DOI: 10.24996/ijs.2026.67.2.40

ISSN: 0067-2904

TOMATOCNN-RF: Machine Learning Based- Hybrid Approach for Tomato Leaf Disease Diagnosis

Payman Hussein Hussan *, Syefy mohammed mangj
Department of Computer Networks and Software Techniques, Babylon Technical Institute, Al-Furat Al-Awsat Technical University, Kufa, Iraq

Received:4/1/2024 Accepted: 26/ 2/2025 Published: 28/2/2026

Abstract

Tomatoes are among the most significant vegetables in the world. It is regarded as a cornerstone of the economies of numerous countries. Tomato crops are susceptible to multiple illnesses that can diminish yields. The early diagnosis of diseases in tomato leaves substantially enhances tomato production and allows farmers to address these challenges more efficiently. Conventional techniques dependent on visual assessments are laborious, time-consuming, and susceptible to human mistakes. Machine learning (ML) and deep learning (DL) techniques have emerged as potent instruments for automating and enhancing the precision of illness diagnosis to tackle these difficulties. This paper presents a hybrid methodology (TOMATOCNN-RF) that integrates a convolutional neural network (CNN) with a Random Forest (RF) for the detection and classification of illnesses in tomato plants. The model was trained on a portion of the publicly available Plant Village dataset, comprising 10 categories of tomato illnesses. We developed the architecture of the hybrid TOMATOCNN-RF model by constructing a new architecture CNN network and substituting the fully connected layer with a Random Forest classifier. The ablation study trials indicate that the hybrid approach attained a classification performance of 98%.

Keywords: Agricultural Technology; Convolutional Neural Networks; Random Forest; Hybrid Modeling; Tomato crop Disease Detection

TOMATOCNN-RF:النهج هجين قائم على التعلم الآلي لتشخيص أمراض أوراق الطماطم

بيمان حسين حسن*, سيفي محمد منجي
قسم تقنيات شبكات وبرمجيات الحاسوب, المعهد التقني بابل, جامعة الفرات الأوسط التقنية, الكوفة, العراق

الخلاصة

تُعَد الطماطم من أهم الخضروات في العالم. وتُعَد حجر الزاوية في اقتصادات العديد من البلدان. محاصيل الطماطم عُرضة لأمراض متعددة يمكن أن تقلل من الغلة. ويعزز التشخيص المبكر للأمراض في أوراق الطماطم بشكل كبير إنتاج الطماطم ويسمح للمزارعين بمعالجة هذه التحديات بكفاءة أكبر. إن التقنيات التقليدية التي تعتمد على التقييمات البصرية شاقة وتستغرق وقتًا طويلاً وعرضة للأخطاء البشرية. وقد ظهرت تقنيات التعلم الآلي والتعلم العميق كأدوات فعّالة لأتمتة وتعزيز دقة تشخيص الأمراض لمعالجة هذه الصعوبات. تقدم هذه الورقة منهجية هجينة (TOMATOCNN-RF) تدمج شبكة عصبية ملتوية (CNN)

* Email : inb.beman10@atu.edu.iq

مع غابة عشوائية (RF) للكشف عن الأمراض وتصنيفها في نباتات الطماطم. تم تدريب النموذج على جزء من مجموعة بيانات Plant Village المتاحة للجمهور، والتي تضم 10 فئات من أمراض الطماطم. قمنا بتطوير بنية نموذج TOMATOCNN-RF الهجين من خلال إنشاء شبكة CNN ذات بنية جديدة واستبدال الطبقة المتصلة بالكامل بمصنف غابة عشوائية. تشير تجارب دراسة الاستئصال إلى أن النهج الهجين حقق أداء تصنيف بنسبة 98%.

1. Introduction

We depend on food plants like we do on oxygen. Crops are essential for food, and food is vital for life. The tomato is the most prevalent and recognizable food in our daily life, serving as a vital horticulture crop throughout all global regions. The tomato, belonging to the Solanaceae family and originating from the Andean region of South America, has, according to "Food and Agriculture Organization Statistics", become one of the most significant and widely cultivated horticultural crops globally over the past fifty years [1]

Due to its genetic characteristics, the tomato is vulnerable to numerous plant diseases induced by pathogens, including fungi, bacteria, phytoplasmas, and viruses [2] . Consequently, a critical research focus in precision agriculture is the identification of diseases through images of plant leaves. Advancements in artificial intelligence, image processing, and graphical processing units can significantly improve the precision of plant protection and growth techniques. Since many plant diseases exhibit distinct visual symptoms, such as changes in color, texture, and shape, learning models must accurately detect and recognize these unique signs[3].

Transfer learning-based deep learning models are employed for disease categorization because they can yield adequate outcomes [4]. Consequently, to develop a specific Convolution Neural network-based model, it is essential to compare it with transfer learning models to validate the enhancement. Agarwal et al. [5] developed a streamlined and efficient CNN model of three convolutional layers, three max pooling layers, and two fully connected layers. The proposed model, when compared to VGG-16 and InceptionV3, attained an enhanced accuracy of 91.20%. Jiang et al. [6], utilized 1000 images to study three distinct tomato leaf diseases. The activation function of ResNet-50 was substituted with Leaky-ReLU, and the kernel size of the initial convolutional layer was modified to 11x11. Robert et al. [7] presented a method for anomaly detection and identifying plant leaf diseases in tomato crops. Fully-Region-Convolution Neural Networks were employed to detect plant irregularities, whereas CNNs are applied for disease identification. Upon establishing the disease's classification, they proposed a specific treatment. This prototype can be utilized for any Lycopersicon crop as it autonomously gathers images through the box. It can detect various diseases accurately. The algorithm is effective for the fungal disease affecting pepper plant leaves.

Mia et al. [8] employed CNN-based transfer learning and conventional machine learning methods, evaluating their efficacy in classifying cucumber diseases. 525 picture samples representing six illness categories were acquired and pre-processed in the investigation. The sample size was augmented to a total of 4,200 photos by applying data augmentation techniques. In the subsequent step, numerous machine-learning methods were implemented, and a comparative analysis was conducted. The CNN-based MobileNet attained the best accuracy of 93.23% among all models.

The classification of tomato leaf disease conducted via Deep CNN, incorporating residual blocks and attention extraction modules as implemented by Zhao et al [9]. The dataset consisted of ten classes, including one healthy class, with an initial sample size of 4585 images. SEResNet50 has achieved greater accuracy than ResNet50.

Xian et al. [10] employed the ELM classification algorithm with a single-layer FFNN to evaluate photos of tomato plant leaves. Images underwent pre-processing in the HSV color space, and features were retrieved using Haralick textures. The ELM classifier underwent training and evaluation on a portion of the Plant-Village dataset. The results achieved an accuracy of 84.94%, surpassing other models like SNN and DT, demonstrating the effectiveness of ELM in plant disease classification. Balakrishna et al. [11] implemented multiple machine learning and deep learning approaches, including Fuzzy-SVM, CNN, and R-CNN. They employed photos of tomato leaves afflicted by six diseases, in addition to healthy specimens, utilizing techniques such as image scaling, color thresholding, and gradient local ternary pattern for feature extraction. The R-CNN classifier attained a remarkable accuracy of 96.735%, surpassing alternative approaches. This study underscores the significance of early disease diagnosis in enhancing agricultural output and mitigating future losses.

Motivation

Despite prior research indicating encouraging results in classifying tomato leaf diseases, opportunities for enhancement remain. This improvement would aid practitioners in various applications, including ensuring agricultural output and food security. It substantially decreases crop losses, enhances production quality, and reduces pesticide utilization.

Therefore, to develop a high-performance approach aimed at Early diagnosis of tomato leaf diseases, The following questions are addressed: What is the proposed approach for automatic classification of tomato diseases, and how does it aim to significantly improve the accuracy and performance of disease recognition systems? Although the previous studies performed well, it can be concluded that deep learning techniques have not achieved good results in classifying tomato leaf disease. Can machine learning (ML) and deep learning (DL) techniques merge to improve tomato leaf disease detection accuracy? Most modern models face unbalanced dataset classes; how can obtaining a balanced dataset enhance performance and dependability, and what insights were gained from the comparative performance analysis of the ablation study compared to other contemporary approaches?

This paper aims to answer the above questions, which are significant for the early identification of tomato plant diseases. The main contributions that this paper brings out are:

1- Suggesting an effective approach for the automated classification of tomato diseases to substantially improve the accuracy and efficacy of disease recognition systems, thereby aiding sustainable agricultural practices and enhancing tomato crop management.

2- Producing a convolutional neural network called Custom CNN model for multi-scale feature extraction to enable effective disease detection in low-contrast images.

3- A hybrid methodology (TomatoCNN-RF) that integrates a custom convolutional neural network and random forest was suggested to identify nine tomato crop diseases.

4- Enhancing performance and reliability through dataset augmentation to achieve a balanced dataset.

5- Comparative performance analysis of the ablation study with other relevant methodologies.

The remaining portions of the paper are organized as follows: Section 2 delineates the methodology employed for the study. Section 3 presents the findings and performs an extensive array of comparative evaluations among the networks used, succeeded by an in-depth discussion and analysis of the results in Section 4. Conclusions are finally offered in Section 5.

2. MATERIALS AND METHODS

This section describes the proposed methodology for the diagnosis of tomato diseases from the RGB images is described. The schematic representation of the pipeline of the proposed method is shown in Figure 1.

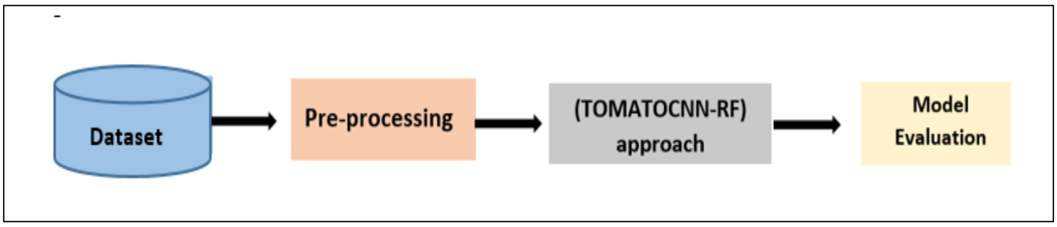

Figure 1: The Proposed Methodology

The various phases involved in our research approach are illustrated in the above figure. The early stages are called Dataset and Pre-processing; they entail acquiring the essential data and carrying out all required pre-processing activities. The third phase is the hybrid approach (TOMATOCNN-RF) incorporating Custom CNN and RF, which focuses on training a custom CNN by using the data that has been pre-processed and extracting significant features as vectors. Then, we train a machine learning classifier (Random Forest) using features captured for the infected leaves to classify the images of tomato leaf diseases into ten classes. In the concluding phase, which entails performance analysis and model evaluation, we assess the effectiveness of our method and examine the efficiency of the trained model.

2.1 Dataset Description

This proposed methodology employed a publicly accessible, research-focused PlantVillage dataset for categorizing tomato leaf diseases. This adaptable dataset comprises around 54,000 photos of 14 varieties of crops (fruits and vegetables). Our research utilized 16,012 colored and labeled images of tomato leaves, encompassing both healthy samples and nine distinct leaf diseases: Bacterial Spot, Early Blight, Late Blight, Leaf Mold, Septoria Leaf Spot, Two Spotted Spider Mite, Target Spot, Yellow Leaf Curl Virus, and Tomato Mosaic Virus. Figure 2 depicts the allocation of photos across each category. Every image in the collection possesses a resolution of 256 × 256 pixels, utilizes the RGB color system, and is formatted as JPG. Figure 3 displays photos from ten classes from the PlantVillage dataset.

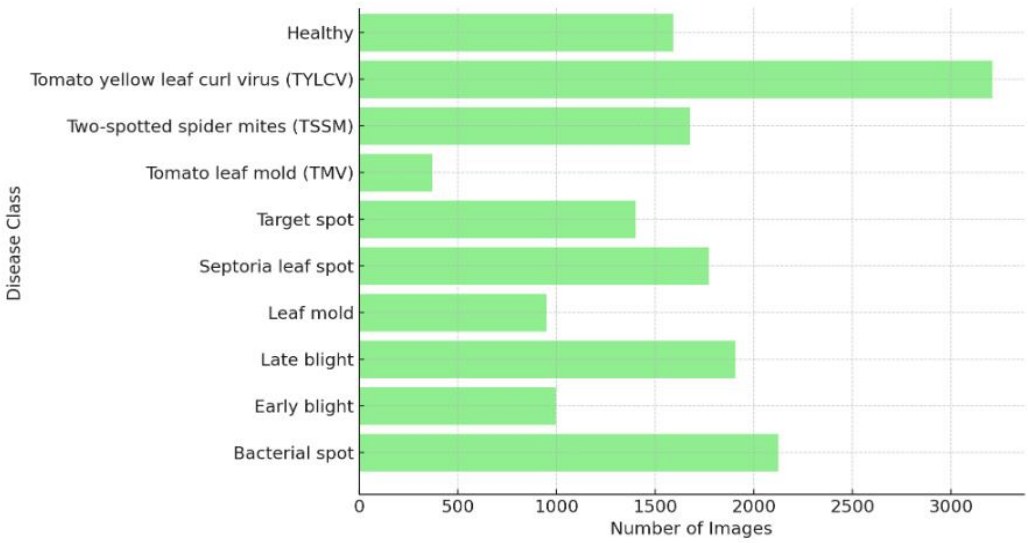

Figure 2: Distribution of images across tomato disease classes

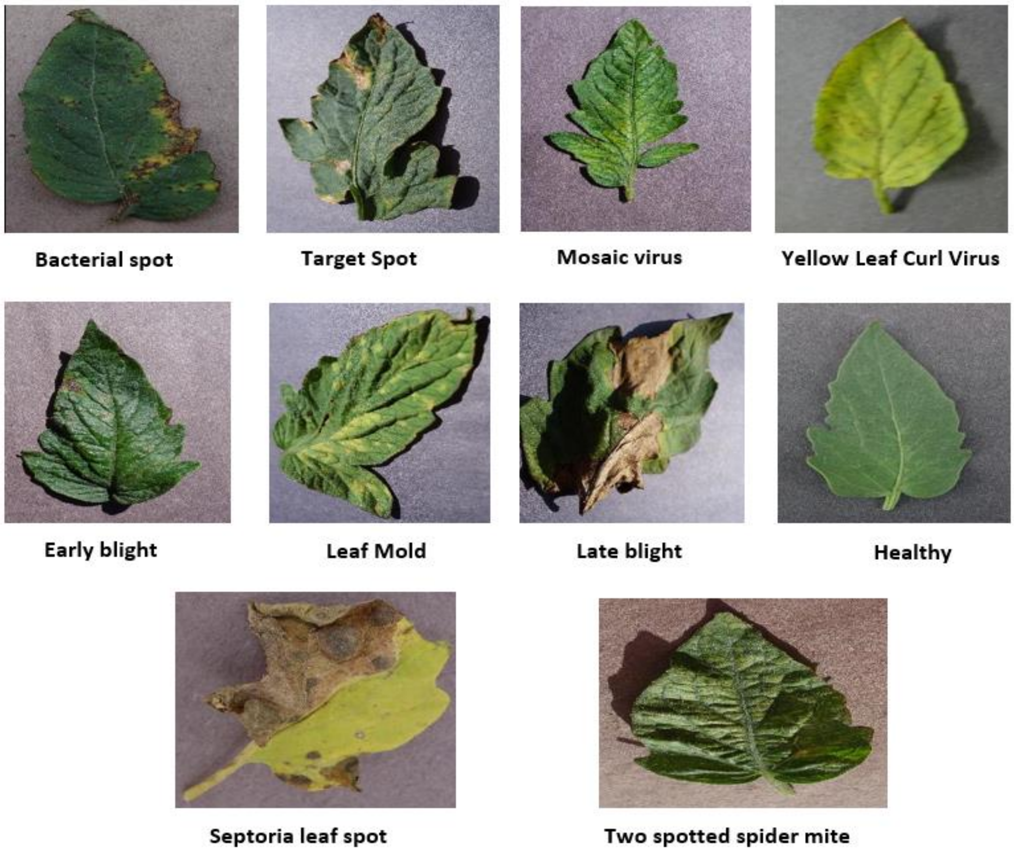

Figure 3: Sample images of the PlantVillage dataset

2.2 Dataset Pre-processing

Our proposed approach utilizes different image pre-processing techniques to improve the reliability and robustness of the hybrid model. This involves normalizing pixel values from 0 to 255 to minimize variations and resizing to a uniform dimension of 128 × 128 × 1. Subsequently, the image's pixel intensity is normalized from its original range of 0–255 to the interval [0, 1] by uniformly dividing each value by 255, enhancing the numerical stability of computations.

Ultimately, an imbalance issue appeared in the training dataset due to certain classes containing more photos than others. Training a deep learning model on an imbalanced dataset may lead to inadequate generalization. This may happen since the model could develop a bias towards the majority classes, resulting in low performance for the minority classes. To mitigate the problem of an imbalanced dataset, a commonly employed technique called data augmentation was utilized. Data augmentation functions as a regularization technique to prevent overfitting and improve the model's resilience by adding variability to the dataset. Data augmentation aims to artificially increase the number of items in the dataset. This study utilized data augmentation across eight categories to generate improved images from the original dataset, thereby equilibrating the training dataset by augmenting the picture quantity in those categories.

Different augmentation techniques have been applied to increase image samples for (Early blight, Late blight, Leaf Mold, Septoria leaf spot, Two-spotted spider mite, Target Spot, Tomato mosaic virus, and healthy) images, which include rotating images by 90 degrees, Randomly shifting them horizontally by 10% of the total width, shifting them vertically by

10% of total height, and applying vertical flipping [12], we obtained a balanced dataset containing approximately 22,600 images, partitioned into 18160 for training and 4440 for testing. Figure 4 illustrates a random sample image and its various augmented methods.

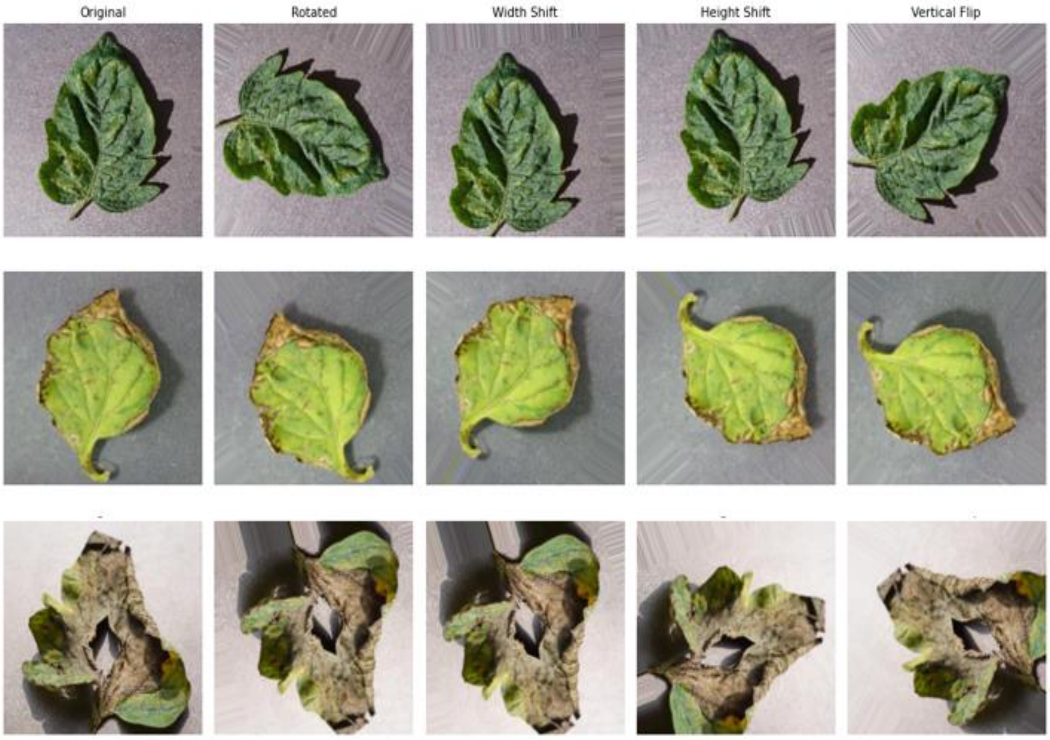

Figure 3: Images of Tomato Leaves After Applying Augmentation Methods

2.3 Proposed hybrid approach(TOMATOCNN-RF) architecture

The proposed hybrid approach, TomatoCNN-RF, for detecting and classifying tomato diseases is a combination method based on Custom Convolution Neural Network and Random Forests. Its block diagram is illustrated in Figure 5.

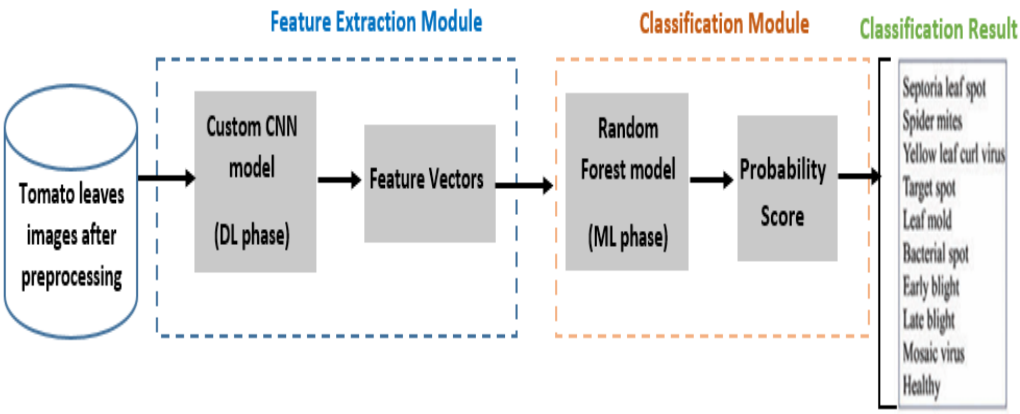

Figure 5: Block diagram of the hybrid approach (TomatoCNN-RF)

In the proposed hybrid (TomatoCNN-RF) approach, the "Custom CNN" part is used as a feature extractor, and the Random Forest is used as a classifier. The feature extraction process is crucial to an automatic image classification methodology. Our proposed approach employed a Custom CNN deep learning model for feature extraction due to the CNN deep learning model's hierarchical structure, which facilitates the extraction and acquisition of high quality information at every layer. Additionally, it might diminish the intricacy of the network architecture.

Our proposed custom CNN aims to extract high-level features of tomato leaf infections from images. The architecture is designed with multiple stacked convolutional layers to achieve this objective. The proposed custom Convolutional Neural Network (CNN) begins with an input layer designed for 128×128×3 images (RGB channels) and consists of two convolutional stacks; the initial stack contains two Conv2D layers, each employing 32 filters, followed by a Batch Normalization layer to improve training stability and efficiency. Two further Conv2D layers, each including 64 filters, are incorporated in the second stack, followed by Batch Normalization layers. The network utilizes MaxPooling2D layers after each convolutional stack to decrease spatial dimensions, reducing the image size from 128×128 to 64×64 and 32×32. Feature maps are generated sequentially by applying multiple convolutional and pooling layers to the tomato data. The tomato leaf dataset feature maps are transformed into an array with one dimension by a flattening layer. This converts the 3-D feature maps of dimensions 32×32×64 into a 1D array, preparing the data for the fully connected or output layers. This design emphasizes feature extraction via convolution and dimensionality reduction by pooling, enabling efficient processing for image classification tasks. Table 1 illustrates the specifications of the topology of our custom CNN model.

Table 1: The Proposed Custom CNN Model Summary

Layer	Output shape	Parameters
Input layer	(128,128,3)	0
Conv2D (32 filters, 3x3)	(128,128,32)	896
BatchNormilization	(128,128,32)	128
Conv2D (32 filters, 3x3)	(128,128,32)	9248
BatchNormilization	(128,128,32)	128
MaxPooling2D	(64,64,32)	0
Conv2D (64 filters, 3x3)	(64,64,64)	18496
BatchNormilization	(64,64,64)	256
Conv2D (64 filters, 3x3)	(64,64,64)	36928
BatchNormilization	(64,64,64)	256
MaxPooling2D	(32,32,64)	0
Flatten	(65536)	0

A Random Forest (RF) [13] model is considered to be trained to distinguish feature vectors into ten different classes of diseases using decision trees to deal with the data combinations feasibly. The integration of CNN-extracted features with the RF classifier aims to harness the strengths of both approaches: the feature extraction by the CNN and the classification's capability of the RF's robustness, as shown in Figure 5.

In constructing the architecture of our hybrid TomatoCNN-RF model, A random forest classifier utilizes many decision trees, and it will be trained using the extracted feature vectors from custom CNN as input, where the top layer of the custom CNN, specifically the fully connected layer with the softmax activation function, was substituted with a Random Forest model. The features extracted by the CNN layers are input into the Random Forest classifiers following their transformation and pass through the initial fully connected layer.

Our suggested methodology could benefit greatly from using Random Forest instead of a classification strategy, since it trains multiple decision trees using randomly selected training data. Two purposes are accomplished by randomization in a random forest: first, it ensures that each decision tree is trained on a separate set of data points; second, it selects each test within a decision tree from a random subset of features. Random forests can produce predictions by taking the results of each decision tree and picking the one with the highest frequency of occurrence. This action aims to decrease prediction variance by increasing forest tree variety. Figure 6 illustrates the RF classifier, which partitions the input information into several decision trees for the purpose of classification. The decision tree comprises multiple nodes that accept subsets of the dataset. Each decision tree forecasts the result during training. A majority vote on all classified trees determines the final decision.
Each random tree in this research consists of 200 decision trees. The main reason for selecting this number of decision trees is that Random Forest reduces variance as the number of trees increases. However, adding more trees does not significantly improve accuracy, and an excessively large number of trees can increase computation time without significant accuracy gains.

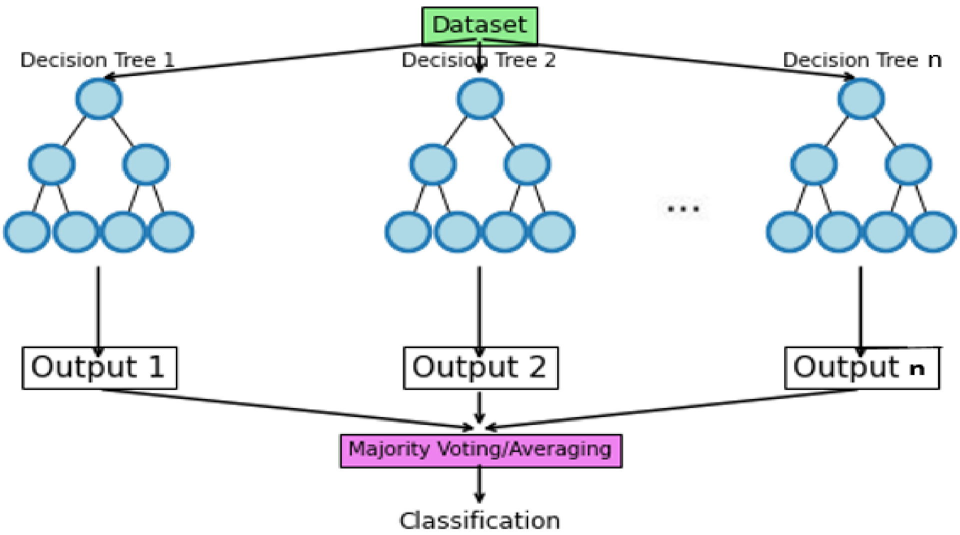

Figure 6: Random Forest working architecture

3. Experimentations

3.1 Model Training and Testing

This research used tomato leaf disease from the public PlantVillage dataset for tomato leaf disease detection. The training set constituted 80% of the total post preprocessing images. In comparison, the testing set included the remaining 20%, where the training dataset contains 18160 sample images, and 4440 images are present in the test dataset.

In the training phase of our methodology, the suggested CNN model is trained with the Adam optimizer, which has been utilized to optimize the training weights, a learning rate of 0.002, a batch size of 16, and the need for around 70 epochs to achieve convergence. We have considered that n-estimators are 200, and the criterion is entropy when training a random forest. The codebase was developed using Python and the TensorFlow libraries. The study employs a system featuring 16 GB of RAM, an Intel Core i7 processor operating at 2.30 GHz, and a dedicated graphics processing unit with 8 GB of RAM.

3.2 Evaluation Metrics

We employed accuracy, precision, recall, and F1-score metrics to evaluate the efficiency of the proposed model. These metrics are calculated as shown:

$$\text{Accuracy}=\frac{\text{TP}+\text{TN}}{\text{TP}+\text{TN}+\text{FP}+\text{FN}} \tag{1}$$

$$\text{Precision}=\frac{\text{TP}}{\text{TP}+\text{FP}} \tag{2}$$

$$\text{Recall}=\frac{\text{TP}}{\text{TP}+\text{FN}} \tag{3}$$

$$\text{F1-Score}=2 \text{ x } \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{4}$$

Where, TP denotes true positive, indicating that the model accurately recognized a positive item as positive. TN denotes true negative, indicating that the model accurately classified a negative instance as negative. FP denotes false positive, indicating that the model erroneously classified a negative item as positive, while FN signifies false negative, meaning the model mistakenly categorized a positive item as negative.

4. RESULTS AND DISCUSSION

This section outlines the experimental assessment of the proposed methodology for evaluating the system's performance. The analysis compares the proposed model's performance with previous works using their approaches to the same public dataset.

4.1 Results of Hybrid Approach(TOMATOCNN-RF)

An ablation investigation involves removing or altering a portion of the architecture to assess how each architecture component represents the proposed methodology. To be more precise, after removing each model, the performance of the TomatoCNN-RF model is assessed. It is essential to assess the stability of the TomatoCNN-RF architecture in this ablation research to ascertain how the Random Forest classifier affects the system’s performance.

We deleted it in these trials to verify the effect of the Random Forest classifier on the model’s performance as a classifier. Table 2 provides an overview of the ablated models’ performance.

Table 2: Ablation study results

Experiment	Accuracy	Precision	Sensitivity	F1-score
TomatoCNN-$RF_{\text{Custom CNN}}$	88%	89%	89%	89%
TomatoCNN-R_{RF}	93%	92%	92%	92%
TomatoCNN-$RF_{\text{CustomCNN+RF}}$	98%	98%	97.7%	98%

According to Table 2, the TomatoCNN-$RF_{\text{CustomCNN+RF}}$ model can enhance accuracy by 10%, accurately detecting diseased and healthy leaves compared with the system using Custom CNN as a classifier. Additionally, it improves accuracy by 5% compared to using Random Forest as the backbone classifier in our approach.

In the context of tomato leaf disease detection, the primary goal is to avoid missing any leaf infected with contagion. Our Model's Precision, at 98%, underscores its ability to reliably detect positive cases and reduce the occurrence of false positives compared with using custom CNN and RF independently as a classifier in our methodology, which produces around 9 % erroneous positive predictions when applied to the test set.

Furthermore, the suggested hybrid model obtained an F1-score value of 98%. A high F1 score signifies that the hybrid model demonstrates excellent precision and recall, accurately identifying positive instances and minimizing false positives and negatives.

Figures 7 and 8 present confusion matrices for the test dataset to visually assess the classification accuracy of the two models in the ablation study. In this two-dimensional matrix, the diagonal values indicate the proportion of accurate forecasts, while the off-diagonal parts denote the fraction of inaccurate predictions. Figure 7 illustrates the confusion matrix, which underscores the superior efficacy of our hybrid model (TOMATOCNN-RFCustom CNN + RF) in distinguishing between classes. It attains precise predictions across most classes, highlighting its robust generalization skills. In comparison to the predictions of the TOMATOCNN-RFCustom CNN model, as illustrated in Figure 8.

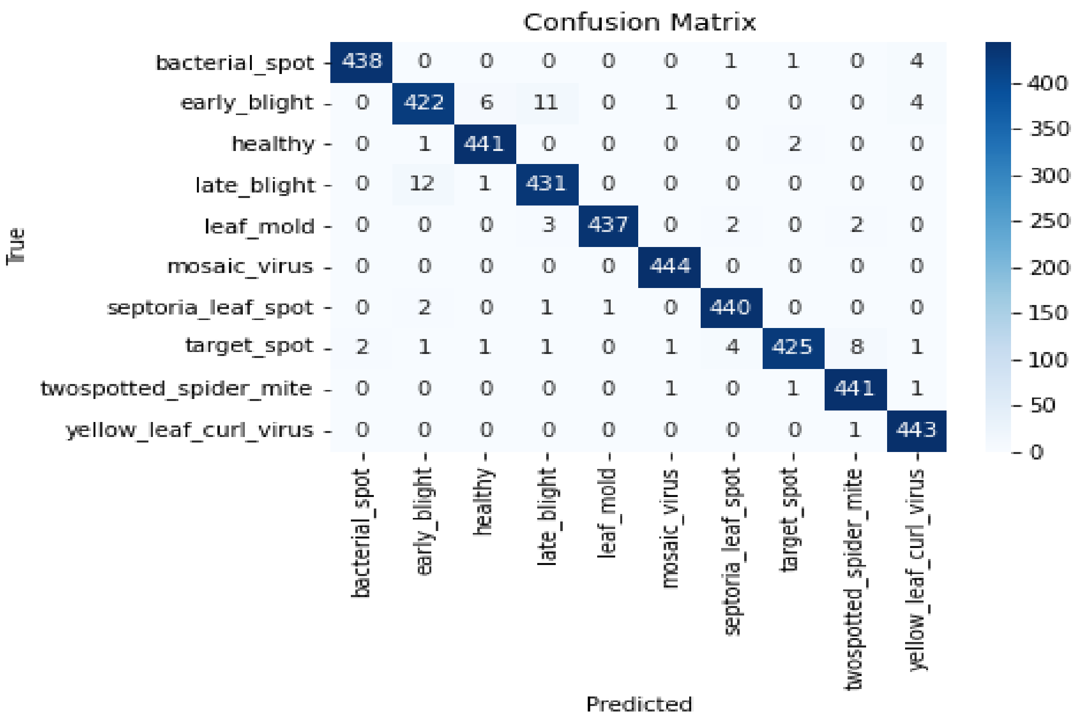

Figure 7: Confusion matrix of the TOMATOCNN-$RF_{Custom\ CNN\ +RF}$ model.

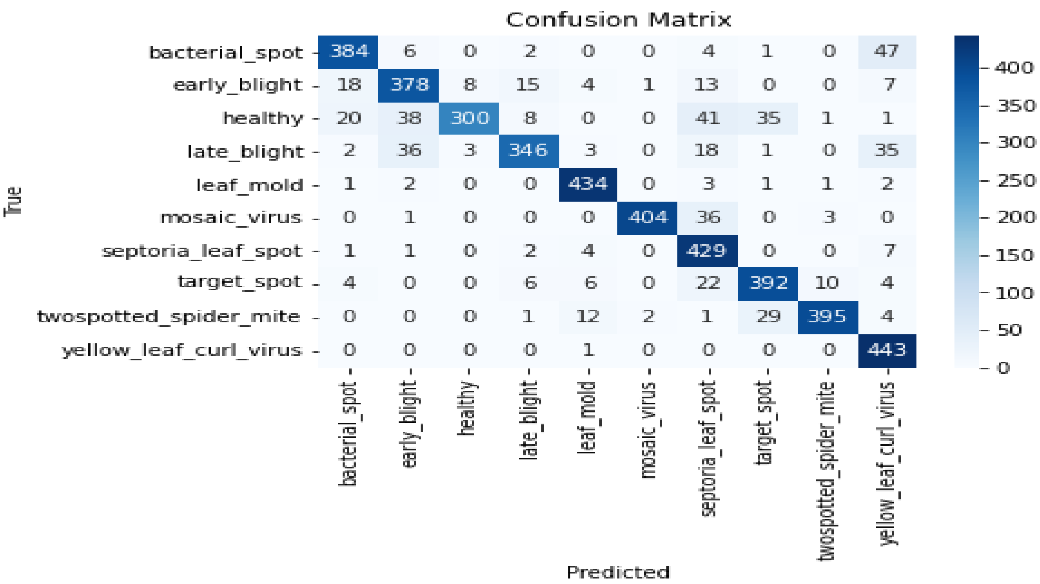

Figure 8: Confusion matrix of the TOMATOCNN-RFCustom CNN model.

Tables 3 and 4 demonstrate class-wise evaluation metrics generated by the two models. Between these, the proposed Random forest-based hybrid model demonstrated the most promising evaluation metrics values, which indicates a higher performance than the model based on a Custom CNN classifier for diagnosing tomato leaf diseases.

Table 3: Class-wise Metrics of TOMATOCNN-RF$_{\text{Custom CNN}}$ model.

Class	Precision%	Recall%	F1-score%	Support
Bacterial spot	89	86	88	444
Early blight	82	85	83	444
healthy	96	68	79	444
Late blight	91	78	84	444
Leaf mold	94	98	96	444
Mosaic virus	99	91	95	444
Septoria leaf_spot	76	97	85	444
Target spot	85	88	87	444
Two spotted spider mite	96	89	93	444
Yellow leaf curl virus	81	100	89	444

Table 4: Class-wise Metrics of TOMATOCNN-RF$_{\text{Custom CNN+RF}}$ model.

Class	Precision%	Recall%	F1-score%	Support
Bacterial spot	100	99	99	444
Early blight	96	95	96	444
Heal thy	98	99	99	444
Late blight	96	97	97	444
Leaf mold	100	98	99	444
Mosac virus	99	100	100	444
Septoria leaf spot	98	99	99	444
Target spot	99	96	97	444
Twospotted spider mite	98	99	98	444
Yellow leaf curl virus	98	100	99	444

The classification results of the TOMATOCNN-RF hybrid model on the test dataset are depicted in Figure 9. The figure shows that our hybrid model correctly classified most classes.

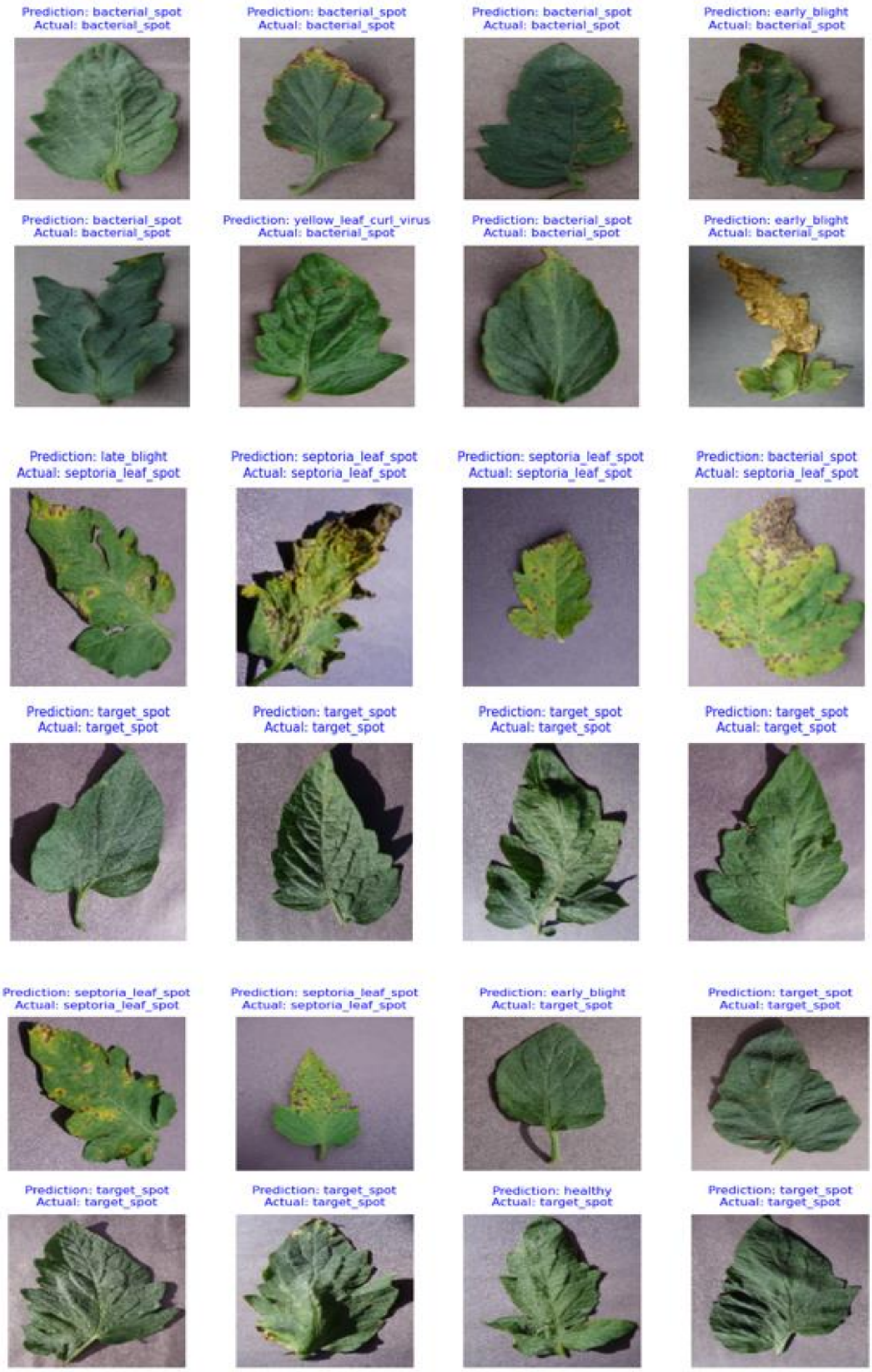

Figure 9: Diagnosis results of the test dataset

4.2 Comparison of the TOMATOCNN-RF model with Previous Works:

To underline the significance and contribution of the work presented in this research, the proposed model is compared to current literature results. Table 5 summarizes the comparison conclusions.

Table 5: Comparison of the TOMATOCNN-RF model with previous studies for diagnosis of tomato leaf diseases using the PlantVillage dataset

Study	Year	Architecture	Accuracy
[14]	2020	YoloV3	92.39%
[15]	2020	Extreme Learning Machine(ELM)	89.19%
[16]	2020	DCGAN with CNN	94.33%
[17]	2021	ELM	84.94%
[18]	2022	VGG	95.71%
[19]	2023	CNN	95.0%
[20]	2023	DIMPCNET	94.4%
[21]	2023	Ensemble Machine learning	95.58%
[22]	2023	Transfer learning	90.0%
TOMATOCNN-RF		Custom CNN + RF	**98.0%**

5. CONCLUSIONS

Our proposed hybrid TOMATOCNN-RF for early and automatic disease diagnosis of tomato crops is an application of artificial intelligence in agriculture, particularly in illness identification and plant health administration. This research employed a custom convolutional neural network with random forests to identify and categorize infections in tomato plants. The model was trained using images taken from the vPlantVillage dataset. The experimental results in the ablation study show that the hybrid TOMATOCNN-RF outperforms the Custom CNN with an accuracy of 98.0%. The proposed approach outperforms various modern studies models, showing its efficiency in accurately identifying tomato diseases.

In the future, we will expand this approach to encompass a wider variety of plants and create a smartphone application to provide farmers with a real-time disease diagnosis tool, enabling them to implement critical measures promptly and cost-effectively.

References

[1] B. Khan *et al.*, "Bayesian optimized multimodal deep hybrid learning approach for tomato leaf disease classification," *Scientific Reports 2024 14:1*, vol. 14, no. 1, pp. 1–30, 2024, doi: 10.1038/s41598-024-72237-x.

[2] S. Panno *et al.*, "A review of the most common and economically important diseases that undermine the cultivation of tomato crop in the mediterranean basin," *Agronomy*, vol. 11, no. 11, pp. 1–45, 2021, doi: 10.3390/agronomy11112188.

[3] M. M. Elsharkawy *et al.*, "Systemic resistance induction of tomato plants against tomato mosaic virus by microalgae," *Egypt J Biol Pest Control*, vol. 32, no. 1, 2022, doi: 10.1186/s41938-022-00538-2.

[4] P. H. Hussan and I. H. Ali, "A Comprehensive Survey on Covid-19 Disease Diagnosis :," vol. 040001, no. October 2022, 2024.

[5] M. Agarwal, A. Singh, S. Arjaria, A. Sinha, and S. Gupta, "ToLeD: Tomato Leaf Disease Detection using Convolution Neural Network," *Procedia Comput Sci*, vol. 167, no. 2019, pp. 293–301, 2020, doi: 10.1016/j.procs.2020.03.225.

[6] Di. Jiang, F. Li, Y. Yang, and S. Yu, “A Tomato Leaf Diseases Classification Method Based on Deep Learning,” *Proceedings of the 32nd Chinese Control and Decision Conference, CCDC 2020*, pp. 1446–1450, 2020, doi: 10.1109/CCDC49329.2020.9164457.

[7] R. G. De Luna, E. P. Dadios, and A. A. Bandala, “Automated Image Capturing System for Deep Learning-based Tomato Plant Leaf Disease Detection and Recognition,” *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, vol. 2018-Octob, no. October, pp. 1414–1419, 2018, doi: 10.1109/TENCON.2018.8650088.

[8] J. Basavaiah and A. Arlene Anthony, “Tomato Leaf Disease Classification using Multiple Feature Extraction Techniques,” *Wirel Pers Commun*, vol. 115, no. 1, pp. 633–651, 2020, doi: 10.1007/s11277-020-07590-x.

[9] G. Paramesh and G. V. S. Narayana, “Tomato Leaf Disease Diagnosis Based On Improved Convolution Neural Network,” *2022 2nd International Conference on Computer Science, Engineering and Applications, ICCSEA 2022*, 2022, doi: 10.1109/ICCSEA54677.2022.9936244.

[10] T. S. Xian and R. Ngadiran, “Plant Diseases Classification using Machine Learning,” *J Phys Conf Ser*, vol. 1962, no. 1, 2021, doi: 10.1088/1742-6596/1962/1/012024.

[11] N. Balakrishna, G. Sunitha, A. Karthik, K. R. Madhavi, and J. Avanija, “Tomato Leaf Disease Detection Using Deep Learning: A CNN Approach,” *2022 International Conference on Data Science, Agents and Artificial Intelligence, ICDSAAI 2022*, vol. 13, no. 1, pp. 305–312, 2022, doi: 10.1109/ICDSAAI55433.2022.10028922.

[12] B. Gülmez, *A Comprehensive Review of Convolutional Neural Networks based Disease Detection Strategies in Potato Agriculture*, no. 0123456789. Springer Netherlands, 2024. doi: 10.1007/s11540-024-09786-1.

[13] N. Horning, “Random Forests: An algorithm for image classification and generation of continuous fields data sets,” *International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences 2010*, pp. 1–6, 2010.

[14] J. Liu and X. Wang, “Tomato Diseases and Pests Detection Based on Improved Yolo V3 Convolutional Neural Network,” *Front Plant Sci*, vol. 11, no. June, pp. 1–12, 2020, doi: 10.3389/fpls.2020.00898.

[15] A. Bhatia, A. Chug, and A. Prakash Singh, “Application of extreme learning machine in plant disease prediction for highly imbalanced dataset,” *Journal of Statistics and Management Systems*, vol. 23, no. 6, pp. 1059–1068, 2020, doi: 10.1080/09720510.2020.1799504.

[16] Q. Wu, Y. Chen, and J. Meng, “Dcgan-based data augmentation for tomato leaf disease identification,” *IEEE Access*, vol. 8, pp. 98716–98728, 2020, doi: 10.1109/ACCESS.2020.2997001.

[17] T. S. Xian and R. Ngadiran, “Plant Diseases Classification using Machine Learning,” *J Phys Conf Ser*, vol. 1962, no. 1, 2021, doi: 10.1088/1742-6596/1962/1/012024.

[18] A. S. Paymode and V. B. Malode, “Transfer Learning for Multi-Crop Leaf Disease Image Classification using Convolutional Neural Network VGG,” *Artificial Intelligence in Agriculture*, vol. 6, pp. 23–33, Jan. 2022, doi: 10.1016/J.AIIA.2021.12.002.

[19] S. G. Paul *et al.*, “A real-time application-based convolutional neural network approach for tomato leaf disease classification,” *Array*, vol. 19, no. August, p. 100313, 2023, doi: 10.1016/j.array.2023.100313.

[20] D. Peng, W. Li, H. Zhao, G. Zhou, and C. Cai, “Recognition of Tomato Leaf Diseases Based on DIMPCNET,” *Agronomy*, vol. 13, no. 7, 2023, doi: 10.3390/agronomy13071812.

[21] S. M. Javidan, A. Banakar, K. A. Vakilian, and Y. Ampatzidis, “Tomato Leaf Diseases Classification using Image Processing and Weighted Ensemble Learning,” *Agron J*, 2023, [Online]. Available: https://api.semanticscholar.org/CorpusID:256384889

[22] S. Lingwal, K. K. Bhatia, and M. Singh, “Deep Convolutional Neural Network Approach for Tomato Leaf Disease Classification,” in *Machine Learning, Image Processing, Network Security and Data Sciences*, R. Doriya, B. Soni, A. Shukla, and X.-Z. Gao, Eds., Singapore: Springer Nature Singapore, 2023, pp. 199–208.