# Tomato Pest and Disease Identification Based on Improved Deep Residual Network and Transfer Learning

**[1]Peng Linli, [2]Tjong Wan Sen, [3]Hasanul Fahmi, [4*]Rusdianto Roestam**
[1,2,3,4]Department of Informatics, Faculty of Computing, President University, Indonesia
Email: [1]penglinli@gmail.com, [2]wansen@president.ac.id, [3]hasanul.fahmi@president.ac.id, [4]rusdianto@president.ac.id

**Article Info**
***Article history:***
Received Nov 10th, 2024
Revised Dec 24th, 2024
Accepted Jan 29th, 2025

***Keyword:***
Attention Mechanism
DropBlock Regularization
Pests and Diseases
PlantVillage
ResNet50

**ABSTRACT**
Tomatoes are a vital global crop, but their yield can be severely impacted by various diseases like leaf mold and spotted wilt. Early and accurate diagnosis of these diseases is crucial for implementing timely treatments, thereby reducing crop loss. Traditional manual diagnosis often suffers from low accuracy, high costs, and time consumption. To address these issues, this study introduces a method for identifying tomato pests and diseases using an improved residual network and transfer learning. A dataset comprising images of seven common tomato diseases and healthy leaves was created. This study introduces an improved residual network and transfer learning method to accurately identify tomato pests and diseases. The enhanced ResNet50 model, with an attention mechanism and focal loss, achieved 98.10% recognition accuracy. This research not only facilitates early disease detection, reducing crop loss but also minimizes pesticide use, thereby enhancing environmental sustainability and agricultural productivity worldwide.



***Corresponding Author:***
Rusdianto Roestam,
Department of Informatics, Faculty of Computing,
President University,
Jababeka Industrial Estate, Cikarang, Bekasi, West Java, Indonesia.
Email: rusdianto@president.ac.id

**DOI: http://dx.doi.org/10.24014/ijaidm.v8i1.34038**

## 1. INTRODUCTION

### 1.1 Background and Significance of the Research

Tomato, an important cash crop, is widely grown worldwide and loved by consumers for its rich nutritional value. It also holds a significant position in agricultural production due to its economic value [1]. However, tomatoes often face various pests and diseases during cultivation, severely limiting yield and quality, causing substantial economic losses [2].

The traditional approach to treating tomato pests and diseases is large-scale pesticide spraying to control outbreaks and prevent their spread, ensuring that tomato yields do not continue to decline [3]. However, excessive pesticide use can lead to serious environmental pollution and increased pesticide residues in tomatoes. If diseases can be accurately identified in the early stages of pest outbreaks and treated with targeted methods, tomato yield reduction can be effectively controlled while minimizing environmental pollution and pesticide residues. Traditional tomato pest control relies on the expertise of agricultural plant protection technicians or experienced tomato growers [4].

Thanks to the breakthrough development of computer technology such as machine vision, the task of rapid and accurate diagnosis of tomato pests and diseases has been supported by theoretical technology [5]. After tomato plants have been subjected to certain pests and diseases, the color, shape and other texture characteristics of the corresponding parts of the leaf surface will be visible to the naked eye, different from the healthy state of the changes in this texture state is the most intuitive diagnosis of whether the tomato plant has a disease and what disease it has [6]. Digital image recognition under deep learning is a technology that can quickly extract image features, this technology through a large amount of data for training and learning, you can learn how to extract the relevant features of the image, and save the parameters of the extracted features for future continued use [7].

This paper focuses on the diagnosis and recognition of common tomato pests and diseases using deep residual networks and transfer learning techniques. The practical significance of this research includes:Early and accurate diagnosis of tomato pests and diseases, allowing for timely preventive measures to improve tomato fruiting rate and increase production. Targeted treatment of pests and diseases, reducing pesticide residues and promoting environmental protection. Deep learning technology does not rely on skilled plant protection personnel, making it widely accessible and applicable in most regions of the country.

### 1.2 Literature review

Deep learning, which has evolved over several decades from artificial neural network research, is a derivative of traditional machine learning with significant prospects in various fields, including biomedical research and stock analysis [8]. It encompasses digital image processing technologies that perform tasks such as object classification and gesture tracking by simulating the visual neural networks of highly intelligent organisms. This enables the extraction of low-dimensional features from images to form higher-level representations [9].

Cheng et al. (2017) developed a residual neural network model for pest detection in complex backgrounds [10]. This study was pioneering in enhancing pest detection accuracy; however, it did not specifically address disease identification and was limited to a narrow dataset. She et al. (2020) improved upon this with an optimized SSD model for identifying pests on rice leaves [11]. While effective, the model's generalizability to other crops, such as tomatoes, was not explored. Zhang et al. (2018) utilized transfer learning to classify diseased cotton leaves with an accuracy of 89.51% [12]. The study demonstrated the potential of transfer learning but suffered from a lack of diversity in the types of diseases examined. Ferentinos et al. (2018) achieved 99.53% accuracy in detecting plant leaves using VGG16 [13]. This study showcased high accuracy but was constrained by the computational intensity of the VGG16 model. Lastly, Jenifa et al. (2019) classified cotton leaf diseases with a 96% accuracy using a deep convolutional neural network [15]. Though impressive, the study did not address the issue of real-time detection, which is critical for in-field applications.

Our research distinguishes itself by focusing on an improved deep residual network, ResNet50, specifically tailored for tomato pests and diseases. Unlike previous studies, our approach integrates transfer learning from a pre-trained PlantVillage dataset with an enhanced model that includes focal loss and attention mechanisms. This combination allows for not only superior classification performance but also computational efficiency. Our model achieves an average recognition accuracy of 98.10%, which is a significant improvement over previous models. Additionally, we address the challenge of real-time detection by optimizing our model for deployment on less powerful computational platforms, making it suitable for practical, in-field use. This research, therefore, presents a comprehensive solution that bridges the gap between high accuracy and practical application in agricultural settings, setting it apart from previous studies that focused solely on laboratory settings or specific aspects of pest and disease detection.

## 2. RESEARCH METHOD

### 2.1 Theoretical Basis

#### 2.1.1 Deep Learning

Deep learning, first conceptualized in the mid-20th century [22-25], faced initial setbacks due to limited data and inadequate computing power [26-28]. It surged in the 21st century with increased data access and advanced computing, leading to applications in healthcare [29], finance [30], security [31], and criminal investigation [32]. Deep learning is widely used across various domains [33-36].

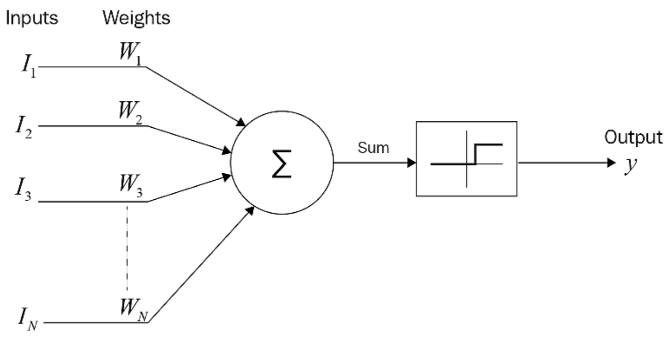



**Figure 1.** McCulloch-Pitts NeuronModel structure

The McCulloch-Pitts Neuron model [37] (Figure 1), designed to mimic brain signal processing, uses multiple inputs and a single output. It handles input through linear weighting and threshold functions, producing binary outcomes. Despite its utility in binary classification tasks like spam detection, its manual weight settings and low accuracy in complex tasks reveal its limitations [38-39].

### 2.1.2 Neural network

Traditional neural networks, often shallow and fully connected, use neurons as basic units [38-41]. These neurons, resembling McCulloch's model, multiply inputs by weights and sum them. Network Architecture can view Figure 2.

A fully connected network comprises many interconnected neurons [42], supporting multiple inputs and outputs. The network's structure is determined by neuron connections across layers. Forward propagation allows layer-based computation, using outputs, weights, and biases from previous layers to calculate inputs for the next [43].

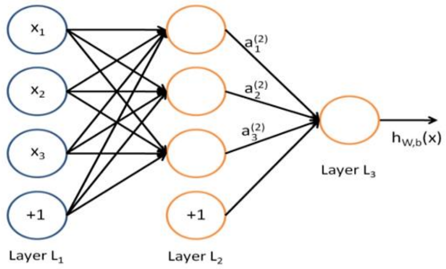



**Figure 2.** Network architecture

### 2.1.3 Convolutional neural network

Classifying a 256x256x1 image with a simple three-layer fully connected neural network requires 65,536 neuron nodes, with the second layer containing 512 nodes resulting in approximately 33,554,944 parameters [44-47]. Due to limited feature extraction capabilities and risk of overfitting, such networks are ineffective for image recognition tasks.

Convolutional neural networks (CNNs) address these issues by using fewer parameters and providing enhanced performance in image classification [48-50]. CNNs only connect neurons to a local region of the previous layer (typically 3x3 or 5x5), making the network more efficient. CNNs include a convolutional base for initial processing and a classifier with fully connected layers. The convolutional layer, crucial for extracting key features, uses kernels to process input matrices and generate feature maps, optimizing image classification tasks, Convolutional Neural Network Structure Diagram can view Figure 3.

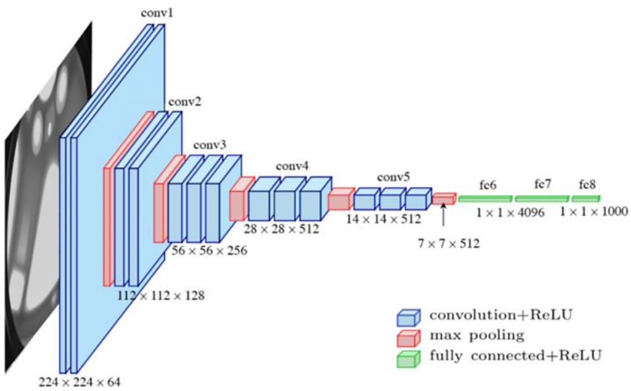



**Figure 3.** Convolutional Neural Network structure diagram

### 2.1.4 Classical convolutional neural network

LeNet-5 resizes raw images to fit the input layer, extracts local features using convolutional layers with activation functions (e.g., ReLU), and downsamples outputs using pooling layers to reduce spatial dimensions and complexity [51]. Fully connected layers combine features for classification or regression tasks. LeNet-5 is known for its simplicity and is widely used for handwritten digit recognition.

AlexNet, similar to LeNet-5, uses multiple convolutional layers with ReLU activation and pooling layers to learn complex features and reduce spatial dimensionality [52]. Fully connected layers map features to output categories. AlexNet introduces Dropout to combat overfitting and significantly advanced deep learning in computer vision by winning the 2012 ImageNet Challenge [53].

VGG stacks multiple 3x3 convolutional kernels to extract features and enhance network depth while maintaining a constant sensory field [54]. It uses 2x2 max pooling layers after each convolutional layer and fully connected layers for classification or regression. VGG's performance is attributed to its increased depth, embodying the "depth" concept in deep learning.

GoogLeNet introduces the Inception module, which uses parallel convolutional kernels of various sizes (1x1, 3x3, 5x5) and pooling operations, combining their outputs for subsequent layers. This design improves adaptability to multi-scale features and increases network width and depth. GoogLeNet includes auxiliary classifiers to address gradient vanishing and accelerate training [55-56]. It replaces fully connected layers with a global average pooling layer, reducing parameters and enhancing model generalization.

### 2.1.5 The flow chart of the research methodology

The flowchart presents the development of a tomato pest and disease identification model. It starts with research, including field studies, information review, and analyzing current research status. After a feasibility study, the project proceeds with dataset construction, involving image acquisition and pre-processing, can view detail in Figure 4.

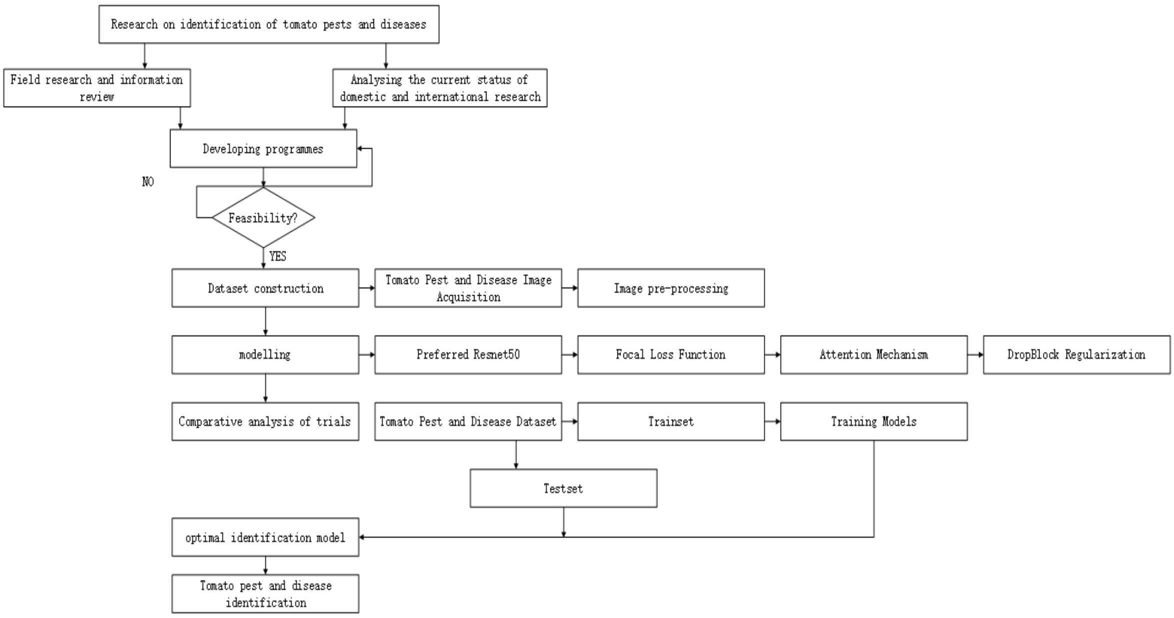



**Figure 4.** Research methodology flowchart

The modeling phase uses a Preferred ResNet50 architecture, enhanced with Focal Loss Function, Attention Mechanism, and DropBlock Regularization to address class imbalance, improve feature recognition, and prevent overfitting.

The model undergoes comparative analysis using the Tomato Pest and Disease Dataset, followed by training and testing. Finally, the model is optimized, resulting in the Tomato Pest and Disease Identification Program, designed to accurately identify and classify tomato pests and diseases.

## 2.2 Dataset

### 2.2.1 PlantVillage Dataset

This study uses the PlantVillage dataset for transfer learning. It is a specialized database of plant disease images, which is widely used in crop disease and plant disease research. By providing a large number of labeled plant disease images, it supports the application of deep learning, machine learning, and other techniques in agriculture, especially in smart diagnosis, crop health management, and pest prediction. The dataset is maintained by Spandan Mohanty et al. and is continuously updated to include more plant species and diseases. The dataset contains nearly 50,000 high-resolution color images covering a wide range of crops (e.g., tomatoes, apples, bananas, etc.) and the various pests and diseases they can suffer from. DiagramMapping of the PlantVillage Datase can view Figure 5.

**Figure 5.** Diagram mapping of the plantvillage datase

#### 2.2.2 Self-constructed Image Dataset of Tomato Pests and Diseases

From June to August 2023, images of tomato pests and diseases were collected at a tomato planting site in Wenzhou, China, using a Nikon Z9 camera (2688 × 1792 pixels). Data collection occurred twice a week at three time slots: 10:30-11:30 a.m., 3:00-4:00 p.m., and 8:30-9:30 p.m., to capture images under varying light conditions. A total of 4,325 images were collected, categorized as follows: Normal Leaf: 552, Bacterial: 568, Early Blight: 505, Late Blight: 520, Leaf Mold: 502, Mosaic Virus: 532, Septoria Spot: 555, Yellow Virus: 591. Examples of tomato pest and disease images can view Figure 6.

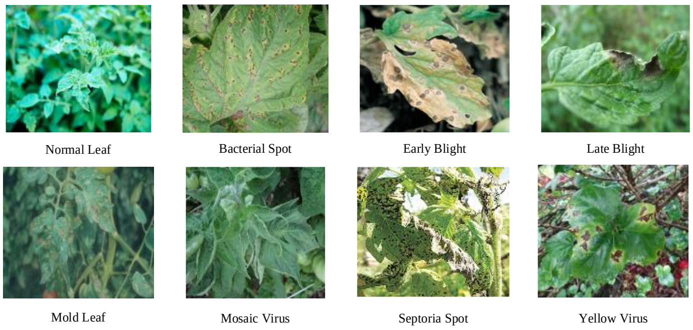



**Figure 6.** Examples of tomato pest and disease images

To ensure fair model evaluation, the dataset was divided into training and test sets in an 8:2 ratio. To address variability from weather conditions, shooting angles, and equipment noise, the training and test sets were enhanced. Techniques employed included Gaussian blur, brightness, sharpness, contrast, and saturation adjustments to simulate weather variations, along with rotation, offset, scaling, shear, and perspective transformations for angle adjustments. Gaussian and pretzel noise were added to mimic equipment noise. The resulting number of enhanced images for each pest and disease category is summarized in Table 1 For model training, images were resized to 224 × 224 pixels using bilinear interpolation, and normalization was applied to improve convergence speed and stability.

**Table 1.** Overview of the retrieval task datasets

| Categories | Original image | Augmented image | |
|---|---|---|---|
| | | Training set | Test set |
| Normal Leaf | 552 | 5740 | 1436 |
| Bacterial Spot | 568 | 5907 | 1477 |
| Early Blight | 505 | 5252 | 1313 |
| Late Blight | 520 | 5408 | 1352 |
| Mold Leaf | 502 | 5220 | 1306 |
| Mosaic Virus | 532 | 5532 | 1384 |
| Septoria Spot | 555 | 5775 | 1443 |
| Yellow Virus | 591 | 5522 | 1381 |

### 2.3 Model Establishment

#### 2.3.1 ResNet Model

Extracting high-level features from images using deep networks can enhance classification accuracy, but increased network depth may lead to issues like gradient disappearance, degradation, and accuracy saturation. The ResNet50 model addresses network degradation by introducing residual modules, allowing the model to compute residuals layer by layer. Instead of directly matching each layer to a given mapping, ResNet trains layers to match a residual mapping, enabling error propagation via shortcut connections. This design facilitates faster convergence and deeper training for improved accuracy.

In this study, the ResNet50 model serves as the backbone network, structured into six main stages. Stage 0 consists of three layers: a convolutional layer with 7×7 kernels and 64 filters, followed by a normalization layer and a ReLU activation function, and concludes with a 3×3 max pooling layer (stride 2) for downsampling, primarily for preprocessing and feature extraction. Stages 1 to 4 consist of residual blocks that extract high-level features, while stage 5 utilizes these features for classification results via global average pooling and fully connected layers, can view Figure 7.

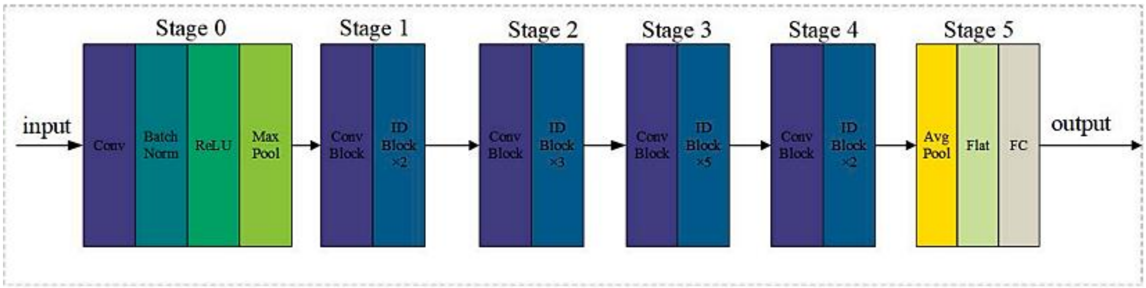



**Figure 7.** The overall structure of ResNet50 model

### 2.3.2 Loss Function

The focus loss function is a solution to the problems of unbalanced data samples and mining difficult samples in target detection. The core idea is to make the model pay more attention to hard-to-classify samples and small number of samples, and solve the problem of data sample imbalance by increasing the weight of small number of samples and hard-to-classify samples and decreasing the weight of easy-to-classify samples. The expression for the focus loss function is equation 1.

$$L_{FL} = -\sum_{i=1}^{C} \alpha y_i (1-\hat{y}_i)^{\gamma} \lg \hat{y}_i \tag{1}$$

where: $y_i$ represents the actual category. $\hat{y}_i$ represents the predicted value of the category. $\alpha$ is a positive adjustable hyperparameter used to balance the effect of positive and negative samples on the loss value. $\gamma$ (also a positive adjustable hyperparameter) is used as a focus parameter to adjust the rate of weight reduction of the samples that are easier to identify and categorize, and when it is 0, the focus loss function is equivalent to the standard cross-entropy loss function. $(1-\hat{y}_i)^{\gamma}$ is the modulation factor. The focus loss function reduces the weight of more recognizable images by a modulation factor.

### 2.3.3 Attention Mechanism

The attention mechanism is a biomimetic vision approach known for its computational efficiency and strong image content understanding, widely utilized in target recognition, detection, image recognition, and natural language processing. It comprises two modules: the channel attention module and the spatial attention module.

Channel attention allocates resources across convolutional channels, while spatial attention maps spatial information from the original image to enhance key areas critical for classification. The channel attention mechanism begins by aggregating spatial information with average pooling (AvgPool) and max pooling (MaxPool) to create two 1×1×C vectors representing pooled features. These vectors are processed through a fully connected layer with shared weights, and the outputs are merged using element summation to generate channel dimension weights.

The spatial attention mechanism, complementing channel attention, focuses on H×W location information. It generates two H×W×1 feature maps through max and average pooling across channels, which are then combined in a convolutional layer to produce a single-channel spatial attention map that assigns weights to enhance or suppress specific image locations. Structure of the attention mechanism can view Figure 8.

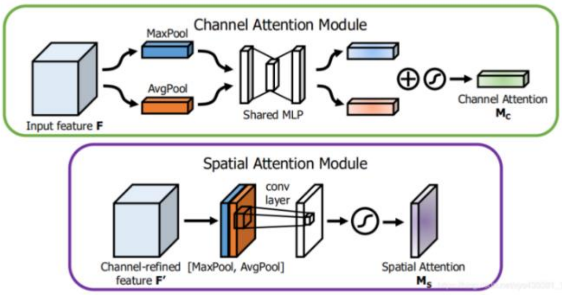



**Figure 8.** Structure of the attention mechanism

#### 2.3.4 Transfer Learning

Transfer learning (Figure 9) is a learning method that transfers the knowledge learned in the source domain to the target domain. In this paper, we use transfer learning method to pre-train a model and apply it as a feature extraction network on tomato pest and disease image dataset, which is used to extract features of tomato pest and disease damage. Transfer learning reduces the amount of training data and the computational power of the experimental platform required to build a deep learning network model, making it unnecessary to retrain and learn the model from scratch.

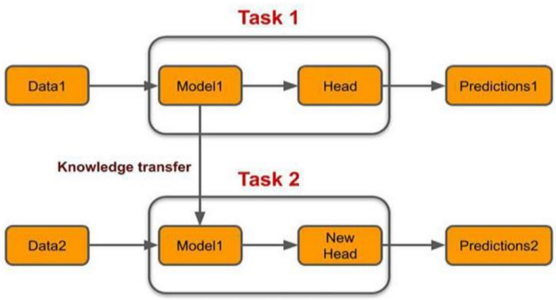



**Figure 9.** Structure of the transfer learning

#### 2.3.5 Building Models

The Deep Residual Network (ResNet50) is selected as the foundational architecture for this study due to its effectiveness in image recognition tasks, attributed to its ability to learn residual functions and alleviate the vanishing gradient problem.

Improvements are integrated into the ResNet50 model to enhance its performance. The focal loss function is introduced to counteract class imbalance by down-weighting well-classified examples and focusing on hard-to-classify ones. The attention mechanism, comprising channel and spatial attention, refines the feature extraction process by concentrating on the most informative parts of the images. DropBlock regularization is employed to prevent overfitting by randomly dropping out regions of the input feature map.

During pre-processing, challenges such as class imbalance and variations in image quality were addressed using data augmentation techniques and image enhancement methods. Stratified k-fold cross-validation was used to ensure the model's generalizability.

The combination of these strategies has resulted in a robust system capable of accurately identifying tomato pests and diseases, with an average recognition accuracy of 98.10%, demonstrating the effectiveness of the proposed improvements in enhancing the ResNet50 model for this specific task, can view Figure 10.


**Figure 10.** Schematic diagram of the improved model structure

### 2.4 Design of Experimental Parameter

Adjusting hyperparameters during training significantly impacts model performance. This paper tests various hyperparameter settings to identify the optimal configuration. The batch size is set to 16, with the SGD (Stochastic Gradient Descent) algorithm used for gradient updates. Three momentum values (0.3, 0.6, and 0.9) and a weight decay of 0.0005 are implemented. Based on optimal momentum, three learning rates for the SGD optimizer (0.01, 0.001, and 0.0001) are chosen using an exponential scale, along with the focal loss function to address class imbalance. The epoch value is set to 100, with test results recorded after each epoch. A DropBlock layer with a 0.4 probability of random parameter dropout is positioned before the fully connected layer.

Model training occurs on a Linux 18.04 Server featuring dual Intel Silver 4210 CPUs (2.2 GHz), 128 GB memory, a 512 GB SSD, a 4 TB HDD, and dual NVIDIA GeForce RTX 3090 GPUs (24 GB memory) for parallel computing. The environment is Anaconda 3 with Python 3.7, and the training framework is Pytorch 1.5.

### 2.5 Evaluation Indicators

In this study, Accuracy (ACC), Precision (P), Recall (R), F1 score and Model Size were used to quantitatively evaluate the performance of the improved model for the identification of tomato. The corresponding equations 2-5.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$P = \frac{T_P}{T_P + F_P} \tag{3}$$

$$R = \frac{T_P}{T_P + F_N} \tag{4}$$

$$F1 - score = \frac{2PR}{P + R} \tag{5}$$

Where TP indicates that the model identifies correctly; FP indicates that the model identifies samples that do not belong to the current class as the current class; FN indicates that the model identifies samples that belong to the current class as other class; and TN indicates that the model identifies samples that do not belong to the current class as other class. Variation of model performance with iteration under different momentum factors can view Figure 11.

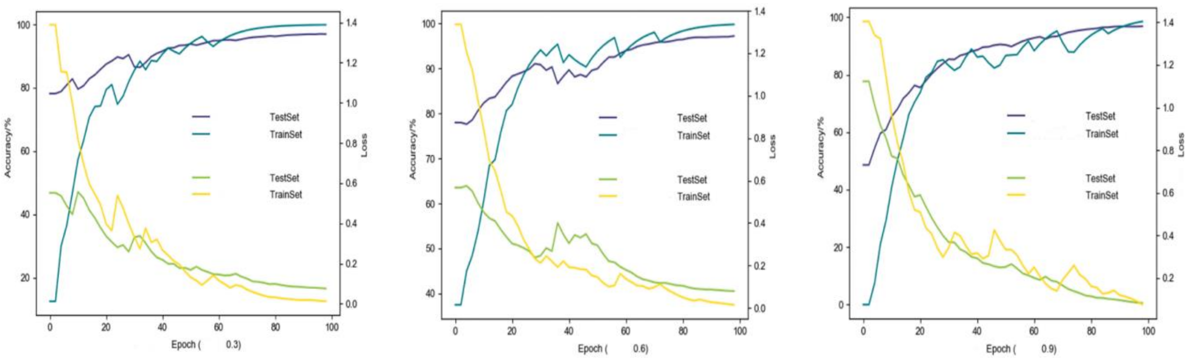



**Figure 11.** Variation of model performance with iteration under different momentum factors

### 2.6 Effect of the Momentum Factor on the Recognition Accuracy

To enhance model recognition accuracy, comparative experiments are conducted with three momentum factor values (m) of 0.3, 0.6, and 0.9, alongside varying learning rates of 0.01, 0.001, and 0.0001. The optimal momentum is determined through these experiments. As illustrated in Figure, momentum factors of 0.3 and 0.6 exhibit slow convergence rates and fluctuating accuracy and loss curves, hindering optimal value attainment. In contrast, a momentum factor of 0.9 achieves recognition accuracies of 97.8% on the test set and 98.6% on the training set, outperforming other configurations. Consequently, m = 0.9 is chosen as the optimal momentum, and the learning rate is tested at 0.001. Results in Table 5.1 indicate that with a momentum factor of 0.9 and a learning rate of 0.001, the training curve shows minimal fluctuation and faster convergence, achieving a peak test set recognition accuracy of 98.1%, making it the most suitable for improving ResNet50 training.

### 2.7 Introducing Different Embedding Methods of Attention Mechanism to Compare Experiments

The improved model selectively embeds the channel attention mechanism and spatial attention mechanism between Stage1 and Stage4: embedding the attention mechanism between shallow network Stage1 and Stage2; embedding the attention mechanism between deep network Stage3 and Stage4; embedding the attention mechanism between shallow network Stage1 and Stage2 and between deep network Stage3 and Stage4, camparison of accuracy can view Figure 12 dan Figure 13 for confusion matrix.

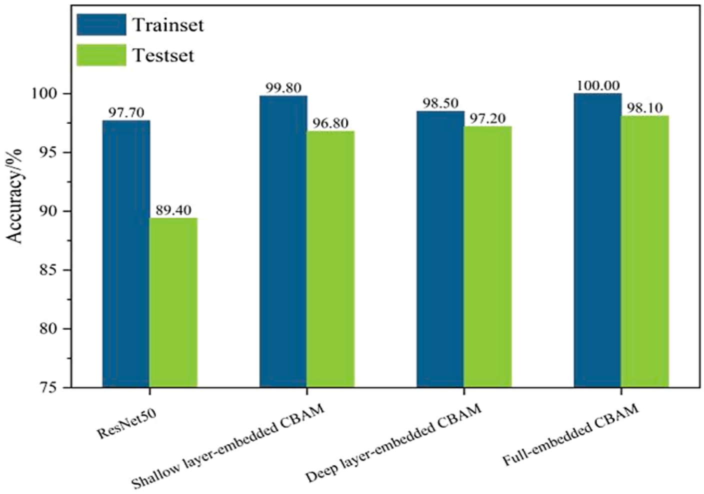
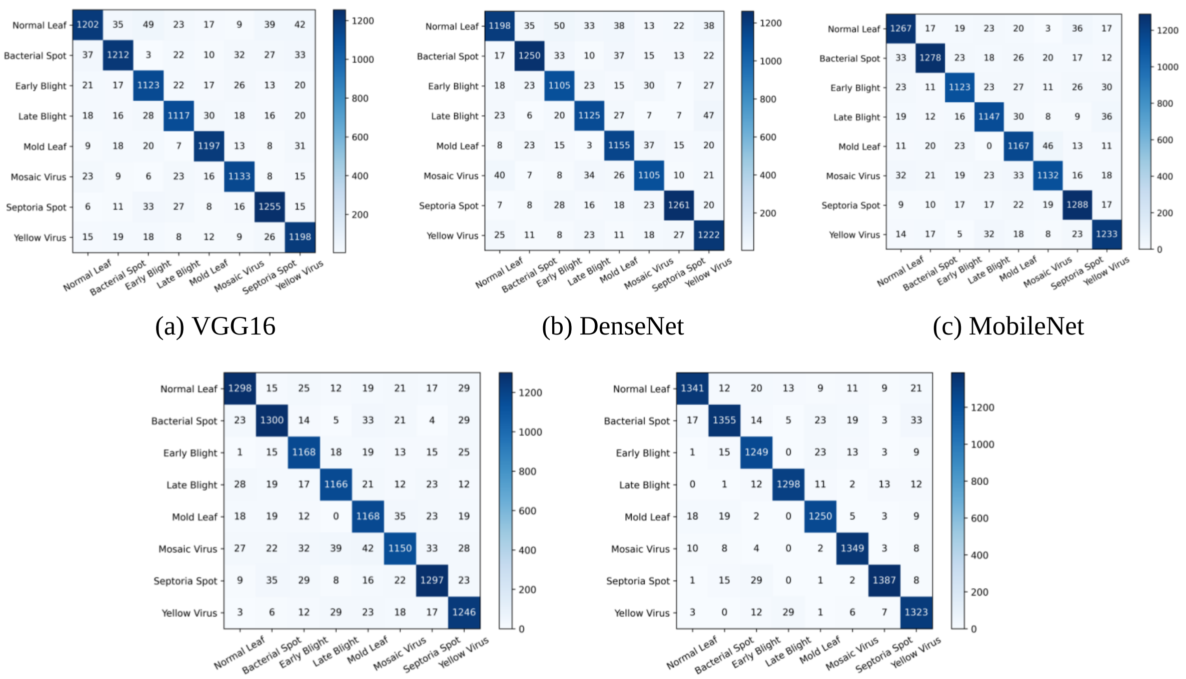



**Figure 12.** Comparison of accuracy under different embedding methods comparedwith the original resnet50 model, the three different embedding




(a) VGG16 (b) DenseNet (c) MobileNet

(d) GoogLeNet (e) Improved Model

**Figure 13.** Comparison of confusion matrix of different models

All three embedding methods enhance recognition accuracy compared to the original ResNet50 model, as the attention mechanism boosts the activation weight of tomato pest and disease areas while minimizing background interference. Using the attention mechanism in both shallow and deep layers simultaneously improves feature extraction for tomato pest and disease damage. The shallow attention mechanism filters relevant textures, while the deep attention mechanism screens and locates high-probability damage areas in high-order semantic dimensions, further enhancing classification accuracy.

## 3. RESULTS AND ANALYSIS

### 3.1 Different network comparison experiments

The improved model shows superior performance in tomato pest and disease recognition, using precision, recall, F1 score, and model size for evaluation. It outperforms classical models like VGG16, DenseNet, MobileNet, and GoogLeNet in terms of F1 score and model efficiency, Specifically, the model's size is notably smaller than VGG16, DenseNet, and GoogLeNet and slightly larger compared to MobileNet, while achieving higher F1 mean values across all comparisons. A confusion matrix further confirms the improved model's enhanced accuracy in distinguishing tomato pests and diseases, marking it as the most effective among the analyzed models.

### 3.2 Discussion

The enhanced ResNet50 model, incorporating focal loss, attention mechanisms, and DropBlock regularization, achieves an impressive 98.10% average recognition accuracy, outperforming models like VGG16, DenseNet, MobileNet, and GoogLeNet. Its success is attributed to effective feature focusing via attention mechanisms and robustness against overfitting with DropBlock.

This research has significant potential to improve pest and disease management in agriculture through accurate and early detection. Despite high accuracy and computational efficiency ideal for real-time applications, the model depends on a well-curated dataset and might perform less effectively in unrepresented environmental conditions.

Future research should consider expanding the dataset to diverse conditions and testing adaptability to other crops. Exploring real-time performance and integration with decision-support systems could increase the model's practical utility. Additionally, integrating the model with IoT or drone technologies for automated crop monitoring could revolutionize proactive pest and disease management, leveraging the model's capability for real-time, on-site processing.

## 4. CONCLUSION

Methods for Classifying Tomato Pests and Diseases: This study utilized an improved deep residual network to identify tomato pests and diseases. We enhanced the original ResNet50 model by incorporating a focal loss function, a channel attention mechanism, and spatial and DropBlock regularization, achieving an average recognition accuracy of 98.10% on the test set, significantly better than the original model.Model Performance Comparison: Experimental comparisons on a self-constructed tomato pest dataset demonstrated that the improved model outperformed others in complex environments. However, the limited sample size and homogeneous experimental conditions may restrict the generalizability of the results. Future research should gather images from diverse growing environments to expand the dataset and evaluate the model's versatility.Implications for Precision Control: The identification model developed in this study can effectively aid in the precision control of tomato pests and diseases, providing practical value for scientific management. Future work could explore the classification of various crop pests and diseases, further supporting the advancement of smart agriculture.

he dataset for this study was limited, with about 500 images per pest type, which may affect the generalizability of the findings. Models trained on such small datasets may perform poorly on diverse real-world data. To enhance adaptability, we plan to collaborate with global tomato farmers to gather a larger and more varied dataset, improving model applicability.While we employed data augmentation strategies to enhance diversity, traditional methods are restricted by predefined rules and may not capture all data variations. Inappropriate parameter settings can also introduce noise, hindering model performance.Recent advancements in Generative Adversarial Networks (GANs) offer significant advantages for data augmentation. GANs can deeply learn and simulate complex data distributions, generating realistic and diverse samples. This capability is vital for enhancing machine learning model generalization in unknown scenarios. Future studies will apply GAN techniques to augment data, aiming to improve model robustness and adaptability through high-quality synthetic data, ultimately achieving more stable performance across varied situations.

**ACKNOWLEDGEMENTS**

I would like to express my sincere gratitude to my thesis advisor, Dr. Tjong Wan Sen, for his invaluable guidance, continuous support, and insightful feedback throughout this research journey. His expertise and encouragement have been instrumental in shaping this thesis.I am deeply grateful to the faculty and staff of the Faculty of Computing at President University for providing a stimulating academic environment and the necessary resources to conduct this research.Special thanks to Dean Rila Mandala, Ph.D., for his leadership and support of graduate research initiatives. I would like to acknowledge the tomato planting base in Wenzhou, China, for allowing me to collect data for this study. Their cooperation was crucial for the successful completion of this research.

My heartfelt appreciation goes to my fellow graduate students and research colleagues for their friendship, collaborative spirit, and stimulating discussions that have enriched my academic experience.I am eternally grateful to my family for their unwavering support, patience, and encouragement throughout my academic pursuits. Their love and understanding have been my strength throughout this journey. I would like to thank all those who have directly or indirectly contributed to this research. Your support, in various forms, has been invaluable in bringing this thesis to fruition.

**REFERENCES**

[1] Sun, L., Liang, K., Wang, Y., et al. (2024). Diagnosis of tomato pests and diseases based on lightweight CNN model. Soft Computing, 28(4), 3393-3413.

[2] Mrosso, S. E., Ndakidemi, P. A., & Mbega, E. R. (2023). Farmers' knowledge on whitefly populousness among tomato insect pests and their management options in tomato in Tanzania. Horticulturae, 9(2), 253.

[3] Jeong, S., Jeong, S., & Bong, J. (2022). Detection of tomato leaf miner using deep neural network. Sensors, 22(24), 9959. DOI: 10.3390/s22249959.

[4] Liu, J., Wang, X., Miao, W., et al. (2022). Tomato pest recognition algorithm based on improved YOLOv4. Frontiers in Plant Science, 13, 814681. DOI: 10.3389/fpls.2022.814681.

[5] Ali, M. A., Dhanaraj, R. K., & Nayyar, A. (2023). A high performance-oriented AI-enabled IoT-based pest detection system using sound analytics in large agricultural fields. Microprocessors and Microsystems, 103, 104946. DOI: 10.1016/j.micpro.2023.104946.

[6] Liu, S., Fu, S., Hu, A., et al. (2023). Research on insect pest identification in rice canopy based on GA-Mask R-CNN. Agronomy, 13(8), 2155. DOI: 10.3390/agronomy13082155.

[7] Lin, T. L., Chang, H. Y., & Chen, K. H. (2020). The pest and disease identification in the growth of sweet peppers using faster R-CNN and mask R-CNN. Journal of Internet Technology, 21(2), 605-614. DOI: 10.3966/160792642020022102013.

[8] Yan, T., Xu, W., Lin, J., et al. (2021). Combining multi-dimensional convolutional neural network (CNN) with visualization method for detection of Aphis gossypii Glover infection in cotton leaves using hyperspectral imaging. Frontiers in Plant Science, 12, 604510. DOI: 10.3389/fpls.2021.604510.

[9] Pattnaik, G., Shrivastava, V. K., & Parvathi, K. (2020). Transfer learning-based framework for classification of pest in tomato plants. Applied Artificial Intelligence, 34(13), 981-993.

[10] Cheng, X., Zhang, Y., Chen, Y., et al. (2017). Pest identification via deep residual learning in complex background. Computers and Electronics in Agriculture, 141, 351-356.

[11] She, H., Wu, L., & Shan, L. (2020). An improved rice pest recognition method based on SSD network model. Journal of Zhengzhou University (Natural Science Edition), 52(3).

[12] Zhang, J., Kong, F., Wu, J., et al. (2018). Cotton disease recognition model based on improved VGG convolutional neural network. Journal of China Agricultural University, 23(11), 161-171.

[13] Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. Computers and Electronics in Agriculture, 145, 311-318.

[14] Shi, F., Wang, H., & Huang, H. (2022). Research on potato eye detection and recognition based on convolutional neural network. Journal of Chinese Agricultural Mechanization, 43(06), 159-165.

[15] Jenifa, A., Ramalakshmi, R., & Ramachandran, V. (2019). Cotton leaf disease classification using deep convolution neural network for sustainable cotton production. In 2019 IEEE International Conference on Clean Energy and Energy Efficient Electronics Circuit for Sustainable Development (INCCES) (pp. 1-3). IEEE.

[16] Caldeira, R. F., Santiago, W. E., & Teruel, B. (2021). Identification of cotton leaf lesions using deep learning techniques. Sensors, 21(9), 3169.

[17] He, D., Wang, P., Niu, T., et al. (2022). A classification model for grapevine downy mildew severity in fields based on improved residual network. Transactions of the Chinese Society for Agricultural Machinery, 53(01), 235-243.

[18] Li, D., Ahmed, F., Wu, N., et al. (2022). YOLO-JD: A deep learning network for jute diseases and pests detection from images. Plants, 11(7), 937.

[19] Mamdouh, N., & Khattab, A. (2021). YOLO-based deep learning framework for olive fruit fly detection and counting. IEEE Access, 9, 84252-84262.

[20] Shi, W., & Nianqiang, L. (2020). Application of target detection algorithm based on deep learning in farmland pest recognition. International Journal of Artificial Intelligence and Applications (IJAIA), 11(3).

[21] Khan, M. A., Akram, T., Sharif, M., et al. (2018). CCDF: Automatic system for segmentation and recognition of fruit crops diseases based on correlation coefficient and deep CNN features. Computers and Electronics in Agriculture, 155, 220-236.
[22] Bollis, E., Pedrini, H., Avila, S., & Ieee. (2020). Weakly supervised learning guided by activation mapping applied to a novel citrus pest benchmark. In IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Electr Network.
[23] Bracino, A. A., Concepcion, R. S., Bedruz, R. A. R., Dadios, E. P., Vicerra, R. R. P., & Ieee. (2020). Development of a hybrid machine learning model for apple (Malus domestica) health detection and disease classification. In IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM). Manila, PHILIPPINES.
[24] Burhan, S. A., Minhas, S., Tariq, A., Hassan, M. N., & Ieee. (2020). Comparative study of deep learning algorithms for disease and pest detection in rice crops. In 12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI). Bucharest, ROMANIA.
[25] Hu, W. J., Fan, J., Du, Y. X., Li, B. S., Xiong, N. X., & Bekkering, E. (2020). MDFC-ResNet: An agricultural IoT system to accurately recognize crop diseases. IEEE Access, 8, 115287-115298.
[26] Nagar, H., Sharma, R. S., & Ieee. (2020). A comprehensive survey on pest detection techniques using image processing. In International Conference on Intelligent Computing and Control Systems (ICICCS). Anna Univ, Vaigai Coll Engn, Madurai, INDIA.
[27] Wang, Q. Y., He, G. Q., Li, F., Zhang, H. X., & Ieee. (2020). A novel database for plant diseases and pests classification. In 10th IEEE International Conference on Signal Processing, Communications and Computing (IEEE ICSPCC). Univ Macau, Macau, PEOPLES R CHINA.
[28] Ahmed, M., Mahajan, T., Sharma, B. D., Kumar, M., Singh, S. K., & Ieee. (2021). AI-based detection of pest infected crop and leaf. In 3rd International Conference on Signal Processing and Communication (ICPSC). Coimbatore, INDIA.
[29] Albanese, A., Nardello, M., & Brunelli, D. (2021). Automated pest detection with DNN on the edge for precision agriculture. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 11(3), 458-467.
[30] Correa, E., García, M., Grosso, G., Huamantoma, J., Ipanaqué, W., & Ieee. (2021). Design and implementation of a CNN architecture to classify images of banana leaves with diseases. In IEEE IFAC International Conference on Automation (ICA) / 24th Congress of the Chilean Association of Automatic Control (ACCA). Electr Network.
[31] Kibriya, H., Abdullah, I., Nasrullah, A., & Ieee. (2021). Plant disease identification and classification using convolutional neural network and SVM. In 18th International Conference on Frontiers of Information Technology (FIT). Islamabad, PAKISTAN.
[32] Li, L. L., Zhang, S. J., & Wang, B. (2021). Plant disease detection and classification by deep learning—a review. IEEE Access, 9, 56683-56698.
[33] Liu, L., Xie, C. J., Wang, R. J., Yang, P., Sudirman, S., Zhang, J., Li, R., & Wang, F. Y. (2021). Deep learning based automatic multiclass wild pest monitoring approach using hybrid global and local activated features. IEEE Transactions on Industrial Informatics, 17(11), 7589-7598.
[34] Mahenthiran, N., Sittampalam, H., Yogarajah, S., Jeyarajah, S., Chandrasiri, S., & Kugathasan, A. (2021). Smart pest management: An augmented reality-based approach for organic cultivation. In 2nd International Informatics and Software Engineering Conference (IISEC) - Artificial Intelligence for Digital Transformation. Ankara, TURKEY.
[35] Nagar, H., Sharma, R. S., & Ieee. (2021). Pest detection on leaf using image processing. In 11th International Conference of Computer Communication and Informatics (ICCCI). Sri Shakthi Inst Engn & Technol, Coimbatore, INDIA.
[36] Shrimali, S., & Ieee. (2021). PlantifyAI: A novel convolutional neural network based mobile application for efficient crop disease detection and treatment. In 2nd Asia Conference on Computers and Communications (ACCC). Electr Network.
[37] Elijah, O., Rahim, S. K. A., Abioye, E. A., Musa, M. J., Salihu, Y. O., Oremeyi, A. A., & Ieee. (2022). Decision support platform for production of chili using IoT, cloud computing, and machine learning approach. In 4th IEEE-Nigeria-Section International Conference on Disruptive Technologies for Sustainable Development (IEEE NIGERCON). Nile Univ Nigeria, Abuja, NIGERIA.
[38] Grimaldo, G., Rodríguez, H., Cabrera, V. L., & Ieee. (2022). Convolutional neural network model for the detection of diseases and pests in coffee crops. In 8th International Engineering, Sciences and Technology Conference (IESTEC). Technol Univ Panama, Panama City, PANAMA.
[39] Huang, C. H., Chen, B. W., Lin, Y. J., & Zheng, J. X. (2022). Smart crop growth monitoring based on system adaptivity and edge AI. IEEE Access, 10, 64114-64125.
[40] Sharma, R. P., Ramesh, D., Pal, P., Tripathi, S., & Kumar, C. (2022). IoT-enabled IEEE 802.15.4 WSN monitoring infrastructure-driven fuzzy-logic-based crop pest prediction. IEEE Internet of Things Journal, 9(4), 3037-3045.
[41] Sunil, C. K., Jaidhar, C. D., & Patil, N. (2022). Cardamom plant disease detection approach using EfficientNetV2. IEEE Access, 10, 789-804.
[42] Toscano-Miranda, R., Hoyos, W., Caro, M., Aguilar, J., Trebilcok, A., Toro, M., & Ieee. (2022). A classification model of cotton boll-weevil population. In 48th Latin American Computer Conference (CLEI). Quindio Univ, Armenia, COLOMBIA.
[43] Kang, G. W., Hou, L., Zhao, Z., Lang, B. B., & Ieee. (2023). Research on the application of convolutional neural network based on YOLO algorithm in pest small target detection. In 3rd Asia-Pacific Conference on Communications Technology and Computer Science (ACCTCS). Shenyang, PEOPLES R CHINA.

[44] Nazari, L., & Zinati, Z. (2023). Gene expression classification for biomarker identification in maize subjected to various biotic stresses. IEEE-ACM Transactions on Computational Biology and Bioinformatics, 20(3), 2170-2176.

[45] Shafik, W., Tufail, A., Namoun, A., De Silva, L. C., & Apong, R. (2023). A systematic literature review on plant disease detection: Motivations, classification techniques, datasets, challenges, and future trends. IEEE Access, 11, 59174-59203.

[46] Ahmed, A. A., Ahmed, J., & Ieee. (2024). An image-based plant weed detector using machine learning. In International Conference on Computing, Networking and Communications (ICNC). Hi.

[47] Arasi, M. A., Almuqren, L., Issaoui, I., Almalki, N. S., Mahmud, A., & Assiri, M. (2024). Enhancing red palm weevil detection using bird swarm algorithm with deep learning model. IEEE Access, 12, 1542-1551.

[48] Bahrami, M., Pourhatami, A., Maboodi, M., & Ieee. (2024). Tomato leaf disease detection using transfer learning: A comparative study. In 13th Iranian/3rd International Machine Vision and Image Processing Conference (MVIP). Tehran, IRAN.

[49] Haider, I., Khan, M. A., Nazir, M., Hamza, A., Alqahtani, O., Alouane, M. T. H., & Masood, A. (2024). Crops leaf disease recognition from digital and RS imaging using fusion of multi self-attention RBNet deep architectures and modified dragonfly optimization. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17, 7260-7277.

[50] Karthik, R., Ajay, A., Bisht, A. S., Illakiya, T., & Suganthi, K. (2024). A deep learning approach for crop disease and pest classification using Swin transformer and dual-attention multi-scale fusion network. IEEE Access, 12, 152639-152655.

[51] Khan, K. H., Aljaedi, A., Ishtiaq, M. S., Imam, H., Bassfar, Z., & Jamal, S. S. (2024). Disease detection in grape cultivation using strategically placed cameras and machine learning algorithms with a focus on powdery mildew and blotches. IEEE Access, 12, 139505-139523.

[52] Mohyuddin, G., Khan, M. A., Haseeb, A., Mahpara, S., Waseem, M., & Saleh, A. M. (2024). Evaluation of machine learning approaches for precision farming in smart agriculture system: A comprehensive review. IEEE Access, 12, 60155-60184.

[53] Naseer, A., Amjad, M., Raza, A., Munir, K., Samee, N. A., & Alohali, M. A. (2024). A novel transfer learning approach for detection of pomegranates growth stages. IEEE Access, 12, 27073-27087.

[54] Nie, J., Jiang, J. C., Li, Y., Li, J. B., Chao, X. W., & Ercisli, S. (2024). Efficient detection of cotton verticillium wilt by combining satellite time-series data and multiview UAV images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17, 13547-13557.

[55] Rashid, R., Aslam, W., Aziz, R., & Aldehim, G. (2024). An early and smart detection of corn plant leaf diseases using IoT and deep learning multi-models. IEEE Access, 12, 23149-23162.

[56] Ussupov, J., Yessenova, M., Abdikerimova, G., Aimbetov, A., Baktybekov, K., Murzabekova, G., & Aitimova, U. (2024). Analysis of formal concepts for verification of pests and diseases of crops using machine learning methods. IEEE Access, 12, 19902-19910.


**BIBLIOGRAPHY OF AUTHORS**

Peng Linli is a master's student at President University,specializing in voice and image recognition,big data analytics,and business intelligence ,His research interests focus on the development and application of advanced deep learning techniques and transfer Learning for solving complex problems in various domains.

Tjong Wan Sen, born in Kediri, Indonesia. He received Ph.D. degree from Institut Teknologi Bandung in 2009. Since 2010, he has been with the Faculty of Computing, President University, Jababeka, Indonesia, where he is currently a lecturer and researcher. His research interests include automatic speech/speaker recognition, computer vision, artificial intelligence, and embedded systems.

Dr. Hasanul Fahmi is a lecturer at President University with expertise in Computer Science. He holds a Doctorate (Dr.) in Ilmu Komputer from Universitas Sumatera Utara (2021), a Master’s degree (M.Kom) from the same university (2014), and a Bachelor’s degree (S.Kom) from Universitas Islam Negeri (UIN) Malang (2011). His research interests include Artificial Intelligence, Machine Learning, Deep Learning, Data Mining, and Decision Support Systems, among others. Proficient in tools such as Visual Studio Code, Android Studio, Python, Laravel, and PHP, Dr. Fahmi combines his academic and technical expertise to advance innovation and foster academic excellence.

RusdiantoRoestam was born inBanjarmasin,Kalimantan.Indonesia.in 1957.He receivedthe B.S.degree inelectronic engineering fromSurabaya Institute of Technology, Surabaya, Indone-sia.in 1982.and the .Sc.and Ph.D.degresin information technology from the University ofSalford,Manchester,U.K. in 1991 and 1996.respectively. He joined the Agency for the Assessment and Application of Technology,Indonesia.which has been merged with the National Researchand Innovation Agency,Indonesia. for over 20 yearswith a focus on computer network protocols under the Advanced NetworkProtocol Laboratory,In 2017,he ioined the Faculty of Computing, PrsidentUniversity, with a focus on research interest in the Internet of Things (loT)and as the author of several articles published nationally and internationallyrelated to this research field.