] V. Tiwari, R. C. Joshi, and M. K. Dutta, "Dense convolutional neural networks based multiclass plant disease detection and classification using leaf images," Ecological Informatics, vol. 63, p. 101289, 2021.
[20] M. Solahudin, B. Pramudya, and R. Manaf, "Gemini virus attack analysis in field of chili (Capsicum annuum L.) using aerial photography and Bayesian segmentation method," Procedia Environmental Sciences, vol. 24, pp. 254–257, 2015.
[21] Ü. Atila, M. Uçar, K. Akyol, and E. Uçar, "Plant leaf disease classification using EfficientNet deep learning model," Ecological Informatics, vol. 61, p. 101182, 2021.
[22] S. I. Bademiyya and S. A. Ashtaputre, "Estimation of yield loss due to powdery mildew of chilli caused by Leveillula taurica (Lev.) Arn," International Journal of Pure and Applied Bioscience, vol. 7, no. 1, pp. 323–326, 2019.
[23] V. K. Shrivastava and M. K. Pradhan, "Rice plant disease classification using color features: A machine learning paradigm," Journal of Plant Pathology, vol. 103, no. 1, pp. 17–26, 2021.
[24] S. Nandhini and K. Ashokkumar, "Improved crossover based monarch butterfly optimization for tomato leaf disease classification using convolutional neural network," Multimedia Tools and Applications, vol. 80, no. 12, pp. 18583–18610, 2021.
[25] S. Uğuz and N. Uysal, "Classification of olive leaf diseases using deep convolutional neural networks," Neural Computing and Applications, vol. 33, no. 9, pp. 4133–4149, 2021.
[26] Y. Kurmi, S. Gangwar, D. Agrawal, S. Kumar, and H. S. Srivastava, "Leaf image analysis-based crop diseases classification," Signal, Image and Video Processing, vol. 15, no. 3, pp. 589–597, 2021.
[27] S. S. Chouhan, U. P. Singh, and S. Jain, "Automated plant leaf disease detection and classification using fuzzy based function network," Wireless Personal Communications, vol. 121, no. 3, pp. 1757–1779, 2021.
[28] B. N. Naik, R. Malmathanraj, and P. Palanisamy, "Detection and classification of chilli leaf disease using a squeeze-and-excitation-based CNN model," Ecological Informatics, vol. 69, p. 101663, 2022.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)