Plant Disease Detection A Review on Various Image Processing Technologies and ML

Reenu Susan Joseph
Department Of CSE
College Of Engineering Kidangoor
Kottayam, India

Nisha C A
Department Of CSE
College Of Engineering Kidangoor
Kottayam, India

Asha Vijayan
Department Of CSE
College Of Engineering Kidangoor
Kottayam, India

***Abstract*—Plants are an inevitable part of our life. They are suffering from various diseases, which affect almost all parts of plants such as Root, Stem, Leaf, and Fruit. Proper Disease diagnosis of a plant is vital and it saves time and money. It is challenging to monitor plant diseases manually. In India, 70% of the Indian economy depends upon agricultural yield. It is estimated that various pests (insects, weeds, nematodes, animals, diseases) each year cause crop yield losses of 20-40%. For this reason, it's crucial for farmers to find out the crop diseases so they can manage them flawlessly. Most of the plant diseases are identified by its symptoms by it's color, texture, shape etc. Many researchers have developed systems based on various techniques of image processing. Identifying and classifying plant disease by using appropriate image processing and machine learning approaches in the agriculture field is crucial. Thus, A systematic survey based on various advanced techniques in Image processing and Deep learning to detect various plant diseases have been done.**

Keywords—Image processing; feature extraction; Deep learning; Neural network; classifier.

I. INTRODUCTION

Plants are important in our life. They are the primary source for food. Living things including humans need plants. Thus the plants vital for ourselves suffer various diseases from root to leaves. in the agriculture sector each year crop yield losses of 20-40% Due to Diseases and excess use of pesticides. To resolve this problem and increase crop yield can be done by preventing disease spread or early detection of disease. The identification of diseases may reduce the use of pesticides. Most diseases are identified by its features, i.e.; its color, texture and shape etc., these features help to identify the Exact disease. Even though some challenges faced in these techniques such as environmental factors like shadows, noise, illuminations, inferences etc. can be resolved by the various image enhancement techniques. The main challenge is the selection of the appropriate method. Because there is no general technique to detect all diseases of various plants and need more research to achieve it. Thus a detailed survey based on advanced deep learning and image processing methods were performed.

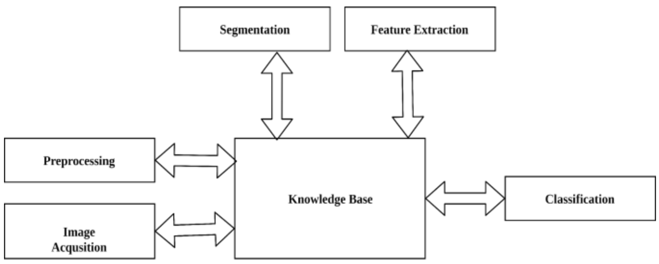

Fig 1.1 Fundamental steps in image processing

Fig 1.1 Shows the fundamental steps in image processing. Knowledge base in our problem is plant disease detection. Captured images are being preprocessed to remove noise or other interference as well as enhancing the given image to provide accurate detection. Segmentation is an important stage of the image recognition system, because it extracts the objects of our interest for detection. Feature extraction removes the irrelevant data from the data set. Final step is classification, categorizing the given image according to our interest. Here to find out the given image belongs to either disease or healthy category. Deep learning is representation learning: the automated formation of useful representations from data. At a high-level, neural networks are either encoders, decoders, or a combination of both: Encoders find patterns in raw data to form compact, useful representations. Decoders generate high -resolution data from those representations. The generated data is either new examples or descriptive knowledge.

Many researchers proposed various Image Processing and Deep learning techniques. Image processing is most commonly used for detecting plant diseases. Machine learning approaches are suitable for identification with uniform background and ideal lab environment. Also, advancement in deep learning methods can be used for identification of plant diseases.

II. LITERATURE SURVEY

There are various techniques for feature extraction. Selecting relevant features from the large set of features can be done by image processing based, machine learning based or deep learning based methods.

A. Image Processing based

G. C. Khadabadi et. al.[1] proposed a system to identify and classify carrot diseases using Discrete Wavelet Transform attained accuracy 88%. Canny edge detection method used as pre-processing to remove noises. This is the most widely used and effective method for edge detection.

Feature extraction done by DWT. DWT is a wavelet transform for which the wavelets are sampled at discrete intervals. By applying DWT the approximate coefficient of the transformed image is obtained. The transformed coefficients are used to calculate statistical features. Mean, Standard deviation, skewness, kurtosis are the various features considered for the classification. The extracted features of each carrot image are stored in a feature vector. These feature vectors used for identification and classification. Probabilistic Neural Network used for classifying carrot diseases. PNN is derived from the Radial Basis Function (RBF) Network. PNN has 3 layers: the first layer is the input layer, accepts input vectors which computes distances from the input vector to the training input vectors and produces a vector whose elements indicate how close the input is to a training input. The second layer is Radial Basis Layer, which evaluates vector distances between input vector and row weight vectors in weight matrix and distances are scaled by Radial Basis Function non-linearly. Third layer is a competitive layer that produces a classification decision. A class with maximum probabilities will be assigned by 1 and other classes will be assigned by 0. Need more testing and training because this method is applied only in a single image of carrot disease and tested 50 carrots. Hence we cannot predict the exact accuracy based on a single image of the proposed system.

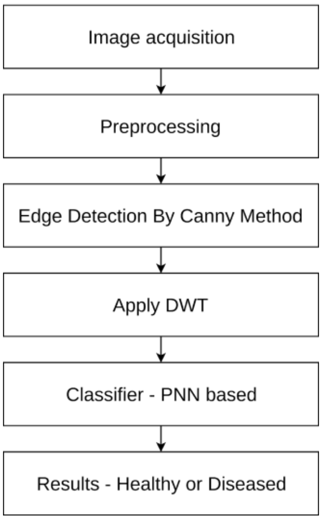

Fig 2.1 General Diagram

S. S. Chouhan et.al [2] introduced a method named Bacterial foraging optimization based Radial Basis Function Neural Network (BRBFNN) for automatic plant diseases. Bacterial foraging technique is a soft computing method which improves speed and accuracy. It also assigns optimal weight to Radial Basis Function Neural Network (RBFNN). RBFNN finds optimal regions of disease present on leaves. Region based algorithm used for feature extraction. The performance of proposed work measured in terms of specificity and sensitivity, is 0.8231 and 0.8357 respectively.

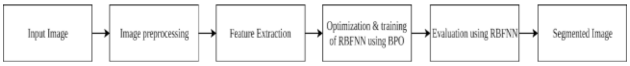

Fig 2.2 BRBFNN for automatic plant leave disease detection

A. Chen et.al [3] proposed a model based on Feature Enhancement and DMS-Robust Alexnet to identify maize leaf diseases attained higher accuracy rate of 98.62% .This paper designs a maize leaf feature enhancement framework which is capable of enhancing the feature of images under a complex environment. DMS-Alex Net is a neural network designed to improve the capabilities of feature extraction. Features are selected using the WT-DIR algorithm. The main difference with AlexNet, DMS-AlexNet combines dilated convolution and multi-scale convolution. PRelu activation function is used instead of Relu or Sigmoid activation function, and AdaBound is taken as the optimizer of the network to improve the effectiveness of learning convergence for the network. Therefore, it is considered best to the identification of maize leaf disease images with desired output.

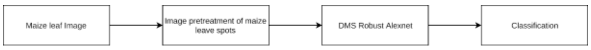

Fig 2.3 DMS Robust AlexNet for maize leave disease detection

B. Machine Learning based

A. Khan et.al.[4] proposed a system to detect plant diseases at earlier stages and implemented for apple disease identification and detection based on strong Correlation and Genetic Algorithm Based Feature Selection. This method uses three pipeline procedures including image enhancement, lesion spot segmentation, and prominent features selection & classification. The initial step is image enhancement done by a hybrid method - conjunction of 3D Box filtering, de-correlation, 3D Gaussian filter, and 3D Median filter. Next step is segmentation of the lesion spot is segmented by strong correlation method and the results optimized by fusion of expectation maximization(EM) segmentation. Finally, feature extraction extract color, color histogram and local binary pattern features from the diseased image and the features are optimized using genetic algorithms and classified by one v/s all m-SVM.

D. Zhang et.al [5] proposed a method for rapid detection of Fusarium Head Blight (FHB), a disease found in wheat using a two level segmentation and width mutation counting algorithm proposed for wheat ear counting. K-Means clustering algorithm used for rough segmentation and Random forest classifier used for fine segmentation and feature selection. They created their own Data set by collecting images from real field environments and obtained accuracy of 93%. The analysis of the proposed algorithm can correctly evaluate the efficacy of fungicides for controlling the wheat FHB disease under the field conditions.

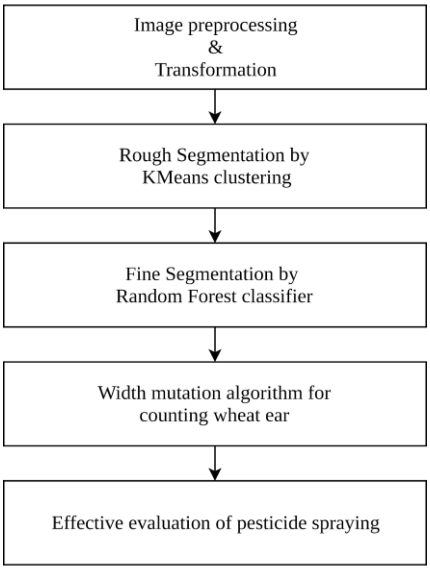

Fig 3.1 Flowchart of Fusarium Blight Detection method

T. N. Pharm et.al.[6] proposed a method for multi-class mango leaf disease classification using deep neural networks. The higher quality image is enhanced using contrast enhancement method, then using wrapper-based feature selection approach using an Adaptive Particle-Grey Wolf meta heuristic (APGWO) was performed to select features out of the originally proposed features. The selected features are used as input for the MLP for the classification task. Finally their proposed method is compared with three popular CNN methods namely AlexNet, VGG16, ResNet-50 and obtained higher accuracy of rate 89.41%. Also these 3 models are enhanced using Transfer learning of accuracy rate 84.88%.

C. Deep Learning based

G. Zhou et.al.[7] proposed a method for detection of rice disease based on fusion of FCM-KM and faster CNN which detects three types of rice diseases namely Rice blast, Bacterial Blight and Sheath Blight obtained accuracy of 97.2%. Address the image noise and interference problem with 2DFM-AMMF noise reduction. Multilevel median filtering saves edge and details but is weak in noise suppression. This problem is resolved by introducing 2Dimensional filter mask combined with Multilevel median filtering(2DFM-AMMF) by adding 2Dimensional filtering mask based on the multilevel median filter. Here the size and shape of the filtering window is determined according to the image information. Faster 2D-Otsu algorithm achieves excellent results by removing the background interference. Otsu algorithm is the one of the best threshold segmentation algorithms which perform excellent segmentation. Optimized K-Means clustering algorithm is proposed. K-Means clustering algorithm optimized by applying dynamic population firefly algorithm. Overcomes the local optimization problem by chaos theory and finds the optimized k value by Max-Min distance algorithm. Faster RCNN used for identifying three rice diseases. Faster RCNN precision and detecting speed is higher compared to other methods. The drawback of Faster RCNN is generating a larger number of boundary boxes. RPN used here for feature selection.

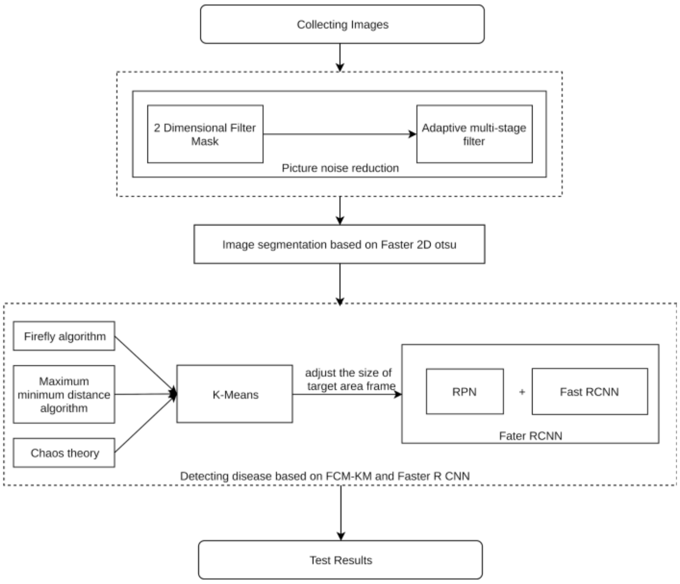

Fig 4.1 General Architecture of rice disease based on fusion of FCM-KM and faster CNN

G. Yang et.al[8] proposed a novel model, self - supervised mechanism to construct a fine-grained classification model of tomato image health/diseases, which can make better use of informative regions of tomato image in self-supervised learning without manual annotation such as bounding boxes/parts. Model consists of 3 networks, a Location network, a Feedback network, and a Classification network, named LFC-Net. Location network locates the informative region on tomato leaves whereas feedback network optimizes the Location network by computing the confidence value of 'm' informative region. Classification network predicts the leaf is either healthy or diseased. The optimized image from the feedback network and previously trained images are concatenated for classification. ResNet-50 used for feature extraction and batch normalization .Also a mechanism for locating tomato regions was proposed. This Multi-network collaboration model attains a high accuracy rate of 99.7%.

S. Huang et.al[9] proposed a detection method for Peach diseases based on Asymptotic non-local means algorithm and fusion of parallel convolution neural network (PCNN) with extreme Learning machine(ELM) optimized by Linear Particle Swarm Optimization. Detects mainly 4 peach diseases - Brown rot, Black spot, Anthracnose, Scab and Healthy peach with accuracy rate 89.02%, 90.56%, 85.37%, 86.70% and 89.91% respectively. So an effective method for peach disease as well as other fruit diseases. But it also consumes a lot of time and cost. This method has some deficiencies in Local area defect extraction and classification accuracy of two similar diseases.

P. Jiang et.al[10] proposed a real time detection of apple disease based on improved convolutional networks. This method detects common apple disease such as Alternaria leaf spot, Brown spot, Mosaic, Grey spot and Rust. In order to ensure the capability of the model, created a data set called Apple Leaves Data set consisting of laboratory and real time images using data augmentation method and image annotation technologies. Real time detection model based on SSD (Single Shot Multibox Detector). First the pre-network modified i.e. , VGGNet modified to VGG-INCEP(VGG net with Inception module) by introducing GoogleNet Inception module to improve the Extraction performance for multi-scale disease spots. Rainbow concatenation method is implemented in R-SSD to improve the detection accuracy for the small objects as well as to extract features . They ensure strong detection performance & robustness to the proposed model. Automatically identify the features of diseased apple images and also detect the 5 types of diseases with a higher accuracy rate of 78.80%.

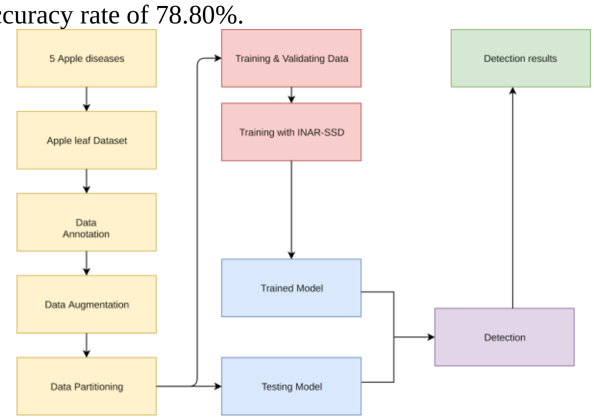

Fig 4.2 General Architecture of INAR-SSD method

J. Sun et.al [11] proposed a multi-scale fusion instance detection method based on CNN to detect plant diseases under complex background of the field which consist of three major parts - Data set preprocessing, fine tuning network, and detection. Data set preprocessing part uses improved retinex, solves the high intensity light problems. In Fine tuning network, improved RPN is used to adjust the anchor box of diseased leaves. Detection takes the optimized anchor as input and detects diseased leaves. Transmission module is also used to transfer the relevant anchor information in the fine tuning network to the detection module and improves detection efficiency. Thus the model attains a higher accuracy rate 91.83%

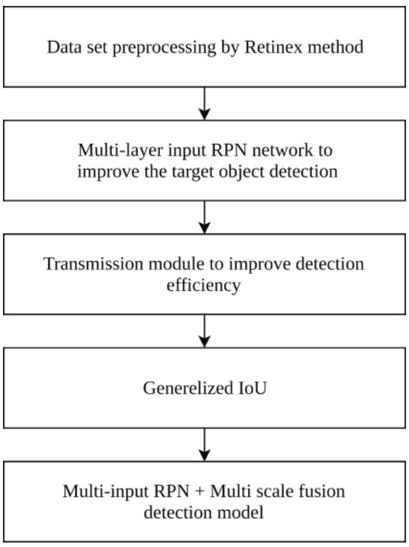

Fig 4.3 Flowchart - Multi scale fusion detection model

Table I shows the plant diseases and various techniques used in these papers according to the ascending order of the year. The impact of the advancement of deep learning technologies and neural networks in recent years are also affected in the research in plant disease detection.

TABLE I. VARIOUS TECHNIQUES AND DISEASES

Author	*Disease*	*Technique used*
G. C. Khadabadi,et. al [1] (2015)	*Carrot Diseases*	*DWT*
S. S. Chouhan, et al[2] (2018)	*Plant leaf Diseases*	*BRBFNN*
A. Chen et al[3] (2020)	*Maize leaf Diseases*	*WT - DIR Algorithm, DMS Robust Alexnet*
G. Zhou, et al [7] (2019)	*Rice blast, Bacterial blight, Sheath blight*	*Faster RCNN*
M. A. Khan et al [4] (2019)	*Black rot , Apple scab , Apple rust*	*Genetic Algorithm, one v/s all m-SVM*
P. Jiang et al[8](2019)	*Alternaria leaf spot, Brown spot, Mosaic, Grey spot, Rust.*	*INAR -SSD*
T. N. Pham et al [6] (2020)	*Anthracnose, Gall Midge, Powdery Mildew*	*ANN*
D. Zhang ,et al [5] (2020)	*Fusarium Head Blight*	*KMeans, Random forest classifier*
S. Huang et al [9] (2020)	*Brown rot , Black spot , Anthracnose, Scab*	*PCNN, ELM based IPSO*
J. Sun, et al [10] (2020)	*Maize leaf Diseases*	*Multi layer RPN*
G. Yang, et al [11] (2020)	*Tomato Diseases*	*LFC-net*

Table II shows the various feature extraction methods and classifier or detection techniques used in the literature. The appropriate methods for selecting features from a large set of features is crucial and the proper detection method or classifier improves the accuracy of the detection.

TABLE II. VARIOUS FEATURE EXTRACTION METHODS AND CLASSIFIERS/ DETECTION TECHNIQUES

Author	*Feature extraction method*	*Classifier/Detection method used*
G. C. Khadabadi,et. al [1] (2015)	*DWT*	*Probabilistic Neural Network*
S. S. Chouhan, et al[2] (2018)	*Region Growing Algorithm*	*BRBFNN*
A. Chen et al[3] (2020)	*WT-DIR algorithm*	*DMS-Robust AlexNet*
M. A. Khan et al [4] (2019)	*Genetic Algorithm*	*one v/s all m-SVM*
D. Zhang ,et al [5] (2020)	*Random forest classifier*	*Random forest classifier*
T. N. Pham et al [6] (2020)	*Wrapper Based method*	*ANN*
G. Zhou, et al [7] (2019)	*RPN*	*Faster RCNN*
P. Jiang et al[8](2019)	*GoogLeNet Inception module with Rainbow concatenation method*	*INAR-SSD*
S. Huang et al [9] (2020)	*location and feedback network*	*Classification network*
J. Sun, et al [10] (2020)	*PCNN*	*ELM based IPSO*
G. Yang, et al [11] (2020)	*MULTI-LAYER INPUT RPN NETWORK*	*fusion of RPN, Transmission module and softmax layer*

III. COMPARATIVE STUDY

Table III deploys the comparison of literature in usage of metrics and data set used.

TABLE III. COMPARISON OF LITERATURE IN USAGE OF METRICS AND DATA SET

Author	*Metrics used*	*Data set*
G. C. Khadabadi,et. al [1] (2015)	*Accuracy*	*Own Data set*
S. S. Chouhan, et al[2] (2018)	*Specificity, Sensitivity*	*Plant Village Data set*
A. Chen et al[3] (2020)	*Accuracy*	*Plant Village, Crop disease dataset*
M. A. Khan et al [4] (2019)	*Accuracy*	*Plant Village data set*
D. Zhang ,et al [5] (2020)	**PA, Precision, F-Measure**	*Own Data set*
T. N. Pham et al [6] (2020)	**Accuracy, Precision, F1-Score, Recall**	*Own Data set*
G. Zhou, et al [7] (2019)	*Accuracy*	*Own Data set*
P. Jiang et al[8](2019)	**mAP**	*Own Data set*
S. Huang et al [9] (2020)	*Accuracy*	*Orchard Data set*
J. Sun, et al [10] (2020)	**mAP**	*NLB Data set*
G. Yang, et al [11] (2020)	*Accuracy*	*Plant Village Data set*

One of the key challenges is the availability of data sets. From the analysis of the table it is clear that most of the researchers created their own data set, which attains higher accuracy rate and chose the evaluation parameter in terms of accuracy

IV. CONCLUSION

In this paper we have done a survey on different plant diseases and various technologies used to detect these diseases. From the review the image processing and machine learning techniques have more potential to detect disease. Therefore, researchers have begun to study about detecting plant leaf diseases using neural networks recently. There are effective methods for specific disease detection of plants. So further research is required for an effective method for general plant disease detection which can extend to assist farmers to identify crop diseases accurately and hence increase the crop yield.

REFERENCES

[1] G. C. Khadabadi, A. Kumar and V. S. Rajpurohit, "Identification and classification of diseases in carrot vegetable using Discrete Wavelet Transform," 2015 International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT), Mandya, India, 2015, pp. 59-64, doi: 10.1109/ERECT.2015.7498988.

[2] S. S. Chouhan, A. Kaul, U. P. Singh and S. Jain, "Bacterial Foraging Optimization Based Radial Basis Function Neural Network (BRBFNN) for Identification and Classification of Plant Leaf Diseases: An Automatic Approach Towards Plant Pathology," in IEEE Access, vol. 6, pp. 8852-8863, 2018, doi: 10.1109/ACCESS.2018.2800685.

[3] M. Lv, G. Zhou, M. He, A. Chen, W. Zhang and Y. Hu, "Maize Leaf Disease Identification Based on Feature Enhancement and DMS-Robust Alexnet," in IEEE Access, vol. 8, pp. 57952-57966, 2020, doi: 10.1109/ACCESS.2020.2982443.

[4] M. A. Khan et al., "An Optimized Method for Segmentation and Classification of Apple Diseases Based on Strong Correlation and Genetic Algorithm Based Feature Selection," in IEEE Access, vol. 7, pp. 46261-46277, 2019, doi: 10.1109/ACCESS.2019.2908040.

[5] D. Zhang, Z. Wang, N. Jin, C. Gu, Y. Chen and Y. Huang, "Evaluation of Efficacy of Fungicides for Control of Wheat Fusarium Head Blight Based on Digital Imaging," in IEEE Access, vol. 8, pp. 109876-109890, 2020, doi: 10.1109/ACCESS.2020.3001652.

[6] T. N. Pham, L. V. Tran and S. V. T. Dao, "Early Disease Classification of Mango Leaves Using Feed-Forward Neural Network and Hybrid Metaheuristic Feature Selection," in IEEE Access, vol. 8, pp. 189960-189973, 2020, doi: 10.1109/ACCESS.2020.3031914.

[7] G. Zhou, W. Zhang, A. Chen, M. He and X. Ma, "Rapid Detection of Rice Disease Based on FCM-KM and Faster R-CNN Fusion," in IEEE Access, vol. 7, pp. 143190-143206, 2019, doi: 10.1109/ACCESS.2019.2943454.

[8] G. Yang, G. Chen, Y. He, Z. Yan, Y. Guo and J. Ding, "Self-Supervised Collaborative Multi-Network for Fine-Grained Visual Categorization of Tomato Diseases," in IEEE Access, vol. 8, pp. 211912-211923, 2020, doi: 10.1109/ACCESS.2020.3039345.

[9] S. Huang, G. Zhou, M. He, A. Chen, W. Zhang and Y. Hu, "Detection of Peach Disease Image Based on Asymptotic Non-Local Means and PCNN-IPELM," in IEEE Access, vol. 8, pp. 136421-136433, 2020, doi: 10.1109/ACCESS.2020.3011685.

[10] P. Jiang, Y. Chen, B. Liu, D. He and C. Liang, "Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks," in IEEE Access, vol. 7, pp. 59069-59080, 2019, doi: 10.1109/ACCESS.2019.2914929.

[11] J. Sun, Y. Yang, X. He and X. Wu, "Northern Maize Leaf Blight Detection Under Complex Field Environment Based on Deep Learning," in IEEE Access, vol. 8, pp. 33679-33688, 2020, doi: 10.1109/ACCESS.2020.2973658.