Proceedings of the 9th International Conference on Agriculture, Vol.7, Issue 2, 2022, pp.31-40

ISSN 2682-7018online
DOI https://doi.org/10.17501/26827018.2023.7202

A STUDY OF PLANT DISEASE DETECTION AND CLASSIFICATION BY DEEP LEARNING APPROACHES

Sandhya Rani D* and Shyamala K

Computer Science and Engineering, Faculty of Engineering, Osmania University, India

Abstract: Agriculture is an important sector in India. In developing countries like India, employment opportunities are provided on large scale. Agriculture comprises many crops and nearly 70% of the population is depending on agriculture. In this study, we have collected more than 30 papers that are published on plant leaf disease identification and classification using Deep Learning algorithms. Plant leaf disease can be occurred by various factors like bacteria, viruses and fungi etc. Identifying the infected leaves is a demanding task for a farmer and also for a researcher. Nowadays, farmers are using pesticides on plants, as a result, it affects human health directly or indirectly and it causes economical loss. Plant Disease Detection techniques are used by the researchers in order to prevent it and the disease will be identified on time. This study will be a helpful resource for the researchers to identify the specific type of plant leaf diseases through deep learning techniques. This paper presents a detailed survey on different papers of various plant leaf diseases depending upon some important criteria like the number of classes (diseases and healthy), size of image dataset, pre-processing, image segmentation, types of classifiers, the accuracy of classifiers, etc.

Keywords: Deep learning, Plant Disease Detection, Classification, Transfer Learning, Classifiers, Image Processing

Introduction

Agriculture is the primary occupation in the world. Indian economy majorly depends on agriculture. The occurrence of Plant diseases has an adverse effect on the growth of the crop and agricultural production. In many developing countries (Savary et.al, 2012) the agricultural economy has a high financial loss due to the severe damage to planted crops in various areas. The major causes of the leaf diseases are genetic disorders, bacteria, fungi, viruses and microorganisms. The majority of the crop production is affected by various diseases like Bacterial Spot, early blight, canker, scabs, Fusarium Wilt, Black Rot, Septoria leaf spot, Mosaic virus etc (Sharada et.al, 2016)

Sometimes monitoring plant diseases manually is a very tedious task of its complicated nature and also takes a lot of time for monitoring. In some cases, farmers will not concentrate on the diseases and face trouble in identifying the diseases which lead to damage to the crop. Farmers have to mainly concentrate on the disease that occurs in a plant at every stage of a plant. This technique of disease identification requires some precautions when choosing pesticides. Due to a lack of understanding or confusion about the severity of the disease, under-usage or heavy-usage of the pesticide may cause damage to the crop. For the above-mentioned reasons, substantial research is being carried out to develop reliable methods for the diagnose diseases (Suganya et.al, 2020) This is one of the

*Corresponding Author's Email: *sandhya.mgu@gmail.com*

fundamental motivations for the proposed methodology to utilize this survey to detect and classify diseases in various plant leaves.

Plant leaf disease identification is a major research area in which deep learning techniques (Sammy et.al, NA) are widely used for the accurate classification. Deep learning is the fast-growing approach and is a sub-part of the Machine Learning family. This approach uses Convolution convolution neural network for the classification of images and gives the best accurate result in working out the real-world problems. In many models, the success rate attained with deep learning is higher than Machine Learning.

This paper is divided into five sections. Section II describes the basic knowledge of Deep Learning. Section III describes survey of Deep Learning models for the plant leaf disease classification. Section IV presents the information regarding the previous research results of various deep learning models. Section V presents general method for detection and the classification of plant leaf diseases. Section VI presents the conclusion.

Deep Learning

Deep Learning was introduced in 1943 and it is a subclass of machine learning and it is similar to Artificial Neural Networks. Deep learning is successfully applied in the domain of Agriculture. The major advantage of Deep Learning is feature extraction. Convolution neural networks (CNN) constitute a class of deep feed-forward Artificial Neural Network and it appears as the technique in many surveyed papers. All the DL models come along with their pre-trained weights, which mean that their network is already been trained by some dataset and therefore learned to provide accurate classification for an application. This process of pre-trained weights is called Transfer Learning Technique.

Deep learning (Lecun et.al, 2015) is one of the important research areas in the field of machine learning. It attained high success in different domains like Classification (Shiravastava et.al, 2019), audio processing (Noda et.al, 2015), natural language processing (Deng et.al, 2014), and Computer Vision (Baykal et.al, 2019; Sradjan et.al, 2016) etc. Among all the Deep learning approaches, convolution neural network (CNN) is one of the successful image classification techniques. These deep Convolution Neural Network models require large dataset and high computation power for efficient training. In order to overcome these limitations Transfer Learning techniques are used where the knowledge of the pre-trained Deep Convolution neural network models are used.

Convolution Neural Network

Convolution Neural Network is a multi-layer feed forward neural network and is the popular deep learning model. CNN is worked with a series of layers. The quantity of features increases as the total number of the layers increases. Convolution neural networks are mainly used for creating a computational model which work mostly on input images of unstructured type and then convert these images to the corresponding classification output labels. These belong to the class of the multilayer neural network's that can be trained to learn the required features for the classification purpose. These images required less pre-processing when compared to the traditional techniques and then perform automatic feature extraction that give the better performance.

Architectures and tools

Various popular deep learning architectures are used now-a-days where researchers build their models rather than starting from scratch. The pre-trained deep CNN models include Alex Net, VGG16, Inception V3, DenseNet169, Google Net etc. The commonly used dataset for pre-training Deep Learning architectures is Image net.

The most popular tools to the researchers to experiment with Deep Learning are Theano, Keras (it is an API on top of Tensor flow and Theano), caffee, Tensorflow, PyTorch.

Survey of Deep Learning Models for the Plant Leaf Disease Classification

Srdjan et.al, (2016) proposed a convolution neural network model. Author says that climatic conditions can change the pathogen development rate. An overall accuracy of 96.3% was achieved by the model. The advantages of this model are it is good to get a more accurate classification. The accuracy can be improved by using fine-tuning and augmentation techniques.

Aravind et.al, (2018) recommended Ttransfer Learning approach for tomato plant leaf disease classification. The pre-trained Alex Net architecture attained 97.49% of accuracy with less computational load and less execution time. Demerits are when mini-batch size is increased, the accuracy is decreased.

Mohommed et.al (2017) proposed Google Net architecture to identify the diseases in tomato crops. The trained CNN model provides an accuracy of 99.18%. GLCM is used to extract features like colour, texture, and edge automatically. Merits include nonlinearity without an explosion of the number of weights and it directly processes the raw images. The disadvantages are the size of the deep CNN model and the computation speed to be minimized.

Santhana Hari (Santhahari et.al, 2019) presented a completely new CNN architecture called Plant Disease Detection Neural Network. Batch Normalization is used to convert and normalize all the values between 0 to 1. The experimental results are collated with MobileNet50 model and this PDDNN architecture has improved accuracy by nearly 7% and attained an accuracy of 86%. Normalization technique is used during computations to reduce the time required for the calculations.

Nithish E kannan, Kaushik M (Nitish et.al, 2020) proposed a pre-trained ResNet-50 model. This model achieved 97% accuracy. Merits include augmentation and transformation techniques are implemented in order to overcome the overfitting problem. Demerits are it is time-consuming and high configuration hardware is used.

Prajwal et.al, (2018) proposed a simple Le-Net model. This model has achieved an accuracy of 94-95% by utilising minimum hardware requirements. Jia et.al (2017) presented a CNN model-VGG16 model. This model is trained with Keras or Tensorflow framework and attained an average accuracy of 89%. The advantage is SVM classifier efficiently performs non-linear Classification but fine tuning the model is time-consuming. Feature extraction is done by CNN while training.

Sardogon et.al (2018) proposed a deep learning model and Learning Vector Quantization algorithm method for tomato leaf disease identification. This method attained 86% accuracy. High accuracy is attained by transfer learning techniques. Future research needs to enhance the rate of recognition in the process of classification by using different filters and different convolutions.

Jiang and Fudong (2020) proposed ResNet model that can predict each category of disease image. The Leaky-ReLU activation function is used in this model. This Model achieved 98.3% accuracy. The main advantage of this model is Robust and is of high potential. Demerits are the simultaneous detection of multiple species and multiple diseases. Future research should be made by detecting multiple species and multiple diseases.

Vimal K.Shrivastava (2021) recommended the VGG16 model and attained an accuracy of 93.11%. Greater number of parameters is used to achieve better performance. Demerits include small and overfitting. Jithy Lijo (2020) proposed Transfer Learning models such as ResNet50, Inception V3, and DenseNet169. The highest accuracy of 98.2% is obtained with the ResNet50 model. The performance of the model is improved by aspect ratio and batch normalization.

H.sabrol (2021) presented a CNN architecture based on supervised learning approaches. Required features extracted from segmented and pre-processed images are provided as input to the CNN model, which then classify the tomato plant leaf disease with an accuracy of 97.3%.

Yang Zhang (2016) recommended a faster RCNN model with Resnet101 based architecture. This model obtained 98.54% accuracy. Merits include this approach gives fast detection speed when compared to the original faster RCNN. Alvaro Fuentes (2020) proposed three types of deep learning meta-architectures called Faster region-based convolution neural network (Faster R-CNN), Single shot multibox detector (SSD) and Region-based fully convolutional network (R-FCN) to effectively recognize different varieties of tomato plant leaf diseases and pests. Data Augmentation and feature extraction techniques are used to improve the accuracy and minimize false positives. The proposed model attained an accuracy of 83.6%.

X.Zhang (2020) proposed Google Net and Cifar10 deep learning models for nine types of Maize leaf disease recognition. The authors applied different methods like addition of dropout operations, and Rectified Linear functions, changing the pooling combinations, and decreasing the number of classifiers. The Google Net model attained an accuracy of 98.9%.

Ronnel R.Atole (2018) proposed a Transfer learning deep neural network model called AlexNet on six hundred images of rice plants. And this model attained 91.23% accuracy. Mohammed Brahimi (2018) tested different state-of-the-art CNN architectures using three learning strategies for plant disease classification on a publicly available plant village dataset. The accuracy achieved by InceptionV3 model is 99.76%. In addition, a visualization tool using saliency maps is proposed in order to better understand the CNN classification model.

Tetila et.al (2020) recommended an approach to classify six diseases of sayobean leaf. These images are collected from UAV (unmanned aerial vehicles). To achieve high accuracy, four deep neural network models are trained with different parameters for fine-tuning and transfer learning. The authors applied different methods like dropout and data augmentation during the training the network to avoid overfitting. The Inception V3 model achieved 99.04% accuracy.

Chen et.al (2020) proposed a transfer learning technique called INC-VGGN architecture to detect maize and rice plant diseases. The accuracy of 97.57% is achieved with the proposed model when compared with the original classification models. Kamal et.al (2019) proposed MobileNet model and it achieved an accuracy of 98.34% with only a few parameters than VGGNet.

Konstantinos and Ferentinos (2018) proposed VGG classification model to detect fifty eight diseases of various plants. The Torch framework is used for the training and testing processes and the model achieved 99.53% accuracy.

Hasson and Ashraf (2022) proposed a new CNN model which makes use of two pre-trained Convolution Neural Network (CNN's) namely EfficientNetB0 and DenseNet121. to identify the diseases from the corn plants. The proposed model achieved 98.56% accuracy with less number of parameters.

Sunil (2021) recommended EfficientNetV2 classification model to detect cardamom plant disease. The author used U2-Net to remove the unnecessary background of the input image by selecting multiscale features from 1724 cardamom plant leaf images of three classes namely, Colletotrichum Blight and Phyllosticta Leaf Spot and healthy category. The proposed model achieved 98.26% detection accuracy.

Discussions

Datasets

Many of the researchers experimented with the DL algorithms with having a large dataset of images that originated from publicly available datasets such as plant village. This dataset has 54309 plant leaf disease images, which covers 14 varieties of vegetable crops and fruits such as potato, soybean, raspberry, pumpkin, strawberry, tomatoes and apple, blueberry cherries, grapes, orange, peach, etc. and it includes 12 healthy crop leaf images. Nowadays, Hyper-Spectral Imaging (HIS) is actively being used for the early detection of plant diseases and it also shows better performance.

Feature Extraction

The features that are managed by feature extraction methods are shown in TABLE I.

Sl.No	FEATURES	METHOD
1	Texture	Gray Level Co-Occurrence matrix (GLCM)
2	Colour	Color Co-Occurrence matrix (CCM)
3	Shape	Minimum Enclosing rectangle (MER)
4	Colour, Texture	Color Co-occurrence matrix (CCM)
6	Other Features	Discrete Wavelet transform(DWT) Scale Invariant Feature transform(SIFT)

The Machine Learning algorithms require a feature extraction mechanism for plant leaf disease detection. However, deep Learning models doesn't require this feature extraction method. For this reason, deep learning framework is being actively used recently for many plant's leaf disease identification and classification.

Table I. Feature Extraction Method

Classification

The comparison of model accuracy is shown in TABLE II

Table Ii. Comparison Of Model Accuracy Of Related Work

Ref.No	Methodology used	No.of.Images	Accuracy
[12]	Alex Net architecture	13262	97.49%
[14]	PDDNN	15210	86%
[15]	ResNet 50	12206	97%
[23]	FasterRCNN-res101	4178	97.18%
[25]	Improved Google Net	500	98.9%
[26]	Alex Net architecture	600	91.23%
[29]	INC-VGGN	966	92%
[32]	Proposed CNN model	15408	98.56%

In Deep Learning models, training and testing should have more images achieve better accuracy.

General Method For Detection And Classification Of Leaf Disease

The block diagram of a proposed system is depicted in Figure 1 for plant leaf disease identification. The required steps in this particular method are 1. Data Acquisition 2. Data Annotation 3. Data Augmentation 4. Model Creation 5. Training and Validation 6. Classification using CNN 7. Performance Verification.

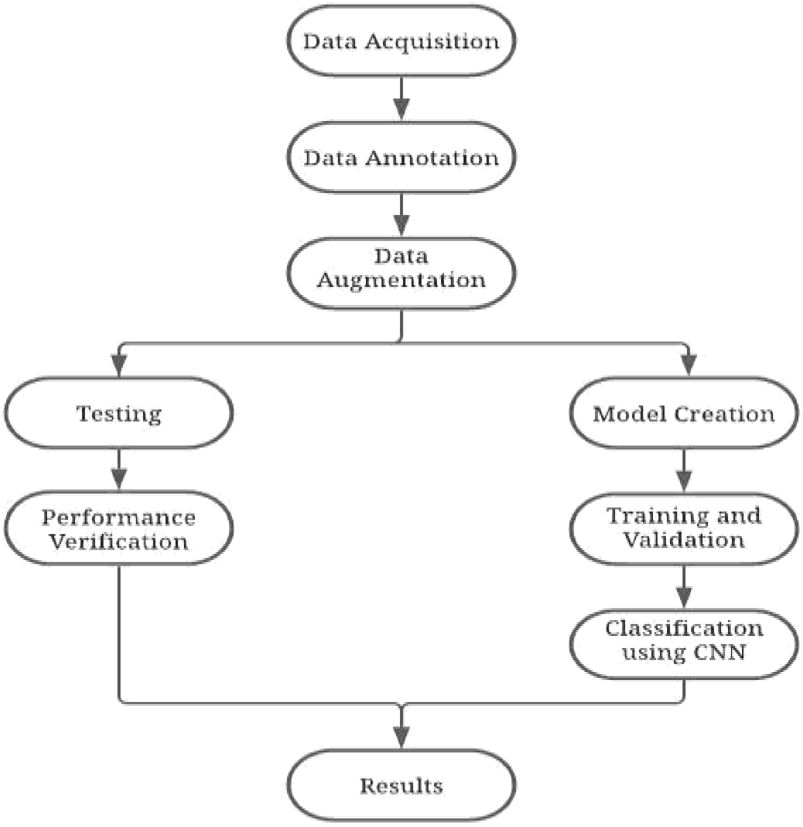

Fig.1. Block Diagram

(i)Dataset Acquisition:

The dataset contains images of several diseases in different crops. The images can be captured with the help of digital camera or they can be obtained from publicly available datasets.

(ii)Data Annotation:

Data Annotation is the process of identifying and classifying objects, elements or any other data present in images, text, and videos and audio. When an object in an image or video is classified, a label or metadata is added as a tag to that object. This process of labelling objects in any data form is called Data Annotation. It helps to recognize data in its correct form. The area of each image which contains the disease with a bounding box and the class in the dataset of images is manually annotated. Some of the diseases appear same by the state of the infection that is present. So, knowledge of detecting the type of disease is given by experienced people. This helps to detect the class of the images and diseased parts of the plant.

An example of annotated bounding box is depicted in Figure 2. The red coloured rectangular box shows the infected parts of the plant and the background area.

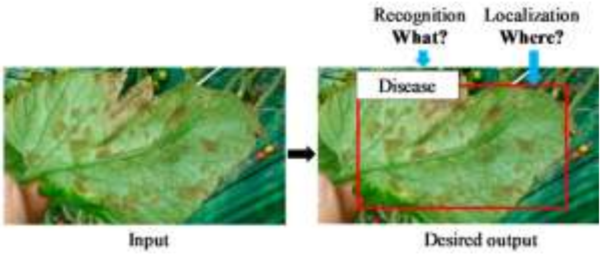

Fig.2. Recognition and localization of plant diseases

(iii) Augmentation:

Augmentation is carried on the images to increase the size of the dataset. This technique is basically used to increase the number of images in the dataset. It consists of geometrical Transformations like cropping, resizing, image rotation, affine transformation, horizontal flip, and intensity transformation (colour, brightness enhancement, contrast. The drawback of Deep Neural Network systems is overfitting problem. Overfitting is referred to as the hyperparameter selection, regularization of the system and the number of images used for training.

(iv) Model Creation:

The Transfer Learning technique is used to minimize the redundancy of developing a new deep learning model every time for various purposes. To overcome this specific problem, pre-trained network model is used that is already trained on thousands of images. Transfer learning can be implemented in two ways on Convolution Neural Network.

(v) Training the Model:

Training the model is determining good values for all the weights and bias from the labeled examples. In supervised learning, deep learning algorithms build a model by examining many examples and attempting to find a model that minimizes loss. The learning rate is initially set to 0.001. In general, different types of hyperparameters are used to train different classification models.

(vi) Classification using CNN:

At the initial stage, the dataset is divided into three datasets: the training set, the validation set, and the test set. By dividing the images randomly, so that 80% of them form the training set. The validation set is separated from test set, which is generally used for selecting parameters and eliminating over fitting. To detect and compare the accuracy of the model, the transfer learning technique is used in pre-existing neural networks. At the time of training the CNN model, after each epoch, the loss that occurred is calculated, gradients are backpropagated concerning the loss and model parameters, and the required parameters are updated with the optimizer. After retraining and modification of the training images, more precise outcomes are obtained with these CNN models.

Conclusion

In this review, a comparative analysis of different transfer learning deep CNN models is provided to detect and categorize plant leaf diseases. More than 30 relevant papers have been identified, observing the specific area, state-of-the-art of DL models employed, the data used from the different sources, data augmentation techniques and image pre-processing tasks adopted and overall performance of the model as per the observations made from the performance metrics of each paper. Our findings from different papers indicate that deep learning gives better performance. The researchers who experiment with computer vision and classification problems for solving agricultural problems can utilize this review of plant disease detection to gain benefits towards sustainable and smarter farming and secure food production. In future work, many classifiers can be build for plant disease classification and also localization of the disease regions can be proposed to help human beings to understand the disease.

References

Savary, S.; Ficke, A.; Aubertot, J.-N.; Hollier, C. Crop losses due to diseases and their implications for global food production losses and food security. Food Sec. 2012.

Sharada P.Mohanty, David P.Hughes, Marcel Salathe," Using Deep Learning for Image-Based Plant Disease Detection". In: Frontiers in plant science, Vol.7, p.1419, September 2016.

K.Suganya Devi, P. Srinivasan, Sivaji Bandhopadhyay. "H2K – A robust and optimum approach for detection and classification of groundnut leaf diseases" , Computers and Electronics in Agriculture, 2020

Sammy V.Militante, Bobby D. Gerardo, Nanette V.Dionisio,"Plant Leaf Disease Recognition using Deep Learning", IEEE Eurasia Conference on IOT, Communication Engineering,ISBN:978-1-7281-2501-5.

Y. LeCun, Y. Bengio, and G, Hinton, "Deep learning," Nature, 521(7553), 2015, 436

V. K. Shrivastava, M. K. Pradhan, S. Minz, and M. P. Thakur, "Rice Plant Disease Classification Using Transfer Learning of Deep Convolution Neural Network," ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Delhi, India, 423, 2019, pp. 631–635.

K..Noda, Y.Yamaguchi, K. Nakadai, H. G Okuno and T. Ogata, "Audio-visual speech recognition using deep learning," Appl. Intell., 42(4), 2015, pp. 722–737.

L. Deng, and D. Yu, Deep learning: methods and applications, Found. Trends Signal Process. 7(3–4), 2014, pp. 197–387.

A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolution neural networks, In Advances in neural information processing systems, Geoffrey Hinton. University of Toronto. Canada, 2012, pp. 1097–1105.

E. Baykal, H., Dogan, M. E, Ercin, S., Ersoz, and M., Ekinci, "Transfer learning with pre-trained deep convolution neural networks for serous cell classification," Multimed. Tools. Appl., 2019, pp. 1–19.

Srdjan Sladojevic, Marko Arsenovic, Andras Anderla, Dubravko Culibrk and Darko Stefanovic," Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification," Hindawi Publishing Corporation Computational Intelligence and Neuroscience, Vol. 2016, Article ID 3289801, 11 pages.

Aravind Krishnaswamy Rangarajan, Aniirudh Ramesh and Raja Purushothaman, "Tomato crop disease classification using pre-trained deep learning algorithm ", In : International Conference on Robotics and Smart Manufacturing (RoSMa2018), Vol.133, pp. 1040–1047,2018

Mohammed Brahimi, Kamel Boukhalfa & Abdelouahab Moussaoui, "Deep Learning for Tomato Diseases: Classification and Symptoms Visualization", In: International Journal of Applied Artificial Intelligence, Vol.31, pp; 299-315, May 2017.

S. SanthanaHari, M Sivakumar, P.Renuga, S.karthikeyan, S.Suriya,"Detection of Plant disease by leaf image using Convolution Neural Network", In: IEEE International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), Vellore,Tamilnadu,2019.

Nithish Kannan E, Kaushik M, Prakash P, Ajay R, Veni S, " Tomato Leaf Disease Detection using Convolution Neural Network with Data Augmentation", Proc.IEEE 5th International Conference on Communication and Electronics System(ICCES), pp.1125-1132, June 2020,

Prajwala TM, Alla Pranathi, Kandiraju Sai Ashritha, Nagaratna, B.Chittaragi,"Tomato Leaf Disease Detection using Convolution Neural Networks", Proc. 11th IEEE International Conference on Contemporary Computing (IC3), Noida, India, pp. 1-5,2-4 August 2018.

Jia Shijie, Jia Peiyi, Hu Siping, Liu Haibo,"Automatic Detection of Tomato Diseases and Pests based on leaf images". In: IEEE China Automation Congress (CAC), Jinan, China, 2017.

Sardogan, Adam Tuncer, Yunus Ozen,"Plant Leaf Disease Detection and Classification based on CNN with LVQ algorithm", Proc. 3rd International Conference on Computer Science and Engineering, pp.382-385, Herzegovina, 2018.

Ding Jiang, Fudong Li, Yuequan Yang, Song Yu,"A Tomato Leaf Diseases Classification Method Based on Deep Learning". Proc. IEEE Chinese Control and Design Conference, Aug 2020.

Vimal K.Shrivastava,Manoj K.Pradhan,"Application of Pre-trained Deep Convolution Neural Networks for Rice Plant Disease Classification",Proc.International Conference on Artificial Intelligence and Smart Systems(ICAIS),San Francisco State Univ,2021 .

Sammy v.Militante, Bobby D.Gerardo,"Plant Leaf Detection and Disease Recognition using Deep Learning", In: IEEE Eurasia Conference on IOT, Communication and Engineering, pp.579-582, Oct 2019

Jithy Lijo,"Analysis of Effectiveness of Augmentation in Plant Disease Prediction using Deep Learning", Proc. 5th International Conference on Computing Methodologies and Communication(ICCMC),pp;1654-1659,April 2021,Erode,India.

Sabrol, H, and K.Satish. "Tomato plant disease classification in digital images using classification tree.", In: International Conference on Communication and Signal Processing (ICCSP), April 2016,India.

Ashok, Surampalli,"Tomato Leaf Disease Detection using Deep Learning Techniques",Proc.IEEE 5th International Conference on Communication on Communication and Electronic Systems(ICCES), June 2020, Coimbatore, India.

Zhang, yang, Chenglong Song and Dongwen Zhan. "Deep Learning-based object detection improvement for tomato disease." IEEE Access, Vol 8, pp 5455-5516, March2020.

R. Ronnel and Daechul Park. "A Multiclass Deep Convolutional Neural Network Classifier for Detection of Common Rice Plant Anomalies", International Journal of Advanced Computer Science and Applications. Vol. 9, No. 1, pp. 67–70, 2018.

M. Brahimi, M. Arsenovic, S. Laraba, S Srdjan, B Kamel and M. Abdelouhab. "Deep Learning for Plants Diseases: Detection and Saliency Map Visualisation", Human and Machine Learning, Human–Computer Interaction Series. Pp. 93-117, 2018

Tetila, C. Everton, BM. Bruno, KM. Gabriel, Adair DSO, A. Marco, PA. Willian, Nícolas A DSB, Gercina GDS and P. Hemerson. "Automatic Recognition of Soybean Leaf Diseases Using UAV Images and Deep Convolution Neural Networks", IEEE Geoscience and Remote Sensing Letters. Vol. 17, No.5, pp. 903–907, 2020

Chen Junde, Jinxiu Chen, Defu Zhang, Yuandong Sun, and Y A Nanehkaran. "Using Deep Transfer Learning for Image-Based Plant Disease Identification", Computers and Electronics in Agriculture. 2020.

Kc Kamal, Zhendong Yin, Mingyang Wu and Zhilu Wu. “Depthwise Separable Convolution Architectures for Plant Disease Classification”, Computers and Electronics in Agriculture. Vol. 165, pp.1-6, 2019

Ferentinos and P. Konstantinos. “Deep Learning Models for Plant Disease Detection and Diagnosis”, Computers and Electronics in Agriculture. Vol. 145, pp. 311–318, 2018.

Hassan Amin and Ashraf Darwish. “End-to-End Deep Learning Model for Corn Leaf Disease Classification.” IEEE Access, Vol.10, pp.31103-31115, 2022.

Sunil C.K Jaidhar C.D.”Cardamom Plant Disease Detection Approach using Efficient Net V2.” IEEE Access, Vol 10, pp.789-804, 2021