A SURVEY ON PLANT LEAF DISEASE DETECTION USING MACHINE LEARNING

Mr. Raghuram A S[1], Panchami C H[2], Pavan V[3] , Yashwanth N[4], Sneha B[5]

Assistant Professor, Dept of CSE, ATME College of Engineering, Mysuru, Karnataka, India[1]

8th Semester Student, Department of CSE, ATME College of Engineering, Mysuru, Karnataka, India[2]

8th Semester Student, Department of CSE, ATME College of Engineering, Mysuru, Karnataka, India[3]

8th Semester Student, Department of CSE, ATME College of Engineering, Mysuru, Karnataka, India[4]

8th Semester Student, Department of CSE, ATME College of Engineering, Mysuru, Karnataka, India[5]

ABSTRACT

The field of agriculture plays a significant role in our lives and is the backbone of our economy. Proper management can result in profitable agricultural production. However, farmers often lack the expertise to identify and address leaf diseases, which can lead to lower production levels. Detecting plant leaf diseases is crucial because the success of agricultural production depends on it. A deep Convolutional Neural Network (CNN) is a potential solution for detecting and classifying plant leaf diseases. The main objective of this research is to detect plant leaf diseases by monitoring large fields of crops and automatically identifying the symptoms of diseases to provide timely medical treatment. In this study, we compare a proposed deep CNN model with a popular transfer learning approach, VGG16. Plant leaf disease detection has a wide range of applications in various fields, including biological research and agriculture institutes. It is also an essential research topic as it can prove beneficial in monitoring large fields of crops and detecting the symptoms of diseases as soon as they appear on plant leaves.

Keyword: - *Deep learning, Convolutional neural networks (CNNs), image classification.*

1. INTRODUCTION

Plant diseases pose a significant challenge to global agricultural productivity, food security, and economic stability. Each year, millions of hectares of crops are destroyed by plant diseases, which leads to a considerable reduction in food availability worldwide. Traditional methods of disease detection and management are tedious and often fall short in addressing the complexity and scale of modern agricultural systems.

In recent years, deep learning techniques, particularly convolutional neural networks (CNNs), have revolutionized the field of image-based disease detection. These algorithms can learn intricate patterns and features from images using vast amounts of labeled data and powerful computational resources. This enables accurate and efficient classification of diseases in plants, offering unparalleled opportunities to develop cost-effective, scalable, and automated solutions for disease monitoring and management in agriculture.

This research paper examines different deep learning methodologies for detecting plant leaf diseases, focusing on three popular architectures: CNNs, DenseNet, and ResNet. These architectures can capture spatial hierarchies, facilitate feature reuse, and address challenges in training deep neural networks. The study evaluates the performance of these models using a curated dataset of high-resolution images of plant leaves showing various disease symptoms across a diverse range of crop species, including apples, cherries, corn, grapes, oranges, peaches, peppers, potatoes, strawberries, and tomatoes. Diseases such as Black Rot, Powdery Mildew, Common Rust, Bacterial Spot, and Early Blight affect these crops.

The study has two primary objectives: firstly, to assess the effectiveness of different deep learning architectures in accurately detecting and classifying plant leaf diseases, and secondly, to provide insights into the strengths and limitations of each architecture in this field. By comparing the performance of CNNs, DenseNet, and ResNet models, the research aims to understand their relative merits and suitability for real-world applications in agricultural disease management.

Moreover, the study evaluates additional performance indicators such as precision, recall, and F1-score, which offer a better understanding of model capabilities, especially in scenarios with imbalanced class distributions or critical misclassifications. The comprehensive analysis aims to provide valuable insights and empirical evidence to researchers, practitioners, and stakeholders in the agricultural sector, informing decision-making processes and resource allocation for disease control and mitigation efforts.

Figure 1: Sample images from the dataset

2. LITERATURE SURVEY

[1] This study introduces a comprehensive approach to detecting leaf diseases in crops using deep learning techniques. The authors utilize Convolutional Neural Networks (CNNs) to analyze images of plant leaves and classify them as healthy or diseased. Their methodology includes dataset acquisition, preprocessing, model selection, training, and evaluation. The research highlights the significance of deep learning in the field of agriculture and showcases promising results in disease detection, which can contribute to the progress of precision agriculture.

[2] This article proposes a new method for identifying plant leaf diseases by using deep learning techniques, which are improved with channel attention and channel pruning mechanisms. The authors introduce attention mechanisms to highlight the most important features in plant leaf images, which increases the accuracy of the model in detecting diseases. They also use channel pruning techniques to optimize model complexity and reduce the computational overhead. The study shows that these enhancements are effective in achieving accurate and efficient disease identification, and can help in precision agriculture and crop protection.

[3] This research focuses on identifying and classifying leaf diseases using deep learning algorithms that have been optimized through various strategies such as hyperparameter tuning, architecture design, and training methodologies. The authors aim to improve the performance of deep learning models in disease detection tasks by overcoming challenges such as overfitting and model convergence issues. Their study highlights the importance of optimization techniques in enhancing the accuracy and reliability of plant disease identification systems. This could facilitate early detection and intervention in agricultural settings, ultimately helping to prevent crop loss.

[4] This research paper introduces a novel approach for detecting and recognizing plant leaf diseases using deep learning. The authors have utilized convolutional neural networks (CNNs) to identify plant leaves in images and classify them as healthy or diseased. Their methodology includes acquiring a dataset, pre-processing it, training the model, and evaluating its performance. This study demonstrates that deep learning techniques can be effectively used for automated plant disease recognition, which can help in the development of efficient and scalable solutions for crop monitoring and management.

[5] This study presents an automated approach for plant disease classification using vision transformers based on deep learning. The use of vision transformers has gained popularity in recent years due to their effectiveness in image classification tasks. The authors utilize these transformers to automatically extract hierarchical representations of plant leaf images in order to classify them into different disease categories. Their research is a significant contribution to the exploration of transformer-based architectures in the field of plant pathology and provides valuable insights into their potential applications and performance in disease recognition tasks.

[6] This review article presents a thorough overview of the current research on using deep learning techniques to detect and classify plant leaf diseases, with a focus on the context of Bangladesh. The authors analyze a variety of deep learning architectures, datasets, and methodologies used in previous studies and discuss their implications for agricultural applications in Bangladesh. Furthermore, the paper suggests a system architecture that caters to the specific needs and challenges of plant disease detection in Bangladesh, highlighting possibilities for future research and innovation in this area.

[7] This research paper showcases a study on the detection and identification of plant leaf diseases using advanced deep-learning models. The authors have used state-of-the-art techniques to analyze images of plant leaves and classify them into various disease categories. Their methodology includes the acquisition of datasets, preprocessing, training of models, and their evaluation. The study is a valuable contribution to the field of precision agriculture, as it demonstrates the effectiveness of deep learning models in automated disease detection and identification. It also provides insights into the potential applications and advancements in agricultural technology.

[8] In this article, the author discusses the use of deep learning techniques to detect plant diseases. The study presents a step-by-step process for training deep-learning models to examine plant leaf images and identify signs of disease. The research highlights the significance of utilizing large-scale datasets and advanced neural network architectures for precise disease detection. Through the application of deep learning algorithms, the study aims to offer an effective, scalable, and efficient solution for monitoring crops and managing diseases in agricultural settings.

[9] This review article provides a comprehensive overview of the recent research on detecting and categorizing plant diseases through deep learning techniques. The authors analyze different deep learning models, datasets, and methodologies used in previous studies and discuss their significance for agricultural purposes. The review emphasizes the progress that has been achieved in precision agriculture by employing deep learning algorithms for automated disease detection.

[10] This review article presents a thorough overview of the latest research on detecting and categorizing plant diseases with deep learning techniques. The authors examine various deep learning models, datasets, and approaches used in previous studies and discuss their significance for agricultural applications. The review highlights the progress that has been made in the field of precision agriculture by utilizing deep learning algorithms for automated disease detection and classification. Furthermore, the paper identifies the challenges and opportunities for future research in this field.

3. METHODOLOGY

[1] The Methodology and the steps involved in the proposed system is shown in the form of block diagram in the below figure-2. Convolutional Neural Network (CNN) models are highly effective in recognizing and classifying objects using image databases. However, there are still some challenges associated with these models, such as the long duration of training and the need for large datasets. To extract low-level and complex features from images, deep CNN models are required, which makes the training process more complicated. Transfer learning is a technique that involves using pre-trained networks, where the model parameters learned on one dataset can be used for other problems. In this section, we will discuss the methodologies used in this work.

.

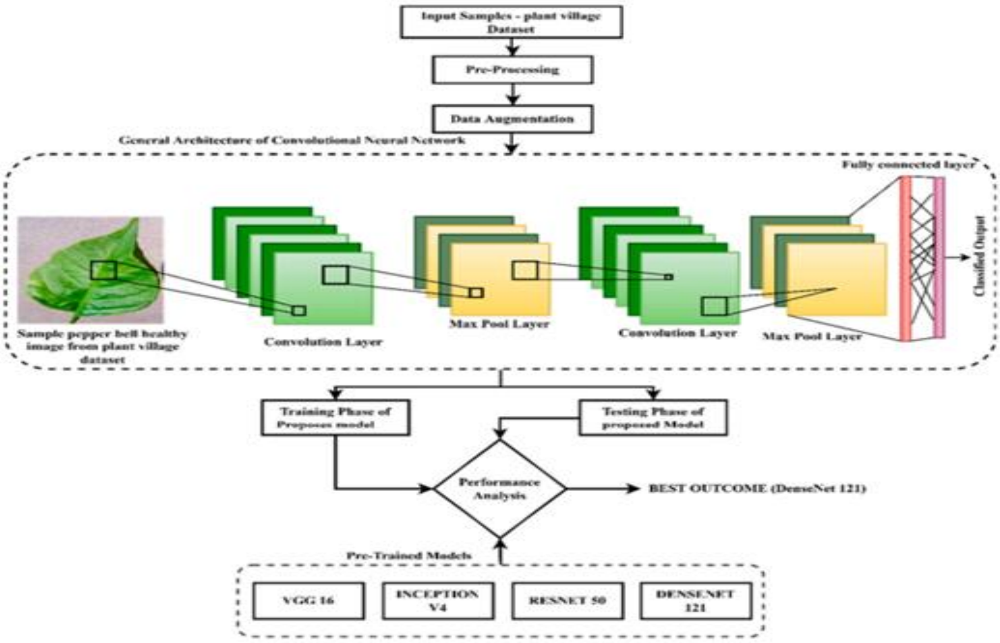

Figure 2: Block Diagram of Proposed System [1]

[2] Figure 3A illustrates the process of the channel attention module inserted into the model. Specifically, we traverse all model layers and insert the channel attention module after each convolutional layer. Then the new model is trained to achieve a better performance effect. Accordingly, Figure 3B depicts removing unimportant channels using channel pruning on the model. We obtain the weight relation in the channel from the well-trained model. Then the L1-normalization of the channel weights is calculated and ranked. The unimportant and associated channels per layer are removed based on a predetermined local compression ratio. Finally, the new model is updated with the remaining channels and retrained to achieve better performance. Figure 3C illustrates the implementation process of CACPNET. In summary, this paper aims to develop a lightweight model with better performance and lower parameters, with the following sections introducing the details on implementing CACPNET.

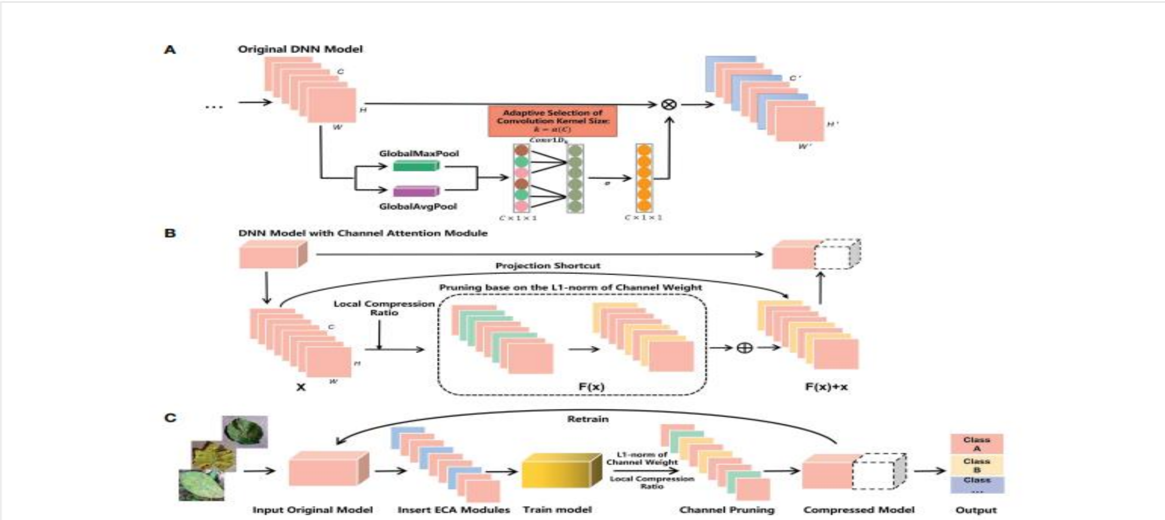

FIGURE 2
Workflow of the CACPNET. **(A)** Insert channel attention modules into the original model. The blue sections are the channel attention modules. Σ is the sigmoid activation function. **(B)** Use channel pruning to compress the DNN model. The green and yellow sections are the removed channels. **(C)** Overall structure of CACPNET.

Figure 3: Workflow of the proposed system [2]

[3] Plant leaf disease detection involves a collection of copy datasets, pre-processing, feature extraction, segmentation, and classification the last three processes are enhanced by the ACO-CNN method. The proposed ACO-CNN process for disease detection is shown in Figure 4.

Data Collection**:** The collected dataset is invented of healthy and infected leaves with the following sets; greening, Canker, melanosis, and blackspots. The proposed model was used to classify the infected leaves from the healthy leaves.

Data pre-processing: After the image is selected the initial step for leaf disease detection is pre-processing. For noise reduction and removal of other objects, the median filter has been applied to enhance the leaf images.

Segmentation: Segmentation is the progression of classifying the plant leaf images into a smaller section of the surface, which is related to characteristics.

Ant colony optimization (ACO): Here the ACO is used to identify the infected pant leaf from the healthy leaf. One of the important parameters that should be modified first in the database is the pheromone rate. All pheromone material of the feature is examined in the matrix(h) with a dimension of G*G such that G as rows and columns denote the complete number of novel feature vectors. Ant colony optimization parameters are modified and the main calculation such as the experimental function F is calculated. Identify the confined best subset and assets, for the next repetition.

Convolution neural network (CNN): The classifiers of the Convolution Neural Network (CNN) are used to identify each leaf disease. It efficiently evaluates graphical pictures and removes the needed features through its multi-layered construction. CNN classifier contains four layers: image input, convolutional layer, Max pooling layer, fully connected layer, and output. The range of image pixel intensity values of plant leaf in the dataset before training convolution neural network model. During the training phase, CNN is the fastest model.

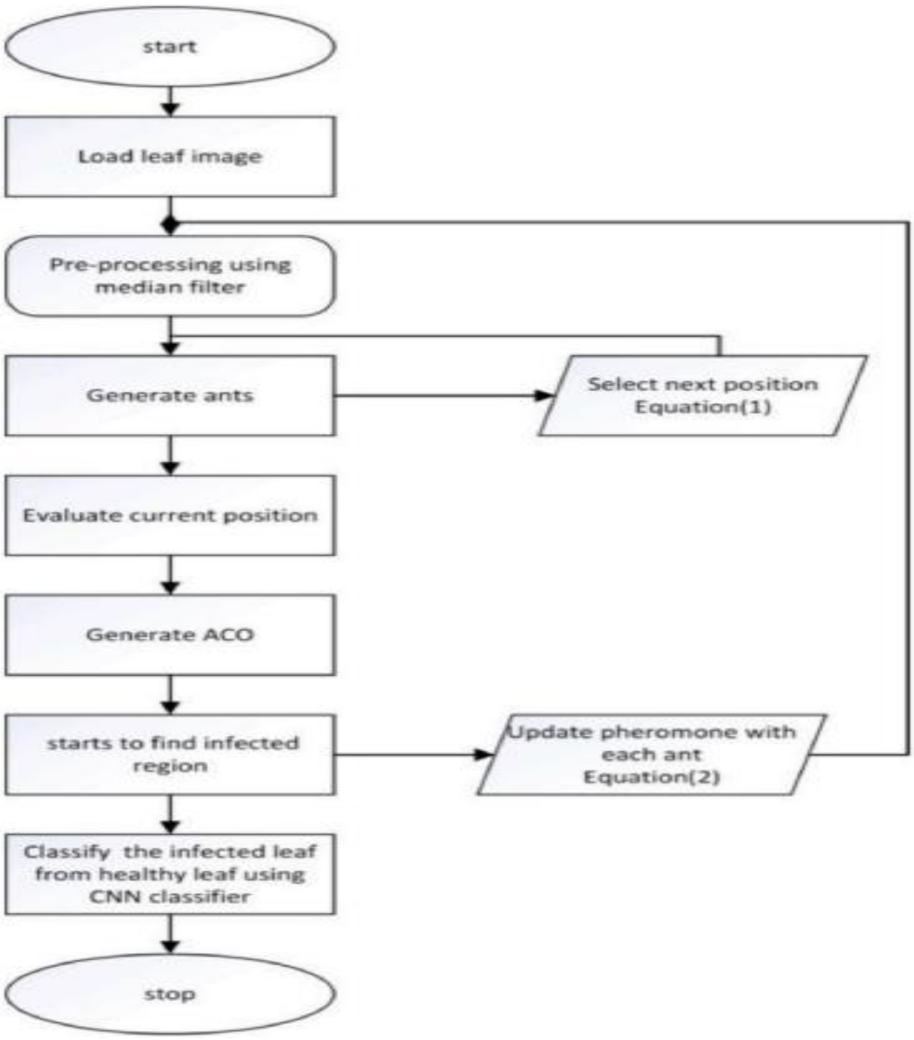

Figure 4: Flow diagram of the proposed method [3]

[4] In this paper, several structures which are described as follows are studied. At first, two main building blocks are defined as CNN and Transformer block:

Convolutional neural networks block; The CNN block consists of two convolutional layers with 3 by 3 kernels. In these two layers, padding and activation are not applied. The output of the second convolutional layer is entered into a leaky ReLU layer. Ten one max pooling layer with 2 by 2 kernel is considered which can be seen in Figure 5a. The idea of using two consecutive convolutional layers with 3 by 3 kernels comes from the VGG structure which suggested using the smaller filters on top of each other has the same receptive field with bigger filters while they have less trainable parameters in comparison with the bigger ones.

.

Transformer block.: The second block is the Transformer block. As it is shown in Figure 5b, the input of the block is fed to the layer normalization layer. The normalization layer is followed by a multi-head attention layer. Four attention heads are used in all the transformer blocks and the projection dimension is 64. Then the output of the multi-head attention layer adds to the input of the block with a skip connection. In the next step, the layer normalization layer is applied again, and the output of this layer goes to the fully connected layer. The final output of the block is calculated by summing the output of the fully connected layers and the input of those layers with another skip connection. It should be noted that in the map part shown in the block diagrams, two fully connected layers with 128 and 64 neurons exist and the activation of each of them is gelu.

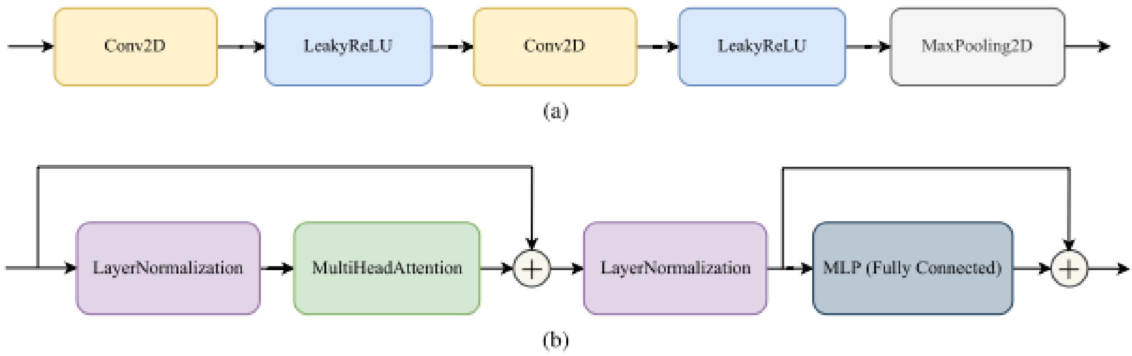

Figure 5: The schematics of the (a) convolutional block structure and (b) transformer block structure.

[7] This is research paper presents the detection of leaf diseases in corn, apple, tomato, rice, and potato leaves by extraction of deep features and texture and color features, followed by feature selection based on BPSO, comparing the two classifications: Bayesian optimized SVM and random forest. Figure 6 provides a general flow chart of the proposed work.

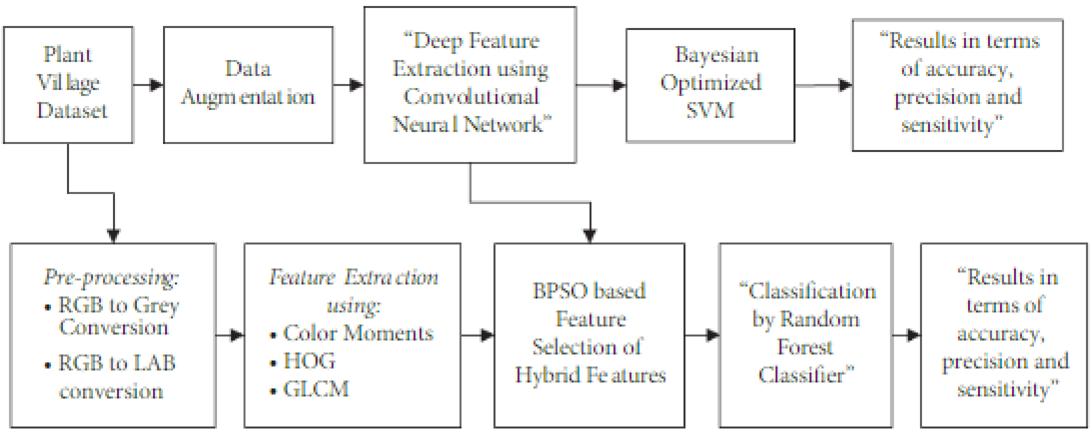

Figure 6: Block diagram for the hybrid approach of plant leaf disease detection

Convolution Operation: "The convolution is an operation applied to two functions with real numbers as arguments.

[7] is research paper presents the detection of leaf diseases in

corn, apple, tomato, rice, and potato leaves by extraction of deep features and texture and color features, followed by feature selection based on BPSO, comparing the two classifications: Bayesian optimized SVM and random forest. Figure 1 provides a general flow chart of the proposed wor Classification of Deep Features by Bayesian Optimized Support Vector Machine: In supervised learning, the classifier is trained through the presentation of a set of examples (input and desired output). A training set is used in supervised methods to educate the model to produce the desired output. 'is training dataset contains both right and incorrect outputs, allowing the model to improve over time. It is expected that based on this knowledge, the classifier will be able to accurately predict the output of new data not previously presented, being able to act in a linear or nonlinear way. Considering the SVM, they divide the feature space into regions using an optimal separation hyperplane positioned exactly in the center between the margins of the two classes. Among the nonlinear functions that can be used in the SVM analysis are quadratic, polynomial, radial basis function (RBF), and Gaussian and two-layer perceptron. 'is technique seeks to maximize the separation margin of samples from two groups. 'e solution to this optimization problem has a broad and established mathematical theory and can be expressed by the following equation:

$$\max W(\lambda) = \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i,j=1}^{N} y_i y_j \lambda_i \lambda_j \left(x_i x_j \right),$$

$$\text{subject to} \begin{cases} 0 \le \lambda_i \le C, \\ \sum_{i=1}^{N} \lambda_i y_i = 0, \quad i = 1, 2, \ldots, N, \end{cases}$$

4. CONCLUSION

We researched to assess the effectiveness of deep learning models in identifying plant leaf diseases. Our study included a diverse dataset with various disease categories across multiple plant species. We experimented with Convolutional Neural Networks (CNNs), DenseNet, and ResNet architectures and found that the CNN model was the top performer, achieving an accuracy of 99%. However, CNNs presented challenges related to overfitting, especially in classes with limited data samples. We mitigated these challenges by using regularization techniques. DenseNet and ResNet architectures also offered competitive accuracies of 90% and 85%, respectively.

Our comparative analysis revealed variations in model performance across different disease classes, emphasizing the importance of robust evaluation metrics beyond overall accuracy. Precision, recall, and F1-score provided valuable insights into the models' abilities to correctly classify diseased leaves while minimizing false positives and negatives.

To further enhance disease detection accuracy and robustness, future research could explore ensemble learning approaches or hybrid architectures. Additionally, addressing challenges such as class imbalance and data scarcity through data augmentation techniques or transfer learning from related domains could be beneficial.

In future work, we will address the problems in real-time data collection and develop a multi-object deep learning model that can even detect plant diseases from a bunch of leaves rather than a single leaf. Furthermore, we are working towards implementing a mobile application with the trained model from this work. It will help farmers and the agricultural sector in real-time leaf disease identification.

Crop protection is a challenging task in organic farming, which requires a comprehensive understanding of weeds, pathogens, pests, and the crop being grown. Identifying leaf diseases at an early stage is crucial for the agricultural industry. The traditional method of using real-time images to identify leaf diseases has been effective, but a new proposed method can help farmers detect and recognize plant leaf diseases. The ACO-CNN optimization approach has been suggested for leaf disease detection, where ACO is used for feature extraction and CNN classifier for organization. This method is used to distinguish the infected leaf from the healthy leaf.

5. REFERENCES

[1] J., A.; Eunice, J.; Popescu, D.E.; Chowdary, M.K.; Hemanth, J. Deep Learning-Based Leaf Disease Detection in Crops Using Images for Agricultural Applications. Agronomy 2022, 12, 2395. https://doi.org/10.3390/agronomy12102395
[2] Chen R, Qi H, Liang Y, Yang M. Identification of plant leaf diseases by deep learning based on channel attention and channel pruning. Front Plant Sci. 2022 Nov 10;13:1023515. doi: 10.3389/fpls.2022.1023515. PMID: 36438120; PMCID: PMC9686387.
[3] Yousef Methkal Abd Algan Leaf disease identification and classification using optimized deep learning
[4] Borhani, Y., Khoramdel, J. & Najafi, E. A deep learning-based approach for automated plant disease classification using vision transformer. Sci Rep 12, 11554 (2022). https://doi.org/10.1038/s41598-022-15163-0
[5] S. V. Militante, B. D. Gerardo and N. V. Dionisio, "Plant Leaf Detection and Disease Recognition using Deep Learning," 2019 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE), Yunlin, Taiwan, 2019, pp. 579-582, doi: 10.1109/ECICE47484.2019.8942686.
[6] Plant Leaf Disease Detection and Classification Using Deep Learning: A Review and A Proposed System on Bangladesh's Perspective
[7]Chau, D.H. et al. (2022). Plant Leaf Diseases Detection and Identification Using Deep Learning Model. In: Hassanien, A.E., Rizk, R.Y., Snášel, V., Abdel-Kader, R.F. (eds) The 8th International Conference on Advanced Machine Learning and Technologies and Applications (AMLTA2022). AMLTA 2022. Lecture Notes on Data Engineering and Communications Technologies, vol 113. Springer, Cham. https://doi.org/10.1007/978-3-031-03918-8_1
[8] Kowshik Plant Disease Detection Using Deep Learning link: https://rspsciencehub.com/article_9834_e9c3d09ae4406e80dc2981886d728cd3.pdf
[9] S. V. Militante, B. D. Gerardo and N. V. Dionisio, "Plant Leaf Detection and Disease Recognition using Deep Learning," 2019 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE), Yunlin, Taiwan, 2019, pp. 579-582, doi: 10.1109/ECICE47484.2019.8942686.
[10] L. Li, S. Zhang and B. Wang, "Plant Disease Detection and Classification by Deep Learning—A Review," in IEEE Access, vol. 9, pp. 56683-56698, 2021, doi: 10.1109/ACCESS.2021.3069646