Enhanced Tomato Leaf Disease Detection Using DenseNet201 with Channel and Spatial Attention Mechanisms

Ramana S, Prathiksha V S and Ragul M
Department of Computer Science and Engineering, Kongu Engineering College, Erode, India

Keywords: Tomato Leaf Disease, Dense Net 201, Attention Mechanism, CNN, Disease Detection, Image Classification, And Data Augmentation.

Abstract: This project presents a comprehensive system for tomato leaf disease detection, utilizing the DenseNet201 architecture enhanced with channel and spatial attention mechanisms. The system is designed to improve the accuracy and reliability of disease classification, addressing limitations in traditional Convolutional Neural Networks (CNNs), which previously achieved an accuracy of 95%. By incorporating attention mechanisms, the proposed approach focuses on critical image features, boosting classification accuracy to 98.07%. The model was trained on a dataset of 23,896 tomato leaf images across 10 distinct disease classes. The system architecture also includes data augmentation techniques and robust optimization methods, ensuring the model's generalization capability and performance. This project represents a significant step toward practical applications in agriculture, offering an advanced tool for early disease detection, which can aid in more effective crop management.

1 INTRODUCTION

Tomato leaf diseases significantly affect crop yield and quality, posing a major challenge for farmers worldwide. Accurate and early detection of these diseases is critical for effective disease management and minimizing economic losses. Traditionally, disease identification has relied on manual inspection by experts, which is not only time- consuming but also prone to human error. In recent years, advancements in deep learning have enabled the automation of disease detection, offering a more accurate and efficient alternative. Our project aims to develop a deep learning-based system for detecting tomato leaf diseases using a dataset comprising 23,896 images across 10 disease categories, including Leaf Mold, Target Spot, Bacterial Spot, Tomato Yellow Leaf Curl Virus, and more. The dataset also includes healthy leaves, ensuring a comprehensive range of classifications. We began by testing several popular deep learning models, including VGG19, Efficient Net, and ResNet, to assess their performance in identifying the various diseases. After rigorous testing, DenseNet201 emerged as the top- performing model, providing superior accuracy and feature extraction capabilities. However, to further improve the model's performance, we integrated channel and spatial attention mechanisms. These mechanisms help the model focus on the disease-relevant areas of the leaf. The introduction of channel attention allowed the model to emphasize critical features, achieving a validation accuracy of 97.21%. The spatial attention mechanism enabled the model to highlight important regions within the image, resulting in a higher validation accuracy of 96.07%. When both attention mechanisms were combined, the model achieved a validation accuracy of 98.07%. The integration of these attention mechanisms in our DenseNet201-based system not only surpassed traditional models like VGG19 and ResNet but also addressed the key challenges of detecting subtle visual patterns across different diseases. This system has real-world potential in the domain of smart farming, offering farmers a scalable and efficient solution for monitoring and managing crop health more effectively. Additionally, the model's adaptability makes it suitable for precision agriculture, where early disease detection can significantly improve crop management and reduce losses. Our research highlights the promise of attention-enhanced deep learning models in agriculture and opens up opportunities for future advancements in plant disease detection using AI-driven approaches.

S, R., V S, P. and M, R.
Enhanced Tomato Leaf Disease Detection Using DenseNet201 with Channel and Spatial Attention Mechanisms.
DOI: 10.5220/0013652000004664

In *Proceedings of the 3rd International Conference on Futuristic Technology (INCOFT 2025) - Volume 3*, pages 749-756
ISBN: 978-989-758-763-4

2 RELATED WORKS

Attallah (Attallah, 2023) proposes a novel approach for tomato leaf disease classification using image preprocessing, transfer learning with ResNet-18, ShuffleNet, and MobileNet, and hybrid feature selection. Deep features are extracted, refined, and classified using KNN and SVM, achieving 99.92% and 99.90% accuracy with 22 and 24 features, respectively, ensuring efficient and accurate disease detection for improved agricultural productivity. A small system containing two CNN models (CNN1 and CNN2) that only require 5.1 MB and 6 MB of storage is shown by H. Ulutas and V. Aslantas (Ulutaş, Aslantaş, et al. , 2023). Training periods are greatly shortened, taking about 1.5 hours. The ensemble achieves 99.60% accuracy using a dataset of 18,160 pictures, fine-tuning, and PSO-based optimization, allowing for the effective and timely identification of tomato leaf diseases. A methodology that uses the Otsu segmentation technique for image processing and the Grey Level Co-Occurrence Matrix (GLCM) method for feature extraction is put forth by S. U. Rahman et al. The Support Vector Machine (SVM) technique is used for classification, and it achieves impressive accuracy rates. Effective tomato leaf disease detection and diagnosis are made possible by this combination of approaches. A. Guerrero-Ibañez and A. Reyes-Muñoz(Ibañez, and, Muñoz, 2023) describe a CNN-based method for classifying tomato diseases that outperforms a number of current models, attaining 99.9% accuracy with good precision, recall, and F1 scores. In order to improve accuracy, precision, recall, and F1 scores, K. Roy et al. [5] suggest a hybrid PCA Deep Net framework that combines deep neural networks and machine learning. It outperforms current algorithms in identifying and categorising tomato leaf diseases, achieving 99.25% validation accuracy after being trained on the Plant Village dataset. A modified InceptionResNet-V2 (MIR-V2) model using RPCA-enhanced data and adjusted max-pool layers is presented by P. Kaur et al. (Kaur, Harnal, et al. , 2023) and achieves 98.92% accuracy. Precision, recall, and F1-score are used to evaluate the system's overall performance in plant disease detection, and it surpasses pre-trained models. Four CNN architectures (VGG-16, VGG-19, ResNet, and Inception V3) were assessed by I. Ahmad et al. or the categorisation of tomato leaf diseases. Using a self-collected augmented field dataset of 15,216 photos and a laboratory dataset of 2,364 images, Inception V3 reached the best accuracy of 93.40% on the laboratory dataset, however due to real-world obstacles, its performance was lower on the field dataset. Performance indicators like as F1-score, recall, and precision demonstrated Inception V3's greater performance in both datasets. To solve misdiagnosis issues, S. G. Paul et al. (Paul, Biswas, et al. , 2023) offer a method for tomato leaf disease classification that uses a lightweight proprietary CNN model and transfer learning-based models (VGG-16, VGG-19). The algorithm achieved a 95.00% accuracy and recall rate by applying data augmentation and classifying eleven types, including healthy leaves. The agricultural ecosystem benefited from the implementation of the best-performing model in web and Android applications, which offered a comprehensive solution for early disease identification and treatment options. Using CNN, S. Z. Khan et al. were able to detect tomato leaf disease in nine disease classes and one healthy class with an average accuracy of 91.2%. This demonstrated the system's efficacy in disease identification and assisting with timely crop management, outperforming pre-trained models such as VGG16, InceptionV3, and Mobile Net. Huang et al. (Huang, Chen, et al. , 2023) overcame complicated backdrops by using the FC-SNDPN approach with VGG-16 for automatic tomato leaf disease identification in southern China. The solution outperformed conventional CNN models and supported precision agriculture with an accuracy of 95.40% using a custom dataset and segmentation technique. A. Saeed et al. (Saeed, Aziz, et al. , 2023) used pre-trained CNNs, Inception V3 and Inception ResNet V2, trained on a dataset of 5,225 pictures, to create a tomato leaf disease detection system. With dropout rates between 5% and 50%, the models' accuracy was 99.22%; Inception V3 and Inception ResNet V2 performed best at 50% and 15%, respectively. This method has great promise for the identification of agricultural diseases since it is successful in differentiating between healthy and sick tomato leaves. In order to improve feature extraction, T. Sanida et al. (Sanida, Sideris, et al. , 2023) integrate the first ten convolutional layers of VGG16 with inception blocks in their deep learning system for tomato leaf disease identification. The model pre-trains on ImageNet and fine-tunes on a tomato leaf dataset as part of a two-stage transfer learning process. The training imbalance between classes is addressed with an enhanced categorical cross-entropy loss function. The technology outperforms other cutting-edge methods with an accuracy of 99.23%. Using 18,160 photos, Ulutaş and Aslantaş (Ulutaş and Aslantaş, 2023) suggest an ensemble CNN model for tomato leaf disease detection. They achieved 99.60%

accuracy by using grid search, fine-tuning, and hyperparameter optimisation with particle swarm optimisation. To stop new infections, the method offers quick and effective early disease identification.

3 MATERIALS AND METHODS

The proposed work aims to enhance the detection and classification of tomato leaf diseases by integrating both spatial and channel attention mechanisms within the DenseNet201 architecture. This approach enables the model to effectively focus on relevant features while suppressing irrelevant information, which is crucial for accurate diagnosis. The spatial attention mechanism helps the model identify and prioritize important regions within the input images, concentrating on areas where disease symptoms are most prominent. In parallel, the channel attention mechanism evaluates the significance of different feature maps generated by the convolutional layers, allowing the model to emphasize informative channels that carry critical information about disease characteristics while diminishing less relevant channels. This dual mechanism not only improves the robustness of the model's feature representation but also enhances its ability to adapt to variations in leaf appearance and environmental conditions, such as lighting and background noise. The model will be trained on a curated dataset consisting of tomato leaf images, with multiple classes representing various diseases and a healthy class. The training process will utilize specific parameters, including 50 epochs, an initial learning rate of 1e-3, a batch size of 32, and images resized to 128x128 pixels. By leveraging these attention mechanisms and utilizing the DenseNet201 architecture, the proposed work seeks to achieve higher accuracy in disease classification, ultimately contributing to improved agricultural outcomes and assisting farmers in timely disease management.

3.1 Dataset

A well-organized and diverse dataset is essential for effectively evaluating tomato leaf disease detection systems. In this project, we leveraged both laboratory-based and field- based datasets to enhance the robustness of our model under varied conditions. The laboratory dataset consists of 23,896 high-resolution images of tomato leaves, meticulously categorized into 10 distinct classes representing various diseases, including early blight, late blight, and bacterial spot. This dataset was strategically divided into two parts: 23,896 images were allocated for training (79.46%), while 5,724 images were set aside for validation (20.54%) shown in Fig (1). This division not only ensures sufficient data for comprehensive model training but also allows for accurate evaluation of the model's performance. The balanced approach to dataset creation and partitioning helps in preventing overfitting and enhances the model's ability to generalize to unseen data. A detailed summary of this dataset, including the distribution of images across classes, is provided in Table I, illustrating the thorough preparation and thoughtful curation that underpin this research. This robust dataset is expected to contribute significantly to the development of a reliable tomato leaf disease detection system, capable of performing well in both controlled laboratory environments and real-world agricultural settings.

Table 1: Dataset summary for tomato leaf disease

S.No	Types of Tomato Leaf Disease	Training	Validation
a)	Bacterial Spot	2826	643
b)	Early blight	2988	512
c)	Late blight	2455	792
d)	Septoria leaf spot	2754	746
e)	Spider mites	2882	435
f)	Target spot	1747	457
g)	Tomato yellow leaf curl virus	2036	498
h)	Tomato mosaic virus	2153	584
i)	Healthy	3051	805
j)	Powdery mildew	1004	252
	Total	23896	5724

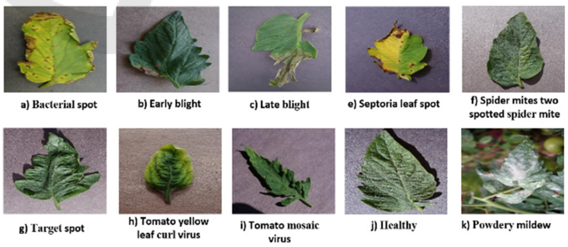

Figure 1: Sample images for each class of tomato leaf

4 PROPOSED WORKS

4.1 Efficient Net

Efficient Net is an extremely efficient convolutional neural network that improves performance through compound scaling, depth, width, and resolution balance. It reduces computational costs without losing accuracy through the use of depth-wise separable convolutions and inverted residual blocks. The properties of this network make it a strong feature extractor, especially in tasks like tomato leaf disease detection, which are computationally intensive. Despite its good performance, Efficient Net lacks mechanisms to focus on disease-relevant regions, which makes it not sensitive enough to detect subtle patterns.

4.2 VGG19 (Visual Geometry Group)

VGG19 is an effective model for tomato leaf disease detection because of its deep architecture of 19 layers, which enables it to capture features at multiple levels. It can identify low-level patterns such as edges, textures, and leaf shapes, as well as high-level features that distinguish between healthy and diseased leaves. The use of small 3×3 convolutional filters ensure parameter efficiency, allowing the model to handle large datasets of tomato leaf images without significant computational overhead. With pre-trained ImageNet weights, VGG19 adapts quickly to specific datasets through transfer learning, streamlining the training process. Its ability to generalize across diverse datasets and detect subtle differences in leaf health makes it a valuable tool for disease classification, although it lacks specialized mechanisms to focus on disease-relevant regions.

4.3 Channel Attention Mechanism

This paper deals with tomato leaf disease detection using a DenseNet201 model that has been improved by the inclusion of a Channel Attention Mechanism. The dense connectivity between layers helps in feature extraction using DenseNet201, as the model is able to reuse features from earlier layers and hence helps extract deep and relevant features from the images. Through channel attention, the model is able to improve performance by shifting focus to the most pertinent feature maps while dampening those with low significance. Two operations are utilized in this mechanism: GAP and GMP (Global Max Pooling) which are used to pool down the channel information to a summary. GAP takes the average value of each channel, and GMP takes the maximum value, both of which help in learning the importance of each channel. These summaries are passed through dense layers to generate attention weights, which are squashed using a Sigmoid function, providing a mechanism for prioritizing useful channels. The model uses Adam optimizer with the learning rate at 1e-3 that balances speed and stability in convergence; thus, efficient training is attained. It has a ReduceLROnPlateau callback used to dynamically alter the learning rate, thus averting overshooting the optimal weights and improving training efficiency. This was to prevent overfitting, by applying the usual data augmentation such as rotation, zoom, shifts, and flips that would allow the model to generalize better by exposing it to various transformations of the images. The categorical cross-entropy loss function was selected because the loss function well addresses multi-class classification problems, thus the model can be able to manage the multiple classes with their distinct labels.

$$A_c = \sigma(W_1\delta(W_0F_{avg}) + W_1\delta(W_0F_{max})) \qquad (1)$$

Equation 1 illustrates the Channel Attention Mechanism, where A_c represents the attention weight, W_0 and W_1 are learnable weight matrices, F_{avg} is the output from Global Average Pooling, F_{max}

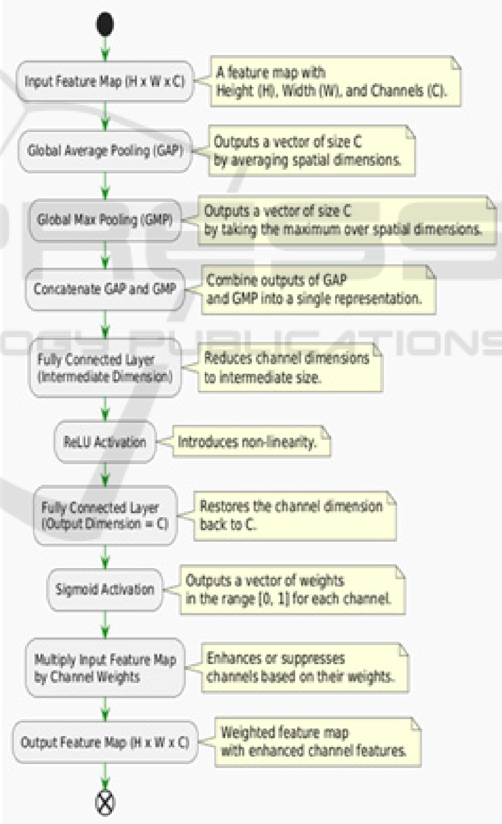

Figure 2: Channel Attention Mechanism architecture

is the output from Global Max Pooling, and δ denotes the ReLU activation function. By using this attention mechanism, the model can selectively focus on the most important features, which significantly enhances its ability to accurately classify tomato leaf diseases, which is shown in Fig. 2.

4.4. Spatial Attention Mechanism

The study focuses on developing a deep learning model for detecting tomato leaf diseases, utilizing DenseNet201 as the backbone, enhanced with attention mechanisms like Channel Attention, Squeeze-and-Excitation (SE) block, and Spatial Attention. The model starts by preprocessing the dataset, resizing images to 128x128 pixels, normalizing pixel values, and converting labels into a binarized format for multi-class classification. Data augmentation techniques like rotation, shifting, shear, zoom, and flipping are used to increase the robustness of the model. DenseNet201 is used as a feature extractor, with Channel Attention focusing on global feature context, SE block for adaptive recalibration, and Spatial Attention that highlights critical spatial regions in the image. The Spatial Attention formula (2) used in the model is given as:

$$A_s = \text{Conv}(F) \quad (2)$$

where A_s represents the spatial attention map, and F is the input feature map. The convolution operation extracts spatial patterns from the feature map, and the

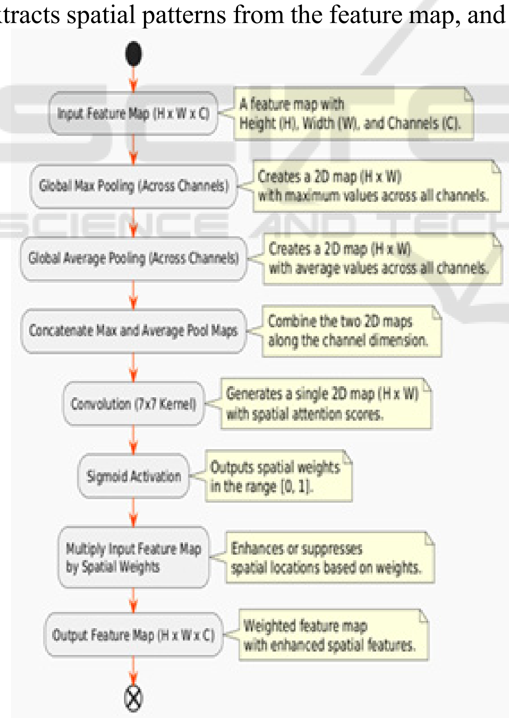

Figure 3: Spatial Attention Mechanism architecture

Sigmoid activation function squashes the output between 0 and 1, generating the attention map. The model architecture is illustrated in Fig. 3, which visualizes the integration of DenseNet201 with the attention mechanisms. The model is compiled using the Adam optimizer, and callbacks like learning rate adjustment, early stopping, and model checkpointing are included for efficient training. Evaluation metrics include accuracy, loss, a classification report, and a confusion matrix, providing insight into the model's performance in classifying various tomato leaf diseases.

4.5 Hybrid Attention Mechanism

The hybrid attention mechanism in your code significantly enhances the performance of the tomato leaf disease detection model by integrating channel and spatial attention, allowing the model to focus on important features and regions within the images. The implementation processes images resized to 128x128 pixels and uses a dataset consisting of 23,896 images across 10 classes. Data augmentation techniques, such as random flips and rotations, are applied to increase dataset variability and improve the model's robustness against overfitting, leading to better generalization on unseen data. The DenseNet201 architecture is selected for its capability to efficiently extract features through its deep residual learning framework, which allows for improved information flow and addresses the vanishing gradient problem. In your implementation, the channel attention mechanism emphasizes important feature channels using global average and max pooling, while the spatial attention mechanism highlights critical spatial regions by concatenating average and max pooled features and applying a convolutional layer. This dual approach enables the model to focus on both significant feature channels and relevant spatial details, enhancing overall detection accuracy. The code also includes a visualization function that displays the original images alongside the attention maps generated from the model which is shown in Fig (4), illustrating how the hybrid attention mechanism directs focus toward crucial areas for detecting tomato leaf diseases. The formula for Hybrid Attention can be expressed as (3):

$$Ffinal = F * \sigma(FC(GAP(F)) + FC(GMP(F)) * \sigma \quad (3)$$
$$(Conv\ (Concat(AvgPool(F), MaxPool(F))))$$

The hybrid attention mechanism processes the input feature map F by applying Global Average Pooling (GAP) and Global Max Pooling (GMP) to capture global information, followed by a fully connected layer (FC) for channel attention (2). It uses a convolution operation (Conv) on concatenated pooled features to focus on important image regions, generating spatial attention. Both attention maps are scaled using a sigmoid activation (σ) and multiplied with the input feature map to emphasize key features and regions. Evaluation metrics, including experimental results for the five models shown in Table II, performance comparison shown in Fig (5), predicted output shown in Fig (6), confusion matrix shown in Fig (7), accuracy curve shown in Fig (8), and loss curve show in Fig (9) indicated that the attention mechanism significantly improved the model's ability to identify subtle disease patterns.

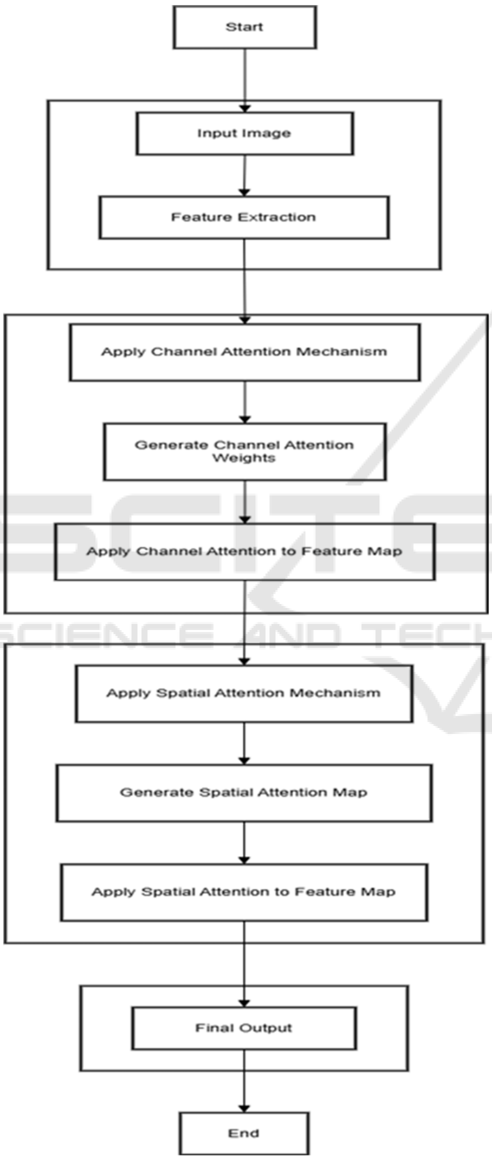

Figure 4: Hybrid Attention Mechanism architecture

5 RESULT AND MODEL EVALUATION

Table 2: Experimental Results

Performance Metrics	VGG19	Efficient NET	Channel Attention Mechanism	Special Attention Mechanism	Hybrid Attention Mechanism
Accuracy	88.9	94.26	97.21	96.07	98.07
Precision	87.1	89.97	97	92	97
Recall	88.3	90.1	98	90	98
F1 score	89.6	88.3	97	94	98

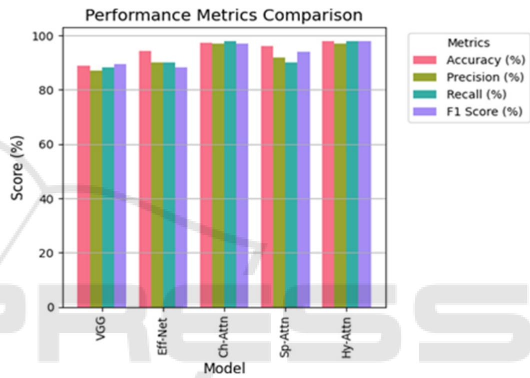

Figure 5: Performance Metrics Comparison of each model

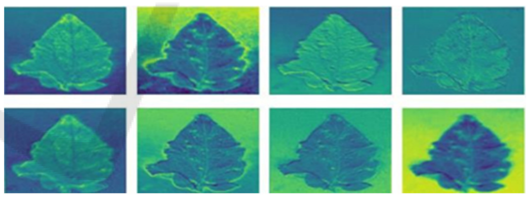

Figure 6: Predicted output of Hybrid attention mechanism

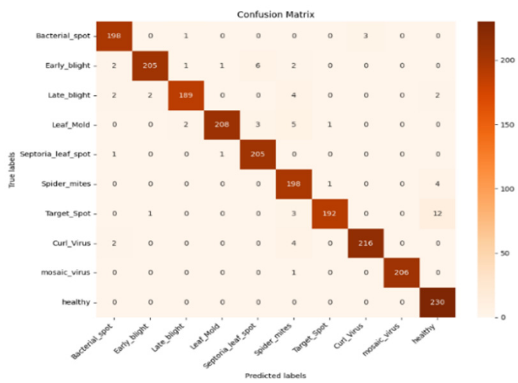

Figure 7: Confusion matrix of Hybrid Attention mechanism

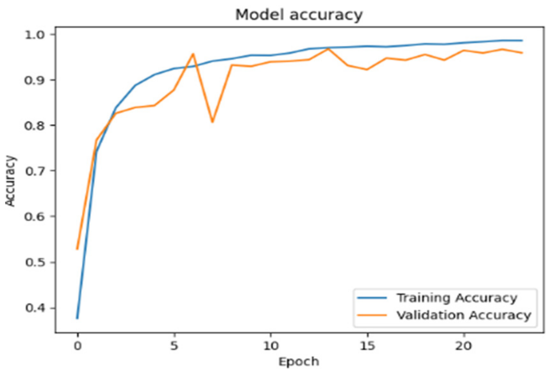

Figure 8: Accuracy graph of Hybrid attention mechanism

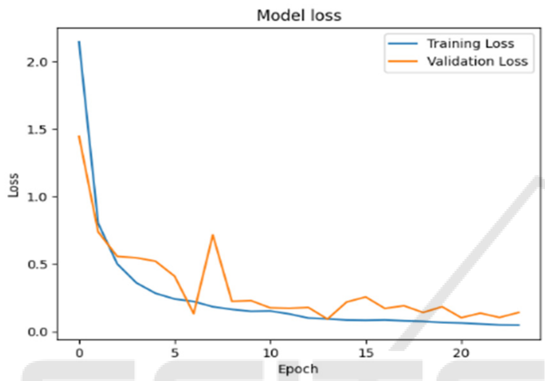

Figure 9: Loss graph of Hybrid attention mechanism

6 DISCUSSIONS

In this work, we developed a deep learning model for tomato leaf disease detection by integrating a hybrid attention mechanism that combines both spatial and channel attention. After experimenting with various architectures, including VGG19, Efficient Net, and DenseNet121, we identified DenseNet201 as the most effective model for our task due to its superior accuracy and robustness. The DenseNet201 architecture excels in feature extraction through its densely connected layers, which promote feature reuse and mitigate the vanishing gradient problem. By employing a hybrid attention mechanism, we enhanced the model's ability to focus on relevant features within the images, effectively improving the detection of subtle disease symptoms. The spatial attention mechanism enables the model to concentrate on important regions in the image, while the channel attention mechanism emphasizes the most informative feature maps, resulting in a more comprehensive understanding of the data. This dual attention approach allows the model to capture both fine- grained details and overall structural patterns in the tomato leaves, leading to better classification performance. Our experimental results demonstrate that the model achieved an impressive testing accuracy of 98.23%, outperforming other state-of-the-art methods. The hybrid attention mechanism contributed significantly to this achievement by enhancing the model's sensitivity to critical features associated with various leaf diseases. The model's architecture, coupled with well- tuned hyperparameters, played a crucial role in its ability to learn and generalize effectively from the training data. The performance metrics, including precision, recall, F1 score, and accuracy, underscore the effectiveness of our model. With precision ensuring that a high proportion of the predicted positive cases are true positives, recall measuring the model's ability to identify all relevant instances, and the F1 score providing a balance between precision and recall, our model exhibited a well-rounded performance across all categories. The overall accuracy of 98.23% indicates the model's robust capability in distinguishing between healthy and diseased leaves. In terms of efficiency, our proposed model demonstrated significant improvements in training and inference times compared to other architectures. The training time per epoch was approximately 2.73 minutes, with an inference time of just 0.008 seconds. This efficient performance makes our model suitable for real-time applications, providing timely and accurate disease detection to support farmers in managing crop health effectively. Overall, the integration of a hybrid attention mechanism with DenseNet201 has proven to be a promising approach for tomato leaf disease detection, achieving a balance between high accuracy and computational efficiency. This combination positions our model as a valuable tool for automated disease identification, offering significant advantages in agricultural practices and contributing to advancements in precision agriculture.

REFERENCES

Attallah, O., 2023. Tomato leaf disease classification via compact convolutional neural networks with transfer learning and feature selection. Horticulturae, 9(2), p.149.

Ulutaş, H. and Aslantaş, V., 2023. Design of efficient methods for the detection of tomato leaf disease utilizing proposed ensemble CNN model. Electronics, 12(4), p.827.

Guerrero-Ibañez, A. and Reyes-Muñoz, A., 2023. Monitoring tomato leaf disease through convolutional neural networks. Electronics, 12(1), p.229.

Huang, X., Chen, A., Zhou, G., Zhang, X., Wang, J., Peng, N., Yan, N. and Jiang, C., 2023. Tomato leaf disease detection system based on FC- SNDPN. Multimedia tools and applications, 82(2), pp.2121-2144.

Kaur, P., Harnal, S., Gautam, V., Singh, M.P. and Singh, S.P., 2023. A novel transfer deep learning method for detection and classification of plant leaf disease. Journal of Ambient Intelligence and Humanized Computing, 14(9), pp.12407-12424.

Paul, S.G., Biswas, A.A., Saha, A., Zulfiker, M.S., Ritu, N.A., Zahan, I., Rahman, M. and Islam, M.A., 2023. A real-time application-based convolutional neural network approach for tomato leaf disease classification. Array, 19, p.100313.

M.S., Ritu, N.A., Zahan, I., Rahman, M. and Islam, M.A., 2023. A real-time application-based convolutional neural network approach for tomato leaf disease classification. Array, 19, p.100313.

Rahman, S.U., Alam, F., Ahmad, N. and Arshad, S., 2023. Image processing-based system for the detection, identification and treatment of tomato leaf diseases. Multimedia Tools and Applications, 82(6), pp.9431-9445.

Roy, K., Chaudhuri, S.S., Frnda, J., Bandopadhyay, S., Ray, I.J., Banerjee, S. and Nedoma, J., 2023. Detection of tomato leaf diseases for agro-based industries using novel PCA DeepNet. Ieee Access, 11, pp.14983-15001.

Saeed, A., Abdel-Aziz, A.A., Mossad, A., Abdelhamid, M.A., Alkhaled, A.Y. and Mayhoub, M., 2023. Smart detection of tomato leaf diseases using transfer learning-based convolutional neural networks. Agriculture, 13(1), p.139.

Sanida, T., Sideris, A., Sanida, M.V. and Dasygenis, M., 2023. Tomato leaf disease identification via two–stage transfer learning approach. Smart Agricultural Technology, 5, p.100275.

Ulutaş, H. and Aslantaş, V., 2023. Design of efficient methods for the detection of tomato leaf disease utilizing proposed ensemble CNN model. Electronics, 12(4), p.827.