International Journal of Research Publication and Reviews, Vol 5, no 6, pp 1753-1759 June 2024

WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Plant Disease Detection, Classification and Analysis Using Convolutional Neural Network

Miss. Kulkarni Kartika B.[1], Prof. Dr. Pushpalata G. Aher [2], Dr. Mohd Maqueem[3], Dr. Pawan R. Bhhaladhare[4]

[1]M. Tech Student, *Department of Computer Science and Engineering,* Sandip University Nashik, India.
[2]Professor, *Department of Computer Science and Engineering,* Sandip University Nashik, India.
[3]Professor, Department of Computer Science and Engineering, Sandip University Nashik, India
[4]HOD, Department of Computer Science and Engineering, Sandip University Nashik, India

A B S T R A C T

Agricultural productivity is a crucial component of the Indian economy, with the cultivation of food and cash crops playing a significant role in both environmental sustainability and human welfare. Each year, various diseases affect crop yields, but the identification and treatment of these diseases are often insufficient. As a result, many plants perish due to a lack of awareness regarding the symptoms and treatments of these diseases. To address this issue, image processing techniques have been employed. In this study, 15 cases were input into the model, including 12 instances of Bell Pepper Bacterial Spot, Potato Early Blight, Potato Late Blight, Tomato Target Spot, Tomato Mosaic Virus, Tomato Yellow Leaf Curl Virus, and Tomato Bacterial Spot. Additionally, there were three cases of Tomato Early Blight, Tomato Late Blight, Tomato Leaf Mold, Tomato Septoria Leaf Spot, and Tomato Spider Mite, along with healthy leaf samples: Bell Pepper Healthy, Potato Healthy, and Tomato Healthy. The model achieved a test accuracy of 94.80%, and various performance metrics were derived to evaluate its effectiveness.

Background: *This study offers an overview of plant disease detection utilizing different algorithms, with a particular focus on a CNN-based approach for identifying plant diseases. Simulations and analyses were conducted on sample images to assess time complexity and the extent of the infected areas.*

Key Word: Convolutional Neural Network (CNN), Leaf Disease, etc.

I. INTRODUCTION

Crop growth and yield are vital for both agriculture and farmers, impacting them economically, socially, and in various other ways. Consequently, it is crucial to monitor crops closely at different growth stages to identify diseases promptly. However, relying solely on human observation can be insufficient and sometimes misleading. Therefore, automatic recognition and classification of crop diseases are essential for accurate identification. This chapter provides a comprehensive overview of the proposed methodology, including the background, problem statement, objectives, and scope.

A significant factor contributing to increased microbial infections is the lack of education among farmers in India. Once a crop becomes infected, it is challenging for farmers to diagnose the actual cause of the disease. Pathogens and pests significantly damage crops, reducing yields of the five main food crops by 10-40%, according to a study published by UC Agriculture and Natural Resources.

In India, where agriculture contributes 16% of the GDP and employs nearly 60% of the population, substantial measures are needed to prevent plant diseases. In 2016, agricultural losses amounted to 13 billion dollars, as reported by the Ministry of Food Processing Industries. Image processing and neural networks can be instrumental in detecting plant diseases. Recent research has demonstrated the effectiveness of neural networks and deep learning in classification tasks related to plant disease detection.

Agriculture is a critical sector in countries like India, where the economy depends heavily on it. It is essential to care for plants from seedling to harvest. Throughout this process, crops face numerous challenges, including weather conditions, diseases, and animal threats. While proper field protection can address animal threats, weather conditions remain beyond human control. The most crucial issue is protecting crops from various diseases, as these can significantly impact crop growth and yield. Early identification of diseases allows for timely intervention with appropriate fertilizers.

Digitalizing the process of disease identification and classification would greatly benefit agriculturists. It would reduce the time needed to identify diseases and increase the precision of classification, ultimately helping protect crops and improve yields.

II. RELATED WORK

1. Sardogan, M., et al. in 2018 [1] presented a model with a combination of convolutional neural networks (CNN) along with learning vector quantization(LVQ) for the identification and categorization of diseases of tomato plant leaves. The presented framework was implemented on the data size of 500 images with the four categories of diseases considered for tomato plant leaves. The convolutional neural network is utilized for the extraction of vital attributes from the images as well as for the classification.

2. Wallelign, S., et al. in 2018 [2] discussed the viability of convolutional neural network architecture for the classification of various plant diseases with the aid of leaf images. The mentioned framework is implemented by utilizing the LeNet, one of the popular CNN architecture, for disease classification in the aspect of soybean plants. The soybean plant leaf images of 12,763 samples are obtained from the standard database called PlantVillage. The mentioned framework able to achieve an accuracy of 99.32% indicating the viability of CNN with plant disease classification utilizing the leaf images.

3. Sladojevic, S., et al. in 2016 [3] concerned the generation of the new-age model for the identification of various diseases of 13 plant diseases out of the healthier plant leaf images. The deep learning architecture called Caffe was utilized for training the data. The results were obtained from the mentioned framework with a precision of 91percent to 98percent.

4. Fuentes, A., et al. in 2017 [4] proposed a framework and can be applied in two stages. At first, the meta architectures of Faster R-CNN, R-FCN, and SSD will be combined to form a single meta-architecture. Lastly, certain methodologies such as VGG- 16, VGG-19, and ResNet-50 will be attached to extract the features from more depth and these models' efficiency was estimated. When compared to many other models, the proposed framework efficiency is better.

5. Arivazhagan, S. and Ligi, S. V. in 2018 [5] proposed a framework based on automated deep learning for the recognition and classification of various diseases in mango plants. The dataset utilized for this framework consists of 1200 images which include both diseased and healthy leaves of mango. The accuracy obtained from the proposed framework is 96.67%. Oppenheim,

6. D. and Shani G. in 2017 [6] proposed a framework based on convolutional neural network architecture for the recognition and classification of various diseases in potato plants. The dataset utilized for this framework consists of 2465 potato images.

A series of steps need to be carefully followed for the process need to be followed in a diseplined manner:

Step-1: Image Acquisition for dataset creation: This step involves exploring various data sources from where data can be extracted for training the model and further how the test image input is to be provided.

Step-2: Image Pre-processing and background removal: This is most important phas, as it involves the quality assurance of the data. In the image pre-processing phase image is processed to desired color format, resized to desired size and images are denoised.

Step-3: Image Segmentation to obtain infected region: Region of interest that is the infected part of the leaf is identified. This is again one of the most crucial step, as entire analysis is dependent on the infected refion identified by the process of segmentation.

Step-4: Extraction of Features from images: On the basis of obtained region of interest which is the infected part of the leaf various image features like standard deviation, mean of red, blue and green channels, the entropy of image is extracted.

Step-5: Evaluate and identification of the affected region: By comparing the extracted region of interests & features which are extracted from the image, an efficient model is derieved.

Step-6: Processed Dataset creation: The data which are processed in previous stages are processed and extracted and converted to a csv file format and stored. This stored data is further utilized for analysis purpose.

Step-7: Training Data Extraction: Randomly the data in csv file is split. The 70% of the split data is used for training the proposed model.

Step-8: Testing Data Extraction: Randomly the data in csv file is split. The 30% of the split data is used for training the proposed model.

Step-9: Classification: Test data has labels such as: Late Blight, Early Blight, and Healthy, based on which classification is performed.

Step-10: Evaluation of proposed model: Depending on the obtained results from the classifier model, the evaluation metrics such as precision, recall, F1-score, and accuracy will be obtained.

(1) (2) (3)

Figure. 1 The three Sample leaves of potato are (1): leaf affected by Light Blight (2): leaf affected by Early Blight (3): leaf unaffected (Healthy)

IV. CONVOLUTIONAL NEURAL NETWORK (CNN)

Convolutional Neural Networks (CNNs), also known as ConvNets, are a type of neural network particularly effective for computer vision tasks such as image recognition and classification. These networks are feed-forward neural networks designed to analyze image data structured in a grid-like topology. Digital images can be represented as matrices of pixel values, where each pixel's value indicates color and intensity.

When humans view an image, the brain processes the visual information through interconnected neurons, each acting as an individual receptive field. Similarly, in CNNs, each neuron processes data within its receptive field. The layers in CNN architectures are arranged to first process simple features like lines and curves. This layered approach allows CNNs to enable computers to "see" by interpreting image data.

The primary four operations in CNNs are:

1. Convolution: This operation involves applying a filter to an input image to create feature maps, highlighting important aspects of the image such as edges and textures.
2. Pooling: This reduces the dimensionality of each feature map while retaining the most critical information, often through operations like max pooling or average pooling.
3. ReLU (Rectified Linear Unit): This activation function introduces non-linearity into the model, allowing it to learn from complex patterns.
4. Fully Connected Layers: These layers are used at the end of the network to combine all the features learned by the previous layers and make the final classification decision.

Through these operations, CNNs can efficiently and accurately analyze and classify images, mimicking the human visual processing system.

- Convolution operation
- Non-Linearity
- Pooling
- Fully Connected layers or Classification

4.1 Convolution Operation

One of the basic building blocks of CNN is the convolution layer. The convolution layers carry the main portion of the computational load. This layer does the dot product between the matrix of learnable parameters known as the kernel and matrix with a limited receptive field. The height and width of the kernel will be spatially small. However, the depth of the kernel extends up to three RGB channels of the input image. While performing the forward pass, the kernel slides through the image's height and width, which is responsible for the image representation of a particular receptive region (Smeda, K., 2019).

4.2 Non-Linearity

To achieve the non-linearity in the CNN network, Rectified Linear Unit (ReLU) activation function is used. ReLU stands for a rectified linear unit. Once the feature set is obtained in the next step, they are forwarded to the ReLU layer (Smeda, K., 2019). It operates on each and every element of the feature set and set the value of all negative pixels to zero. It is represented in Figure 4.2.

4.3 Pooling Layer

Like the convolutional layer, the pooling layer is used to reduce the spatial size of the convolved features obtained after convolution operation. Applying pooling on the obtained features reduces the computational power required for data processing using dimensionality reduction. Moreover, it helps in effective

4.4 Fully Connected Layers or Classification

The fully connected layer is a typical multi-layer perceptron that uses softmax as an activation function in the output layer. The neurons in the fully connected layer are fully connected with the neurons in the preceding and succeeding layer. For the input RGB image, the output of the convolutional and pooling layer delivers the high-level features. Based on the training dataset provided to the network, the fully connected layer utilizes these features for input classification into different classes (Smeda, K., 2019). Also, fully connected layers are usually used for learning a non-linear combination of the features obtained from the previous layers. The working of fully connected layers is described in Figure 4.4.

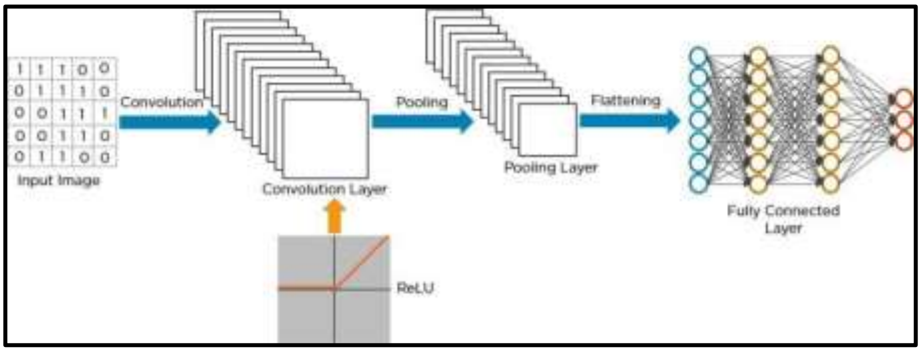

Figure 2 Fully Connected Layers for Classification

V. PROPOSED OUTPUT

At first machine learning algorithm was implemented to detect plant disease. This was done in two phases. I.) Implementing ML algorithm on potato dataset. II) Implementing ML algorithm on the entire dataset.

I. Implementing ML algorithm on potato dataset.

The dataset which is considered in the proposed work is an openly accessed dataset & it was randomly divided into the training dataset consists of 1820 images and the testing dataset consists of 780 images. The otsu algorithm was utilized for the binary image segmentation and infected region identification this was done with the help of preparing an image mask. The Gray Level Co- occurrence Matrix is the main tool that implements the concepts learned from extracted features. utilized for feature extraction, & multi-class support vector machine(SVM) methodology was utilized for the classification of potato leaves. The model derived is evaluated using certain evaluation metrics: precision, recall, F1-score, and accuracy

4.1.1 Input Image:

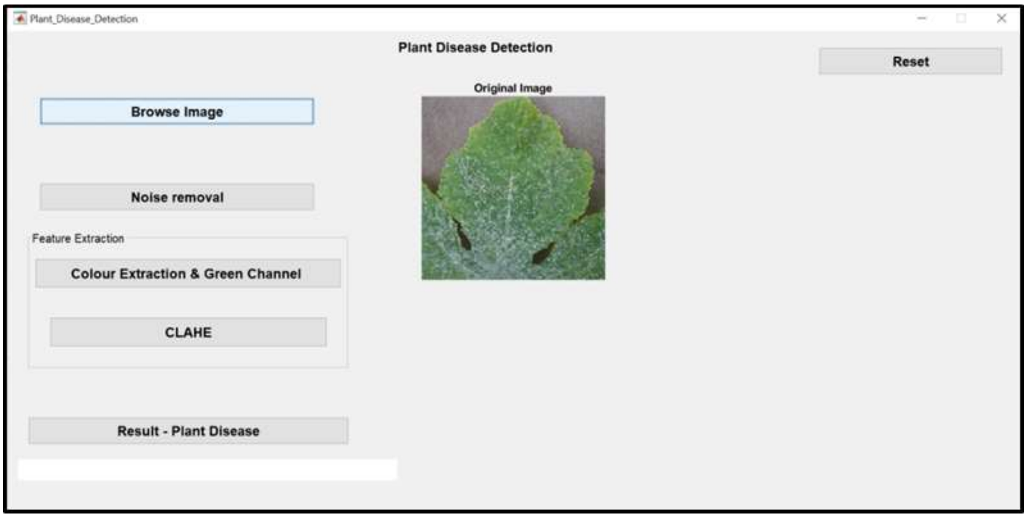

Fig. 3 GUI for Input Image for Plant Disease detection using CNN

4.1.2 Noisy Image :

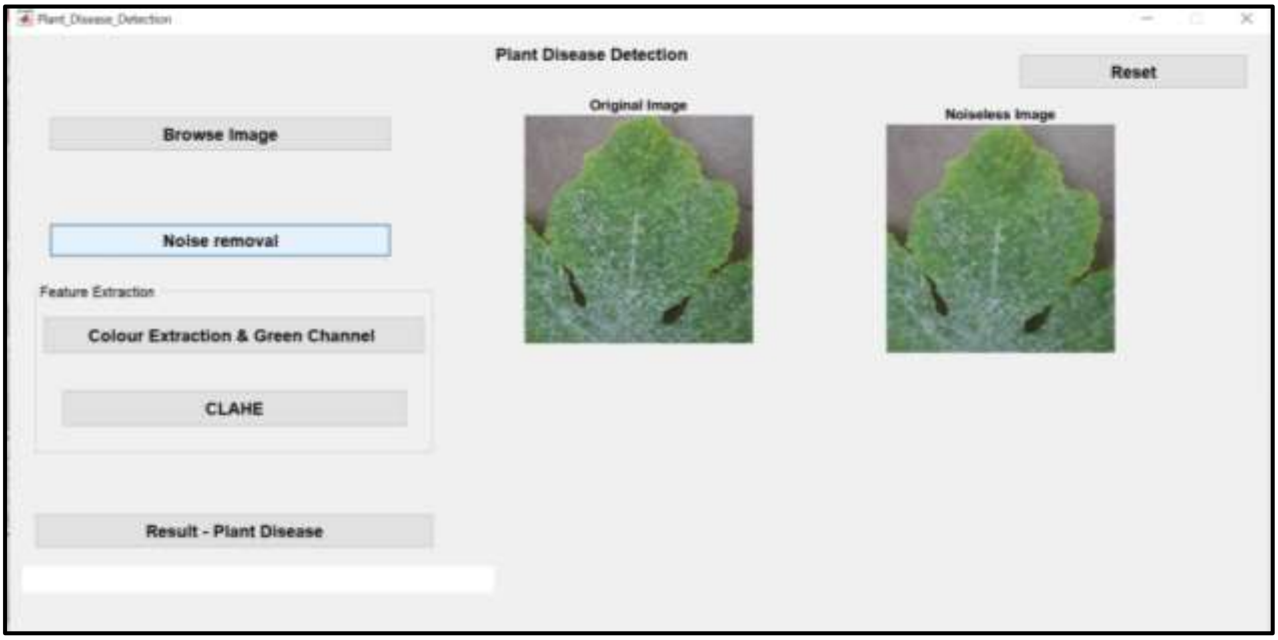

Fig. 4 GUI for Noisy Image for Plant Disease detection using CNN

4..1.3 : Colour extraction & green channel

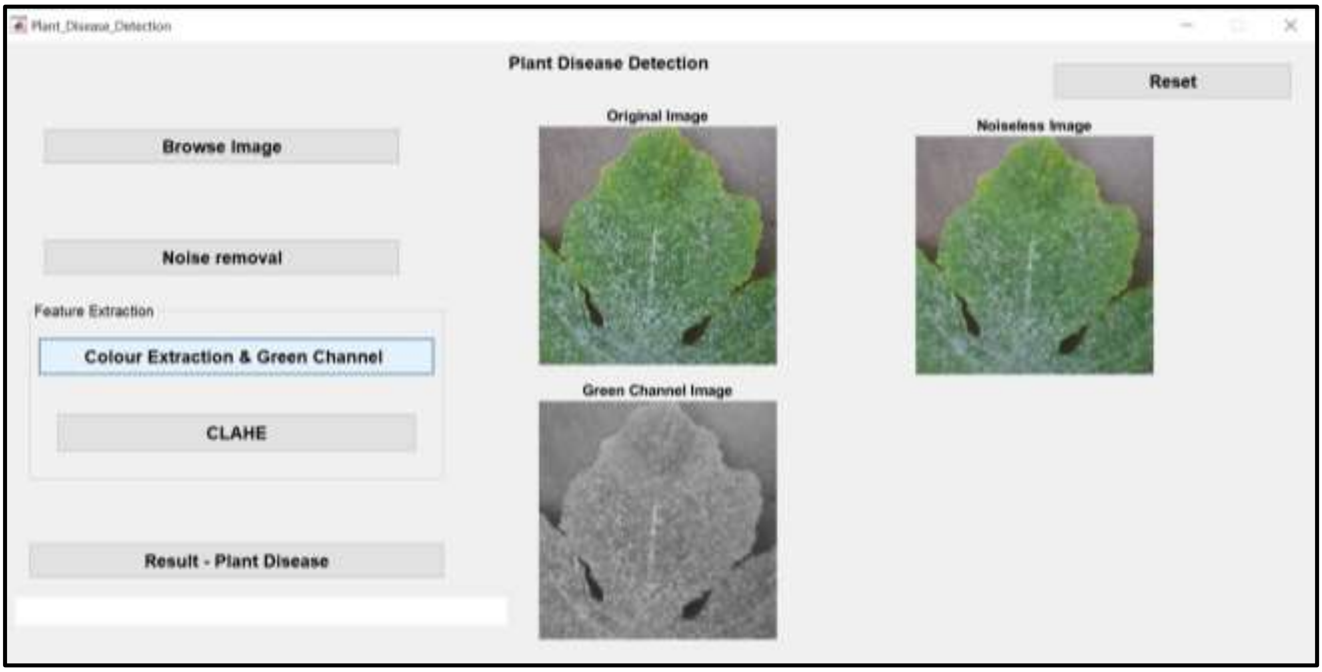

Fig. 5 GUI for Green Channel Extraction for Plant Disease detection using CNN

4.1.4 CLAHE :

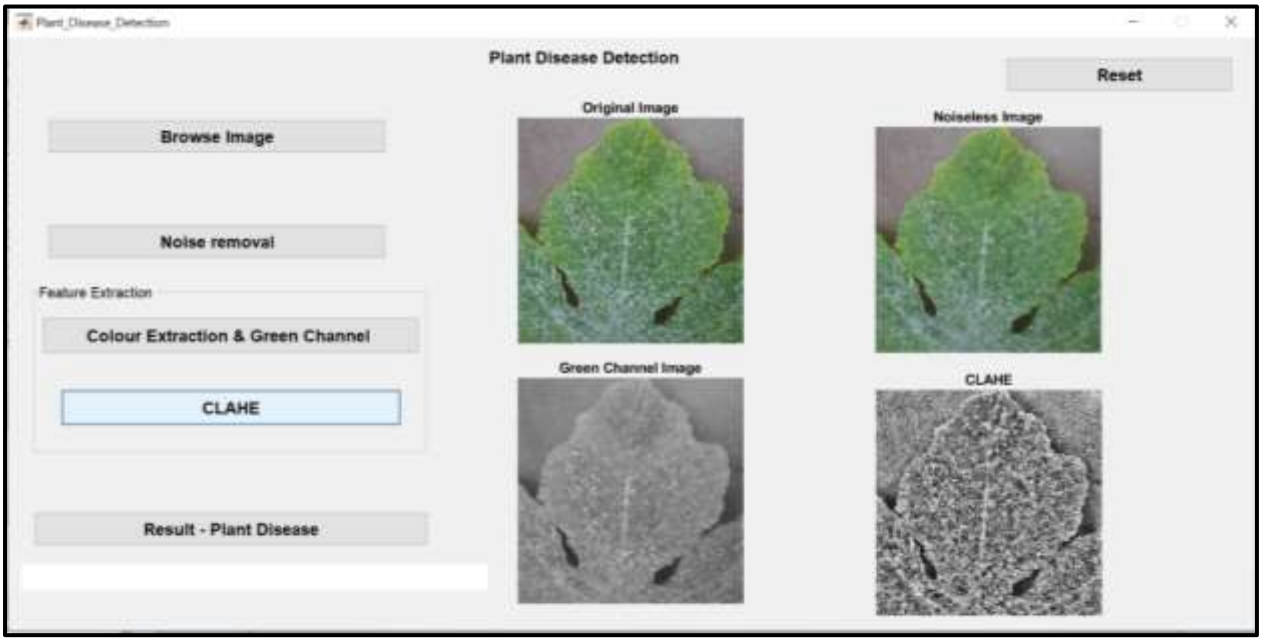

Fig. 6 GUI for CLAHE for Plant Disease detection using CNN

4.2 Convolutional Neural Network Code:

The convolutional neural network (CNN) is a class of deep learning neural networks and works by extracting features from the images. The role of CNN is to reduce the images into a form that is easier to process, without losing features critical towards a good prediction

CNN consists of the following:

Hidden layers consist of convolution layers, ReLU (rectified linear unit) layers, the pooling layers, and a fully connected Neural Network.

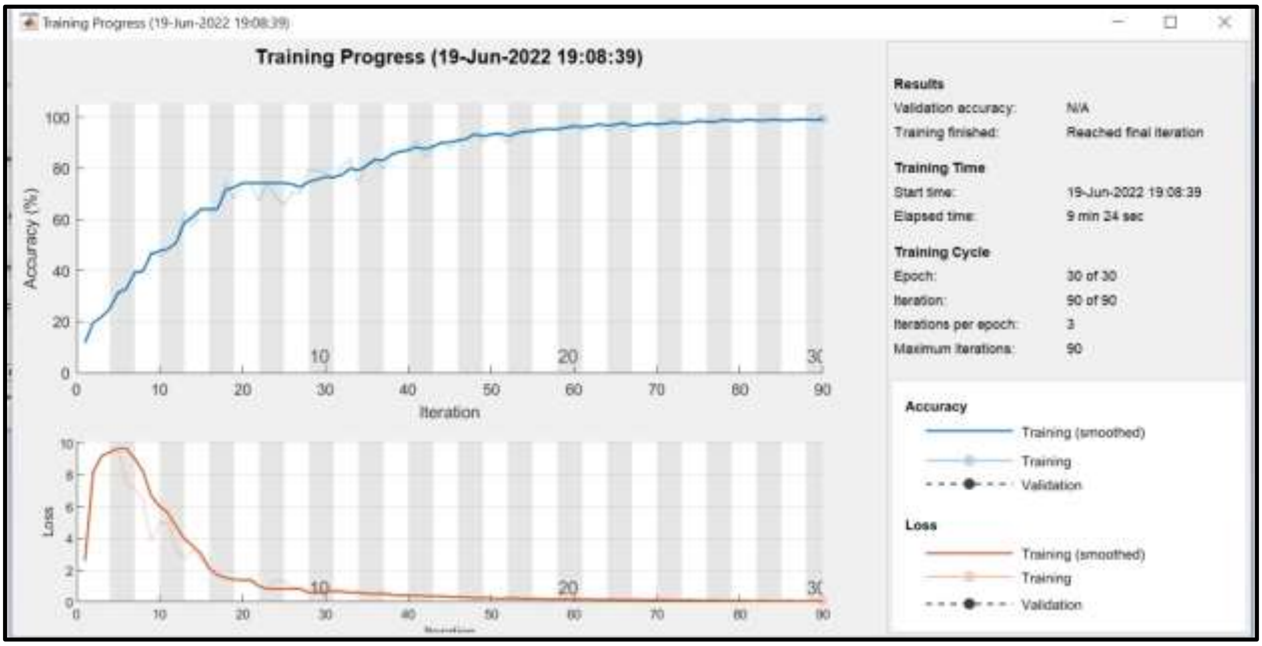

Fig. 7 Training Accuracy And Validation Accuracy Graph.

In most of the researches, the PlantVillage dataset was used to evaluate th performance of the DL models. Although this dataset has a lot of images of several plant species with their diseases, it was taken in the lab. Therefore, it is expected to establish a large dataset of plant diseases in real conditions.

CONCLUSION

The machine learning algorithm initially employed in the proposed work is Support Vector Machine (SVM). SVM performed well when the number of detection categories was limited. However, as the number of disease categories increased, its accuracy diminished. To address this, the proposed framework leverages transfer learning, a cutting-edge approach that enhances model performance while requiring minimal and faster training phases. This approach has proven highly effective in the proposed framework.

The proposed framework achieves superior accuracy using three models, with the Convolutional Neural Network (CNN) based transfer learning model demonstrating slightly higher efficiency compared to the others. This framework is proficient in multi-class classification, effectively identifying various diseases and distinguishing healthy leaves in crops such as pepper, potato, and tomato.

LIMITATIONS:

1. The three most crucial points in the selection of any transfer deep learning model are:
   - The proposed framework is utilized for the classification of diseases across the various species of crops.
   - The proposed framework utilized the concept of deep learning.
   - The proposed framework also adopted a trending research concept of transfer learning and able to achieve a better efficient model. If any of these three are neglected will result in a negative transfer or overfitting problem.
2. Machine learning models are not very efficient in predicting diseases from leaf images when the no of categories is increased.

FUTURE SCOPE:

1. The disease detection system can be integrated in cloud system for efficient result processing.
2. Integration of automated disease detection system with sensos to measure soil product.

REFERENCES:

[1] Sardogan, M., Tuncer, A., and Ozen, Y.: Plant Leaf Disease Detection and Classification Based on CNN with the LVQ Algorithm. In: 3rd Int. Conf. Comput. Sci. Eng. (2018) 382– 385

[2] Wallelign, S., Polceanu, M., and Buche, C.: Soybean plant disease identification using a convolutional neural network. In: Proc. 31st Int. Florida Artif. Intell. Res. Soc. Conf. FLAIRS 2018 (2018), 146–151

[3] Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., and Stefanovic, D.: Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification. Comput. Intell. Neurosci. 2016(2016)

[4] Fuentes, A., Yoon, S., Kim, S. C., and Park, D. S.: A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. Sensors (Switzerland) 17 (2017).

[5] Arivazhagan, S. and Ligi, S. V.: Mango Leaf Diseases Identification Using Convolutional Neural Network. Int. J. Pure Appl. Math., 120(2018) 11067–11079

[6] Oppenheim, D. and Shani, G.: Potato Disease Classification Using Convolution Neural Networks. Adv. Anim. Biosci., 8 (2017), 244–249

[7] Barbedo, J. G. A.: Factors influencing the use of deep learning for plant disease recognition. Biosyst. Eng., 172 (2018) 84–91

[8] Brahimi, M., Boukhalfa, K., and Moussaoui, A.: Deep Learning for Tomato Diseases: Classification and Symptoms Visualization. Appl. Artif. Intell., 31 (2017) 299–315

[9] Shrivastava, V. K., Pradhan, M. K., Minz, S., and Thakur, M. P.: Rice plant disease classification using transfer learning of deep convolution neural network. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch., 42 (2019) 631–635

[10] Ozguven, M. M. and Adem, K.: Automatic detection and classification of leaf spot disease in sugar beet using deep learning algorithms. Phys. A Stat. Mech. its Appl., 535 (2019) 122537

[11] Akshay K, Vani M, “Image Based Tomato Leaf Disease Detection”, Proceedings of the 10th International Conference on Computing, Communication andNetworking Technology, Kanpur, India, 2019.

[12] Sharada P Mohanty, David P Hughes, and Marcel Salath, “Using deep learning for image-based plant disease detection”. In: Frontiers in plant science 7 (2016), p. 1419.

[13] S. D. Khirade and A. B. Patil. “Plant Disease Detection Using Image Processing”, Proceeding of the International Conference on Computing Communication Control and Automation. Feb. 2015, pp. 768–771