IJTRET "Voice of Innovative Minds"
doi.org/10.54473/IJTRET.2024.8103
International Journal of Trendy Research in Engineering and Technology
Volume 8 Issue 1 February 2024
ISSN No.2582-0958

PLANTATION DISEASE DETECTION AND DIAGNOSIS

Hitesh Kumar, Kunal, Shreya Jadon, Saksham Sharma , Mrs. Upasana Tiwari
Apex Institute of Technology
Chandigarh University Mohali, India
corresponding author: hitesh70001@gmail.com,

Received 04 March 2024 Received in revised form 06 March 2024 Accepted 07 March 2024

ABSTRACT

Deep learning, as a branch of AI , has got into the limelight in recent years due to its automatic learning and feature extraction capabilities. It has found widespread applications in various domains, including multimedia processing, voice analysis, and NLP. Moreover, it has emerged as a promising area of research in agricultural plant protection, specifically in tasks like plant disease recognition and weed/bacterial assessment. Leveraging neural network learning for plant disease recognition mitigates the drawbacks associated with manual process of choosing of disease marks attributes and features, processing disease feature selection and extraction more objectively and accelerating research productivity and technological advancements.

This research paper encapsulates the recent advancements in DL technology concerning the identification of plantation diseases. The paper highlights current tech and their complexities in utilizing neural networks and optimized image processing techniques for plantation disease detection. We aim for this work to serve as an effective reference for researchers involved in the study of plant. Additionally, we delve into the existing challenges and issues that warrant resolution.

Keywords — CovNets, Deep learning, Diagnosis, Disease Detection, Machine Learning, Plant disease, Sustainable Agriculture.

I. Introduction

P. Srivastava et. Al.[1] authored the importance of early detection of various disease present in plants. Agricultural production is adversely affected by the presence of plant diseases. If these diseases are not diagnosed early, food insecurity can be severe. Early detection is the cornerstone of powerful plantation sickness prevention and manipulate, and plays an critical function in agricultural manufacturing management and selection-making In current years, the introduction of plant illnesses has shown that a main difficulty.

Infected plants often exhibit visible symptoms or lesion on leaf, stem, petal , or fruits. Each disease usually appears as a specific visual pattern, which can be used for a different diagnostic purpose. Plant leaves are often the main resource for the diagnosis of plant deformities, as many disease symptoms appear initially in the foliage.[2]

Often, agricultural professionals train rely on on-site surveys, or farmers use their own experience to diagnose diseases and pests affecting fruit trees This approach is not only subjective but it is time consuming and not efficient.

Inexperienced farmers are prone to misdiagnosing plant diseases, often resorting to the indiscriminate use of chemical treatments. This can lead to reduced product quality, lower yields, and environmental contamination, resulting in avoidable financial losses. To address these topics, the investigation of pixel analysis and processing methods that caters identification of plant diseases has gained significant traction in research, emerging as a prominent and pressing area of study.

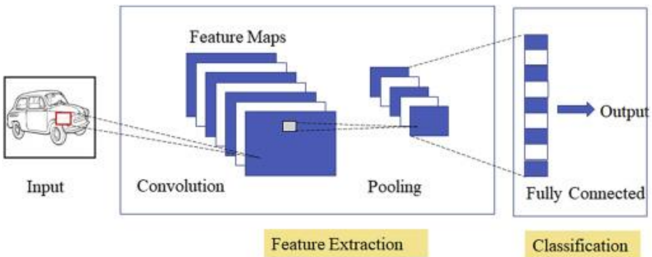

Figure 1 : Image processing using CovNets

In the work authored by Kiran, Ajmeera et. al.[3], modern research on plant diseases using Convolutional Neural Networks (CNNs) was extensively reviewed., highlighting on feature extraction flow based on either techniques made from scratch or deep learning. Their findings concluded that DL techniques have outperformed surface-level classifiers utilizing manually made features. However, their review omitted recent advancements in visualization techniques and did not address the critical aspects of early disease identification and the classification of crop diseases based on limited sample sizes.

This paper aims to address the gaps in existing researches in the field of plantation disease prediction. It provides a state-of-the-art alike of new discoveries in plantation disease diagnosis, including dealing with images, multispectral imaging and neural nets. Our objective is to offer a valuable study for researchers engaged in the area of crop/plantation leaves disease diagnosis through the application of neural networks.

II. LITERATURE REVIEW

A. *Conventional Methods*

The conventional process for employing image detection methods to identify crop diseases is depicted in Figure 2. V Singh et. al.[4] researched various works and few of which employed the centroid based clustering technique to segment lesion regions. They integrated various features, to unfold both RGB value and consistency characteristics from apple lesions. Subsequently, they utilized an improved Support Vector Classifiers to identify and classify 3 classes of apple diseases, gaining a classification accuracy of 93 percent [5].

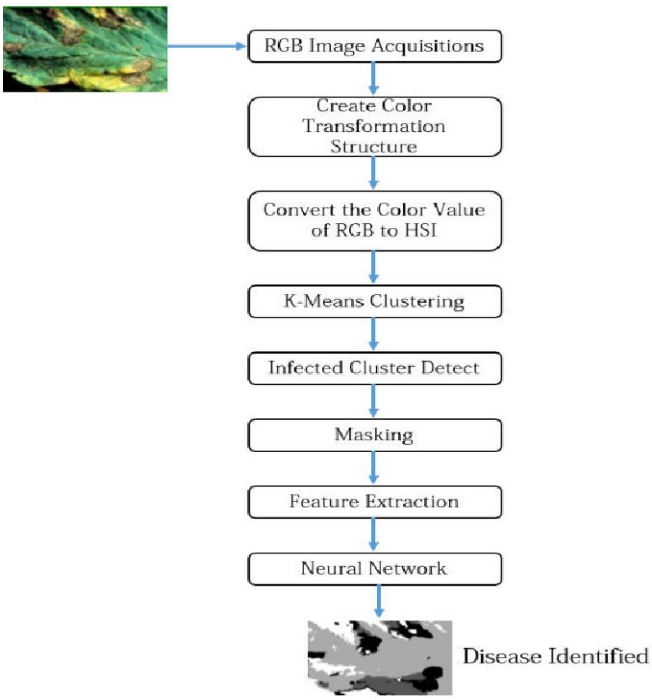

Figure 2 : Classification by Dubey & Jalal

Chai et al.[6] conducted a study which was discussed by S. Zhang, B Wang et. al.[7,7a] that was focused upon 4 tomato foliage diseases. They derived eighteen distinctive parameters involving RGB value, consistency , and structure attributes from tomato leaf spot images. This extraction process utilized PCA and Bayesian Belief Networks . PCA methods were applied to derive these attributes and establish the LDA model. Remarkably, The precision of the two approaches achieved 0.947 and 0.983, respectively.

Chai et al.[6] conducted an analysis where they extracted a total of 63 parameters encompassing morphological, RGB value, and consistency features from rice leaf disease marks. They applied LDA and the Bayesian discriminant method to successfully categorize and identify 3 distinct rice diseases (bacterial leaf blight, blast and stripe blight), achieving an impressive recognition accuracy of 0.942. Here consistency is referred as texture of datapoint.

In summary, it can be deduced that research centered on plantation disease diagnosis using standard techniques has produced commendable outcomes, achieving a high level of accuracy in disease identification. However, there are evident shortcomings and cons, including:

1) The research processes involve intricate and subjective steps, proving time-taking and labor-intensive.

2) The method mainly relies on the segments of disease spots.

3) Its significantly relies on manual feature extraction.

4) Testing the disease detection evaluation of models or algorithms in more complex environments presents challenges.

Therefore, realizing intelligent, fast, and precise recognition of plant leaf diseases is of utmost importance. In past years, there has been notable progress in the application of neural networks technology to plant disease recognition. Deep learning (DL) technology offers transparency to users and reduces the necessity for high expertise in plant preventions and stats. It automatically extracts image attributes and classifies crop disease spots, eliminating the arduous task of manual feature extraction and classifier design associated with traditional image recognition technology. Moreover, DL can capture the intrinsic characteristics of original images and features end-to-end. These unique attributes have brought significant attention to neural net techniques in the field of crop or plantation disease recognition, making it a prominent and widely researched topic. This surge in interest can be attributed to three key aspects: the presence of extensive datasets, the capability and adaptiveness of multi-core graphics hardware (GPUs) , and advancements in training CNN and supporting software modules, such as NVIDIA's CUDA.

B. Data gathering and available datasets

Several generic data sets are available for plant disease diagnosis research:

1) The open dataset known as "PlantVillage" comprises a vast collection of 54,309 images of plant leaves afflicted with diseases. It had fourteen types of fruits and vegetables. The dataset encompasses 26 distinct ailments, comprising various fungal diseases, four bacterial diseases, two mycoses, two bacterial infections, and one mildew disease, along with an additional set of 12 images depicting healthy crop leaves.
2) The dataset known as the "Plant Pathology Challenge" from CVPR FGVC7, accessible on Kaggle[8], comprises a collection of more than three thousand meticulously annotated RGB images. This dataset encompasses 1000+ images featuring apples, 1400 images depicting rough symptoms in cedar apples, 187 images capturing diverse complex diseases (including multiple diseased leaves), and 865 images showcasing good apple leaves.
3) Several other sources consist of authentic pictures compiled by researchers to meet specific research requirements. These datasets include images related to crops such as corn, tea, soybeans, cucumbers, apples, and grapes.

4) Another common method for data acquisition involves growing plants and intentionally inoculating them with viruses. This approach is frequently employed in use cases that utilize multi-spectral images for disease diagnosis.

A. Early works related to CovNets

Researchers worldwide have delved into various approaches for automating the detection of plant diseases, including ML[9,10] , neural[2] and Image Processing techniques. In current section, we present a basic summary of the advanced discoveries in the literature developed for plant diagnosis models.

Sanga et al.[11] pioneered the development of a pathology identification tool for banana trees, using 5 CNN algorithms including InceptionV3, three variants of ResNet (18,50,152) and VGG_16. Additionally, they created a mobile application that enabled farmers to identify banana plant diseases by uploading leaf images from their smartphones, utilizing the InceptionV3 model with 99% confidence. Notably, the ResNet-152 model featured 60 million training parameters.
Shaikh et al. [12] explored the performance of five new-age CNN models, including GoogLeNet, AlexNet, Overfeat, AlexNetOWTBn, and VGG, for plant disease detection with the PlantV dataset. VGG exhibited the best performance, achieving an accuracy of 0.995 while utilizing 0.13 billion training attributes, as reported in the VGGNet paper , conducted an extensive analysis of AlexNet and GoogLeNet CNN models for plantation disease prediction, experimenting with 60 different configurations. Their findings highlighted the superiority of GoogLeNet with two way learning, achieving an accuracy of 0.993 and employing approximately 7 million training parameters, as per their paper.

Punam Bedi and Pushkar Gole. [13] applied various new-age CNN model and classifiers to automatic plantation disease identification using the PlantV dataset. They leveraged Visual Geometry Group-16, ResidualNetwork -50, and GoogLeNet for feature extraction, while classification was carried out using KNN and SVM models. Their investigation revealed that SVM based ResidualNetwork -50 performed better than the other models, achieving an accuracy of 0.98. In line with the ResNet paper, the ResNet-50 model featured around 25 million training parameters (He et al., [14]). Similarly, Albattah et al.[15] introduced an neural based plant disease diagnosis system for potatoes yield. CNN algorithms like VisualGeometryGroup-19, VisualGeometryGroup-16, and InceptionV3 were used for feature engineering, and classifiers like SigmoidRegression, KNN,SVM, and Deep Network were used for got sick By Logistic Regression, VGG-19 proved to be positive, with an accuracy of 0.978. The VGG-19 model consisted of approx. 144mil training attributes, as described in the VGGNet.

Khamparia et. al.[16] presented a deep CAE model for disease detection in plants. They focused on 910 leaf images of three crops: potatoes, tomatoes, and baby corns, which were divideds among 6 groups (five patients and one healthy subject). Their model got a training accuracy of 1 with 0.86 in testing. However, the noticeable difference between training and testing accuracy hinted at potential overfitting. Their system employed around 3.3 million training parameters, significantly exceeding the 9k used in our proposed work. Utilizing CAE and CNN for plantation disease identification, our novel hybrid model achieved higher testing accuracy compared to Khamparia et al.

It is noteworthy that all the aforementioned research works shared a common drawback, which is the utilization of a large amount of training attributes. Training models with a large amount of variables demands considerable time and computational power. In response to this limitation, our research focuses on minimizing the number of training attributes for plantation disease diagnosis without compromising classification accuracy. Consequently, we introduce a novel combined model that employs Convolutional Autoencoders (CAE) to minimise the dimensions of input leaf samples before classification using CNN, thus significantly limiting the number of training parameters—an essential contribution of this study.

III Problem Formulation

Globally, the agro industry plays a crucial role in food security and financial stability. However, plant diseases pose a serious threat to agricultural production and food supply. Conventional diagnostic methods usually require visual inspection, which can be time-consuming, prone to human error, and dependent on higher levels of knowledge In recent years, technology has become a critical component of modern agriculture, and provides opportunities for improved diagnostic techniques.

Recent advancements in technology, including high-resolution imaging, sensor networks, and machine learning, have shown promise in addressing the challenges of disease prediction in plants. Neural network techniques, in particular, have exhibited remarkable success in various CV tasks, including image object recognition and classification. The research gap in this domain pertains to the lack of standardized, scalable, and interoperable solutions for plant disease detection. Existing studies tend to focus on specific crops or diseases, leaving a void for a comprehensive, cross-species approach. Additionally, deploying such systems in real-world agricultural settings introduces technical challenges related to environmental conditions, resource constraints, and data privacy.

This research aims to bridge these gaps by designing, developing, and evaluating a versatile "Plantation Disease Detection System" that leverages deep learning techniques to detect diseases across a wide range of plant species. The primary objectives include creating a comprehensive dataset of plant diseases, developing an accurate and efficient deep learning model, integrating the system into practical plantation settings, and evaluating its performance.

The importance of this research is in its capacity to improve both the quantity and quality of crop yield, reduce manual inspection dependency, & contribute to sustainable agriculture. By providing a blueprint for integrating deep learning techniques into real-world agricultural scenarios, this research strives to address critical challenges in the agricultural industry and transform disease detection systems.

In summary, this research seeks to revolutionize disease detection in plantations by harnessing the capabilities of neural network techniques and addressing the pressing challenges of the agricultural sector.

IV Methodology

Data Collection and Preprocessing

Data Acquisition: We acquired a diverse dataset comprising images of plants affected by various diseases and healthy plants, encompassing various plant species and disease types.

Data Preprocessing: The acquired dataset was preprocessed by standardizing image resolutions, and a split was made and validation was done using multiple CV techniques. Moreover, data augmentation methods, including rotation, flipping, and cropping, were implemented on the training dataset to boost variability and mitigate the potential for overfitting.

Model Selection and Architecture

Selection of ResNet50: We selected the ResNet50 model as our pre-trained convolutional neural network for feature extraction, primarily due to its well-documented performance in image classification tasks and its ability to capture intricate features.

Transfer Learning: We initiated the ResNet50 model with pre-trained weights from a large dataset, specifically ImageNet. To adapt the model to our plant disease detection task, we froze the early layers, retaining their learned features, and fine-tuned the later layers.

Model Development

Building the Classifier: We designed a custom classification head on top of the ResNet50 base. This head comprised fully connected layers with appropriate functions that introduces non-linearity.

Training the Model: We closely observed the performance on the validation set, implementing early stopping to avoid over fitting. For the multi-class classification task, we employed categorical cross-entropy as the chosen loss function.

Model Evaluation

Performance Metrics: We assessed the model's performance using a variety of metrics, encompassing accuracy, precision, recall, F1-score, and the confusion matrix. These measures offered valuable insights into the model's ability to effectively classify plant diseases.

Hyperparameter Tuning: We adjusted different parameters such as learning rate η, batch size, and dropout rates to optimize the model's performance in this regard.

Visualization: The model's training history was visualized through plots of loss and accuracy over epochs, helping us identify trends and potential issues.

Model Testing

Test Set Evaluation: We assessed the model's performance on an independent test set to obtain the final evaluation of its accuracy and generalization.

Real-world Deployment: In anticipation of real-world deployment, we prepared the model for use in plantation settings, which would enable real-time disease detection in the field.

Interpretability and Explainability

Activation Maps: We generated activation maps to gain insights into the regions of the image that were influential in the model's decision-making.

Class Activation Mapping (CAM): CAM was used to highlight the specific areas of the image that contributed most to the disease classification.

V PROPOSED SYSTEM

In this section, we propose the architectural aspects of our novel model designed to detect accurate plant diseases. Our NeuralNet is a fusion of Convolutional Autoencoders (CAE) and Convolutional Neural Networks (CNN), designed to leverage both unsupervised feature learning and supervised classification for improved disease diagnosis in plants.

5.1. ResNet50

The initial stage of our model utilizes a ResNet50 for pre-filtering and dimensionality reduction. The encoder compresses input images into a low-dimensional latent representation, and the decoder reconstructs the original input based on this latent representation.The architecture of the ResNet can be summarized as follows:

Table 1 : ResNet50 Layers

Name	Filters	Size	Remarks
Convolutional Layer 1	32 filters	3x3 kernel	ReLU activation
Max-Pooling Layer 1		2x2 pool size	
Convolutional Layer 2	64 filters	3x3 kernel	ReLU activation
Max-Pooling Layer 2		2x2 pool size	
Convolutional Layer 3	128 filters	3x3 kernel	ReLU activation
Max-Pooling Layer 3		2x2 pool size	

Table 2 : Augmentated Layers

Name	Filters	Size	Remarks
Convolutional Layer 4	64 filters	3x3 kernel	ReLU activation
Up-Sampling Layer 1		2x2 up-sampling	
Convolutional Layer 5	32 filters	3x3 kernel	ReLU activation
Up-Sampling Layer 2		2x2 up-sampling	
Convolutional Layer 6	1 filters	3x3 kernel	, Sigmoid activation (for image reconstruction)
Max-Pooling Layer 3		2x2 pool size	

The encoder extracts essential features from plant images, and the decoder's purpose is to recreate the original image. During training, the CAE learns to capture significant patterns in the input images, which are beneficial for disease detection.

5.2. Convolutional Neural Network (CNN)

The output of the ResNet50 serves as the feature representation for the second part of our model, a Convolutional Neural Network (CNN) for disease classification. The CNN takes the compressed feature representation and learns to classify plant images into good or bad categories.

JJTRET "Voice of Innovative Minds"
International Journal of Trendy Research in Engineering and Technology
Volume 8 Issue 1 February 2024
ISSN No.2582-0958

The architecture of the CNN can be summarized as follows:

Table 3 : CNN Layers

Name	Filters	Size	Remarks
Convolutional Layer 1	64 filters	3x3 kernel	ReLU activation
Max-Pooling Layer 1		2x2 pool size	
Convolutional Layer 2	128 filters	3x3 kernel	ReLU activation
Max-Pooling Layer 2		2x2 pool size	
Fully Connected Layer 1	256 units		ReLU activation
Fully Connected Layer 2	128 units		ReLU activation
Output Layer	2 unit		Softmax activation

5.3. Model Fusion

The outputs of the CAE and the CNN are concatenated, and the combined representation is fed to a final fully-connected layer with Softmax to convert the produced values into probability scores.

The fusion architecture allows the model to leverage both the unsupervised feature learning capabilities of the CAE and the discriminative power of the CNN for accurate disease prediction.

The entire model is trained end-to-end using a labeled dataset of plant images, with loss functions and optimization techniques suitable for each part of the architecture.

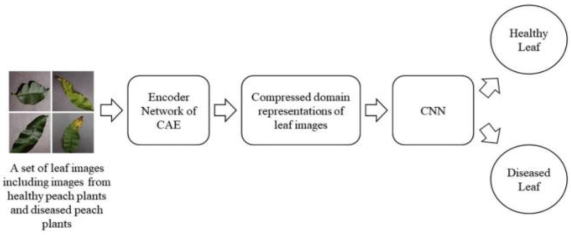

Figure 3 : Flowchart

VI RESULT AND ANALYSIS

This section delves into the outcomes of the experiments conducted as part of the current research. To begin, we present the results obtained from the CAE network. Subsequently, we showcase the outcomes of the presented hybrid model. The evaluation of the CAE network's performance has been carried out using the NRMSE loss metric, computed by comparing the original leaf images with their remodeled counterparts. The NRMSE loss for both training and testing phases stands at 59.7% and 60.7%, respectively. The variation in the Reconstruction Loss over different epochs is depicted in Figure 5.

Figure 4 : Orginal plant image

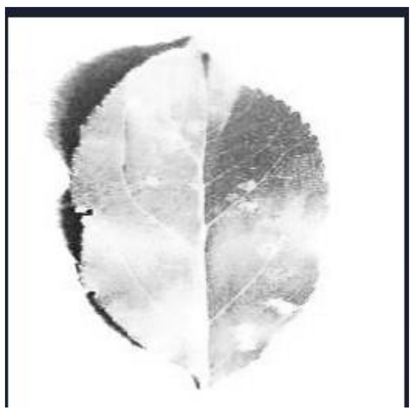

Figure 5 : Grayscale plamt image

IJTRET "Voice of Innovative Minds"
International Journal of Trendy Research in Engineering and Technology
Volume 8 Issue 1 February 2024
ISSN No.2582-0958

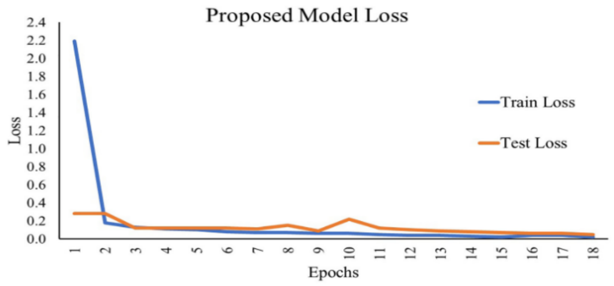

Figure 6 : Epochs Vs. Loss

Returning to the proposed hybrid model, this resulted in an impressive training accuracy of 0.993 and 0.983 in testing, with minimum training inaccuracies and testing loss of 0.02 and 0.05, respectively as shown in figure.

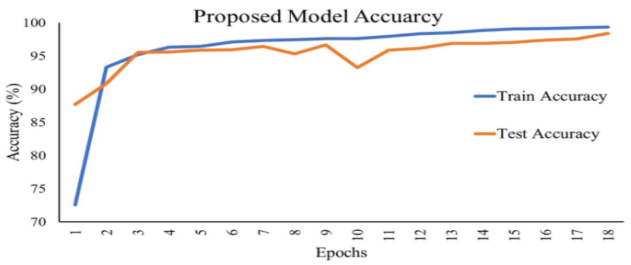

Figure 7 : Epochs Vs. Accuracy

Moreover, evaluation has been done (as per Zaki and Meira Jr,[17]) and different metrics were used and calculated for the proposed model, resulting in a Precision score of 91.0%, a Recall of 93.72%, and an F1-measure of 92.36%.

Table 4 : Comparision of existing & proposed models

Name	Proposed Research	Testing Accuracy	Number of Data Inputs
Sanga, Sophia, et al.[11]	VGC16	90.90%	2.6 Million
Islam, Md Tariqul[2]	AlexNetOWTBn	94.29%	2 Million
Srivastava Prakanshu [1]	CNN	88%	1.01 Million
Proposed Model	ResNet50	94%	64,000

A noteworthy aspect of the proposed work is its efficiency in terms of training parameters. With only 9,914 training parameters, our model stands out as significantly more streamlined in comparison to state-of-the-art systems developed by previous researchers. This translates to a substantially reduced training and prediction time for our hybrid model.

The practical applications of our proposed model are twofold. Firstly, it can be employed for automated plantation disease prediction on resource-constrained model plus swift learning/feeding and identification processes. Secondly, the algorithm should be adopted for use on smartphones, enabling on-device DL for plantation disease diagnosis. This approach minimizes latency and safeguards data privacy for farmers by avoiding the need to transmit plant images to remote servers.

VII CONCLUSION

Identifying plant diseases during their initial phases poses a daunting and intricate challenge. Many scholars have explored a variety of ML or neural network techniques to address plantation disease diagnosis. However, a common issue with these approaches is the excessive use of training parameters, which can lead to suboptimal classification accuracy. In this study, we introduced an innovative combined model designed for automated crop disease identification and diagnosis. This model leverages two distinct neural network techniques: CAE and CNN. The hybrid model's approach begins with the initial creation of compacted depiction of plant images via CAE. These compressed representations are then used for classification purposes through the CNN. The dimensionality reduction accomplished by the CAE results in a substantial decrease in the quantity of features, leading to a noteworthy reduction in training parameters when collated with existing systems. To evaluate the model's effectiveness, we used it for identifying instances of spot or pigment diseases caused by bacteria in nectarine and berry plants. The results were noteworthy, as the model attained a training accuracy of 0.97 and 0.943 in testing, all while utilizing only 64k training attributes. The effective utilization of training features in our devised integrated model substantially reduces both the training time in our detection system and the duration needed for disease identification in crops using the optimized model.

REFERENCES

[1] Prakanshu Srivastava , Kritika Mishra , Vibhav Awasthi , Vivek Kumar Sahu and Pawan Kumar, "Plant disease detection using convolutional neural network." Journal of Advanced Research, 9 (2021) 691-698 DOI: 10.21474/IJAR01/12346.

[2] Md Tariqul Islam. "Plant disease detection using CNN model and image processing." Int. J. Eng. Res. Technol 9(2020) 291-297.

[3] Ajmeera Kiran, et al. "Plant Disease Detection using Image Processing with Machine Learning." 2023 4th International Conference on Electronics and Sustainable Communication Systems (ICESC). IEEE, 2023.

[4] Vijai Singh, Namita Sharma, and Shikha Singh. "A review of imaging techniques for plant disease detection." Artificial Intelligence in Agriculture 4 (2020) 229-242.

[5] Shiv Ram Dubey and Anand Singh Jalal. "Adapted approach for fruit disease identification using images." Image processing: Concepts, methodologies, tools, and applications. IGI Global, 2013: 1395-1409.

[6] Chai, A.L.; Li, B.J.; Shi, Y.X.; Cen, Z.X.; Huang, H.Y.; Liu, J."Recognition of tomato foliage disease based on computer vision technology." Acta Horticulturae Sinica 37 (2010) 1423-1430.

[7] Lili Li, Shujuan Zhang, and Bin Wang. "Apple leaf disease identification with a small and imbalanced dataset based on lightweight convolutional networks." Sensors 22 (2021) 173.
Lili Li, Shujuan Zhang, and Bin Wang. "Plant disease detection and classification by deep learning—a review." IEEE Access 9 (2021) 56683-56698.

[8] Kaggle Plant Pathology 2020 FGVC7 Competition. URL: https://www.kaggle.com/competitions/plant-pathology-2020-fgvc7

[9] Imtiaz Ahmed and Pramod Kumar Yadav. "Plant disease detection using machine learning approaches." Expert Systems 40 (2023) e13136.

[10] Shima Ramesh, et al. "Plant disease detection using machine learning." 2018 International conference on design innovations for 3Cs compute communicate control (ICDI3C). IEEE, 2018.

[11] Sophia Sanga, Victor Mero, Dina Machuve, Davis Mwanganda, "Mobile-based deep learning models for banana diseases detection." arXiv preprint arXiv:2004.03718 (2020).

[12] Tawseef Ayoub Shaikh, Tabasum Rasool, and Faisal Rasheed Lone. "Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming." Computers and Electronics in Agriculture 198 (2022)107119.

[13] Punam Bedi and Pushkar Gole. "Plant disease detection using hybrid model based on convolutional autoencoder and convolutional neural network." Artificial Intelligence in Agriculture 5 (2021) 90-101.

[14] Kaiming He Xiangyu Zhang Shaoqing Ren Jian Sun,' Deep Residual Learning for Image Recognition', 016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),2016, http://dx.doi.org/10.1109/CVPR.2016.90

[15] Waleed Albattah, Marriam Nawaz, Ali Javed, Momina Masood & Saleh Albahli., "A novel deep learning method for detection and classification of plant diseases." Complex & Intelligent Systems 8 (2022)1-18.

[16] Aditya Khamparia, Gurinder Saini, ·Deepak Gupta, Ashish Khanna, ·Shrasti Tiwari, ·Victor Hugo C. de Albuquerque "Seasonal crops disease prediction and classification using deep convolutional encoder network." Circuits, Systems, and Signal Processing 39 (2020)818-836.

[17] Mohammed J. Zaki and Wagner Meira. "Data mining and analysis: fundamental concepts and algorithms." Cambridge University Press, 2014.