Mathematical Statistician and Engineering Applications
ISSN: 2094-0343
2326-9865

# Tomato Leaf Disease Detection and Classification Using Cnn

Pushpa B R [1]

Department of Computer Science, Amrita School of Arts and Science, Mysore
Amrita Vishwa Vidyapeetham, India
preeths1@gmail.com

Aiswarya V V [2]

Department of Computer Science, Amrita School of Arts and Science, Mysore
Amrita Vishwa Vidyapeetham, India
aiswaryajayapreet99@gmail.com

**Article Info**
**Page Number: 2921-2930**
**Publication Issue:**
**Vol. 71 No. 4 (2022)**

**Article History**
**Article Received: 25** March 2022
**Revised: 30** April 2022
**Accepted: 15** June 2022
**Publication: 19** August 2022

**Abstract**
Tomatoes are a major crop of vegetables that are used in various foods. Plant diseases are currently the main challenge to food security, and researchers are attempting to develop an effective approach to identify and diagnose disease at their earliest stage, which aids in prevention of those diseases and thereby improves the agricultural sector. The dataset used in this research is a compilation of all publicly accessible datasets from Kaggle. used 14,531 datasets in 10 different classes. This study suggests an effective CNN (Convolutional Neural Network) model to categorize tomato leaf diseases and detect the name of the disease affecting tomato leaves. A 2-Dimensional Convolutional Neural Network model approach with 2-Max Assembling covers and a fully associated layer When compared to other classification models like SVM, VGG16, Inception V3 and Mobile Net CNN model, the experimental findings reveal that the model is effective enough to detect the disease with an accuracy of 96 %.
**Keywords**- Detection, Classification, Convolutional Neural Networks

## I. INTRODUCTION

The majority of tomato infections begin in the leaves before spreading to the entire plant. A favorable development environment is created via the classification and detection of tomato leaf diseases. But despite our best efforts, our garden occasionally contracts a tomato plant disease. The cost- prohibitive and subjectivity-prone ordinary proficient analysis of tomato leaf disease is problematic. With the hasty growth of computer technology, deep learning, machine learning, and computer vision are now broadly used in the diagnosis of agricultural diseases. conventional methods for computer vision Crop diseases, RGB images are separated depending on characteristics such as color, texture, or shape. However, because many diseases share characteristics, it can be challenging to identify their classification, and disease detection in complex environment is not very accurate. Deep neural networks are used as illustrations of end-to-end wisdom in the Deep learning has various uses and the proposed work is one of those uses. An input, such as an image of an unwell plant, and an output, such as a pair of yield diseases, are mapped using neural networks. In a neural network, the exact nodes take mathematical input from the incoming edges and output numerical results as the outgoing edges. Deep neural networks merely map the input layer to the production layer by stacking several layers of nodes. The difficult part of building a deep neural network is ensuring that the network's structure, tasks (nodes), and edge weights accurately transform the input to the output Deep neural networks are modified by modifying the network parameters to enhance the mapping over the course of training. Convolutional Neural Network (CNN) is a elevated in-depth learning grid that replaces labor-intensive feature extraction and image preprocessing methods with an end-to-end organization, which substantially speeds up the identification process as comparison to its research. Using CNN to forecast tomato leaf disease can increase diagnostic precision and lower labor expenses. Only a little percentage of the leaf's picture size is present in the diseased area for the analysis of tomato leaf. In order to detect disease from the complicated environment, this reading incorporates an thoughtfulness element into the CNN network model. The disease feature channel is the main focus of feature extraction, and any useless feature channel data is removed. An enhanced CNN network model for the precise detection of tomato multiple leaf disease is proposed in this research. Based on their properties, deep learning technique can be used to make a diagnosis fast and precisely. Diseases can be managed using this tactic by applying immunity techniques more quickly and making more attempts to avoid them. The ability to detect many diseases is limited, and the procedure is time- consuming. People used to be able to identify tomato diseases based on personal experience.

## II. LITERATURE REVIEW

[1]. H. Al-Hiary et al, proposed plant diseases, a novel technique based on computer image processing where The Otsu approach was chosen to segment the leaf region for samples that were taken from an existing source. Using the

Sobel operator, the diseased spot boundaries were separated into different zones.

[2]. Dheeb Al Bashish et al, proposed an image processing model which identifies diseases with an accuracy of 93% and very effective in diagnosing diseases. More studies show that accurate feature extraction and segmentation are required.

[3]. Anand H. Kulkarni et al, developed method for precisely identifying plant diseases, Segmentation and filtering are done with Gabor filter by choosing feature values that can quickly differentiate between healthy and unhealthy samples, (ANN) is trained. Investigational results show that cataloguing performance using an ANN with a article set is better, with an accuracy of 91%.

[4]. Robert G.de Luna et al, proposed a model based on CNN and F-RCNN. Performance is attained by the suggested strategy at 91.67%. Deep learning techniques and Leaf diseases are recognized and categorized using image processing methods.

[5]. J Singh H Kaur et al, proposed a discussion of the many methods for identifying plant leaf diseases which identifies disease from the input images which is taken from existing dataset available online as an effective tool for plant disease identification.

[6]. Weinan Shi et al, proposed phenotyping techniques that are more advanced also provide more accurate phenotypic data for plants. Using a multi-view approach, a two- dimensional

(2D) image of the plant can be transformed into a point-cloud model in three dimensions (3D). Using FCN and a masked R-CNN, morphological and incident segmentation on the 2D images was completed.

[7]. Gime´nez-Gallego et al, proposed a study which indicates image is segmented by individually classifying each pixel into the relevant "leaf" or "non-leaf" category. The deep-learning model was found to be the best option. [8]. Surampally Ashok et al, devised to use open-source techniques. image segmentation, and clustering to identify tomato plant leaf disease in order to build a reliable, secure, and accurate system for monitoring leaf disease in plants. [9]. Yang et al, developed the VGG16 model and Mask R- CNN model to section and categorize leaf images with multiple targets and a complex circumstantial after more than 2500 leaf shots with a complex background have been obtained, target pixels, and background pixels have been falsely labelled.

[10]. Nigam et al, proposed a few non-invasive methods that have been applied to identifying plant diseases. the process of converting acquired images into the required output format for data preprocessing. Image pre-processing, which frequently includes image improvement and color space transformation, aims to highlight the region of interest.

[11]. Chohan et al, proposed a deep learning-based model called Plant Disease Detector is developed in this paper. Using images of leaves, the model can identify a number of plant diseases. Model's testing accuracy was 98.3%. In future, the model can be combined with a drone to live identify plant problems.

[12]. P. Sharma et al, developed a method for classifying and diagnosing plant leaf diseases autonomously. More than 20000 images in 19 different modules comprise the dataset. The hyperparameters may be changed to change accuracy and using an even larger dataset with more disease categories.

[13].Parul Sharma et al, proposed a CNN model for 15 different plant species, which had an accuracy rate of more than 93%. Python and Tensor Flow are used to implement the model.

[14]. Joshi et al, proposed Vigna Mungo, a leguminous plant that is mainly grown in the Indian subcontinent, is the subject of research. A viral infection changes a number of leaf image characteristics across the entire leaf structure. [15]. B R Pushpa et al, proposed a system that uses textural traits, methods include Classification, feature extraction, and image enhancement. The characteristics of the leaf pictures are recovered using digital image processing methods, and a comparison is then made between them.

[16]. N Manohar et al, developed An image processing system with automated diagnosis and classification of different rice leaf diseases is shown in a test case on paddy leaves by using segmentation, feature extraction also it recognizes diseases.

[17 ]. V Kanchana et al, proposed a method for accurately detecting illnesses on visible plant parts that includes picture pre-processing, colour conversions, extraction of features, and segmentation. The method is then applied to a classifier..

[18]. N Shobarani et al, devised a study that suggests the needs of various mobile phone

users for automatic recognition, leaf regions should be identified and extracted from complex backgrounds. The extraction of the leaf region from complicated backgrounds is made more challenging by a number of elements, including changing background patterns, clutters, various leaf forms and sizes, as well as shifting illumination as a result of erratic weather patterns. [19]. Chowdhury et al, proposed a system that gathers 18,161 basic and segmented images of tomato leaves to be used in a deep learning architecture based on the efficientnet.

[20]. J. Kan et al, proposed a more effective U-shaped network segmentation approach for leaf image processing. In order to segment leaf images against complicated backgrounds and increase the efficiency of leaf image segmentation.

[21]. Akshay S et al, proposed a neural network approach for classifying apple defects. An artificial bee colony algorithm (FSCABC) and neural feeding network-based categorization system for fruit have been developed (FNN). These implementations show that CNN requires a new model in order to increase model performance and general accuracy.

## III. PROPOSED WORK

The proposed research provides advantage of deep learning, a branch of machine learning and artificial intelligence that also makes it easier to upload numerous datasets. Two of the networks used in deep learning to read and analyze data were Convolutional Neural Network (CNN) for image classification and Open cv for computer vision. We create a sequential model by using CNN to classify data for training and testing before dividing it into smaller groups for further study. The 2-Dimensional Max-Pooling initialization is put into the sequential model, allowing it to build a completely connected layer. The flask Python framework is then used to train this unique model, which identifies the type of disease affecting tomato leaves and yields findings with better accuracy.

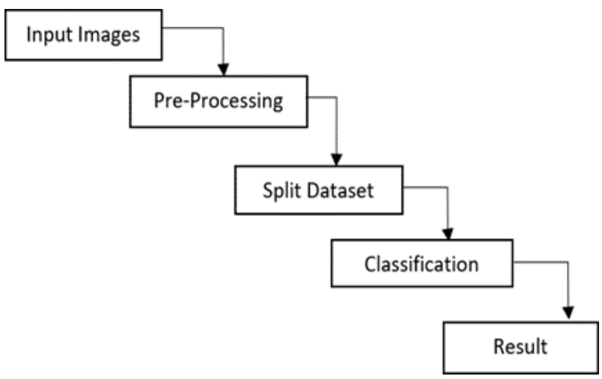


Fig 1. Proposed Method Architecture

### *A. CNN Architecture for Classification*

Convolutional, max-pooling, and fully associated layers comprise CNN architecture for classification. Convolutional and max-pooling layers are used in the feature extraction procedure. Convolution layers are designed to detect features, whereas max-pooling layers are designed to choose features. Max-pooling layers are used when an image doesn't need all the high-resolution information or when a CNN's output with smaller regions extracted after down sampling the input data. Convolutional and pooling layer output are fed into fully connected layers for classification.

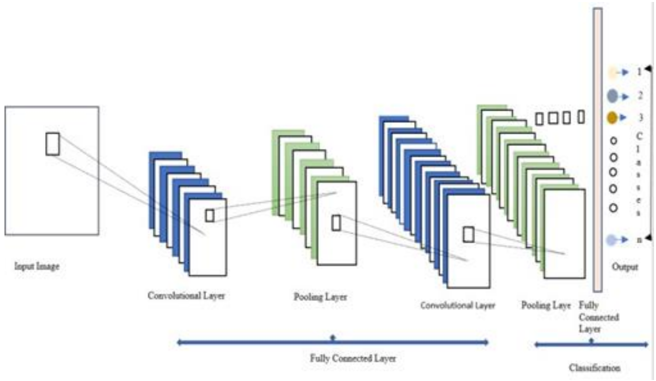


Fig. 2. The basic CNN architecture for image classification

### B. *CNN Architecture for Segmentation*

One of the essential components of computer vision is image segmentation, which deals with images. To facilitate image analysis, it entails breaking a visual input into segments. segments are made up of one or more than one pixel. Using image segmentation, it is not vital to learn of each pixel as a separate unit and instead divides them into more manageable pieces. thus, it is called sections or "tiles". Identifying discrete areas on an image that shouldn't be separated into discrete components is the first step in the segmentation process. The positions of these regions known as seeds which define the tiles.

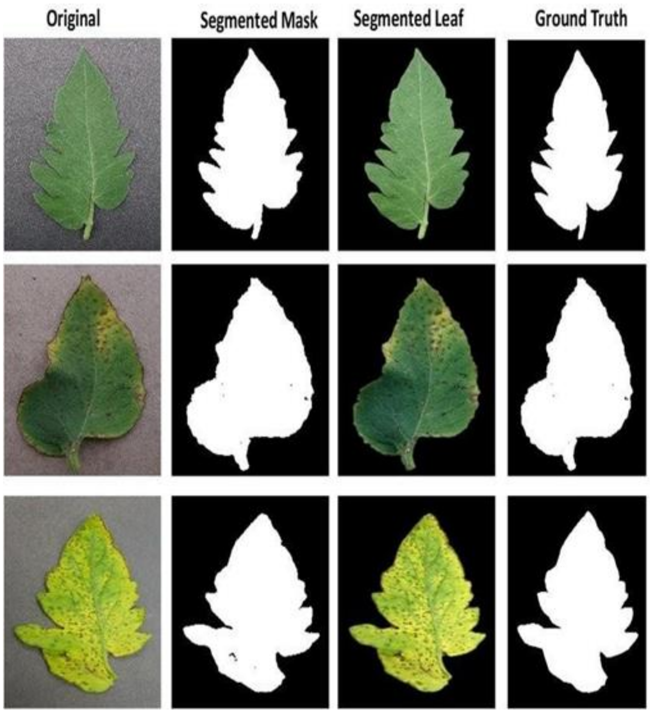


Fig 3. Image Classification, detection, and segmentation

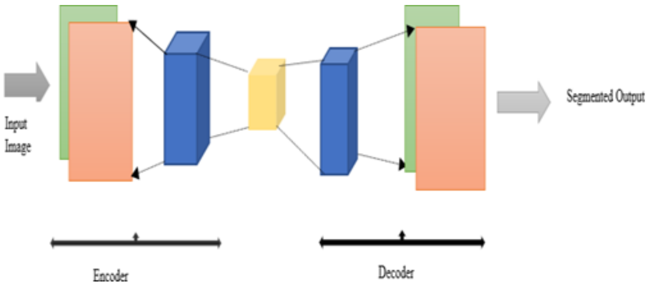


Fig. 4. The basic CNN architecture for image segmentation

## IV. METHODOLOGY

### A. *Building Dataset*

The database, used in the proposed research is the biggest crop database in the world, includes several images of plant diseases. The image data necessary for the study is manually examined after being initially obtained as an image of a tomato leaf in order to prevent issues like image duplication and classification mistakes in the dataset. The data was retrieved from Kaggle that is openly accessible to the public. It is difficult to manually acquire such a large quantity of datasets. where each leaf is meticulously chosen so that the background isn't distorted. 14,531 data were used among 10 classes.

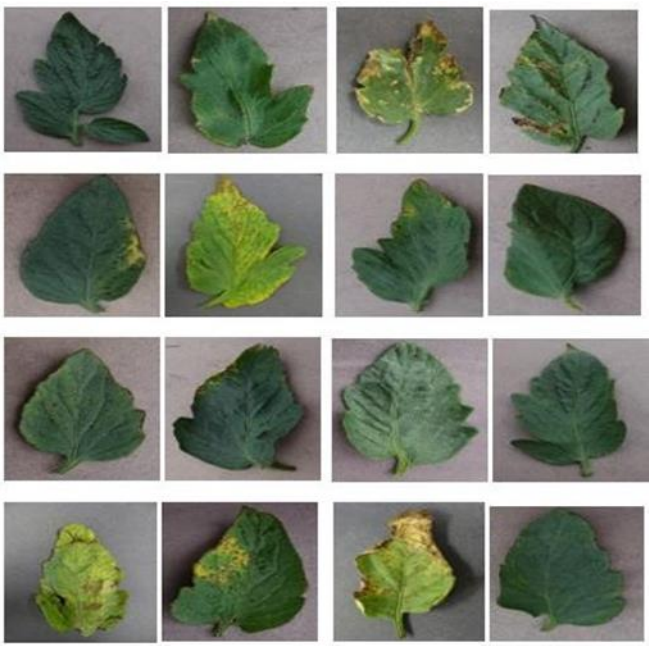

Fig 5. Dataset

### B. *Preprocessing and labelling*

Image pre-processing helps in enlightening the feature of the image before training the model. It also eliminates unwanted details of the image before training. Effective pre-processing enhances the system's accuracy and enables the particular model to choose the necessary characteristics from the picture. While Labelling helps the model to identify the classes of objects within the image.

### *C. The Evaluation Index*

The proposed network is compared to many well-known CNN models in order to gauge its performance. To assess the classification oucome utilised, precision (PPV), recall (TPR), F1 score (F1), and detection speed—the average accuracy evaluation metric used in the area of image classification—are employed (TA)

Precision: PPV=TP/(TP+FP)

Recall: TPR=TP/(TP+FN)

$$\text{F1 score: } F1=2\times((PPV\times TPR)/(PPV+TPR))$$

where TP (true positive) is the number of positive samples actually tested positive, FP (false positive) is the number ofnegative samples actually tested negative, and FN (false negative) is the amount of negative samples really tested negative.

Detection Speed:TA=T/N

Where T is the Total Detection Time for Verification set andN is the Total number of verifications set.

| | Training Accuracy | | | | | | | | | | Testing Accuracy | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 0 | 0.68 | | | | | | | | | | 0.79 | | | | | | | | | |
| 1 | | 0.79 | | | | | | | | | | 0.82 | | | | | | | | |
| 2 | | | 0.81 | | | | | | | | | | 0.83 | | | | | | | |
| 3 | | | | 0.84 | | | | | | | | | | 0.84 | | | | | | |
| 4 | | | | | 0.85 | | | | | | | | | | 0.86 | | | | | |
| 5 | | | | | | 0.86 | | | | | | | | | | 0.88 | | | | |
| 6 | | | | | | | 0.86 | | | | | | | | | | 0.83 | | | |
| 7 | | | | | | | | 0.87 | | | | | | | | | | 0.91 | | |
| 8 | | | | | | | | | 0.89 | | | | | | | | | | 0.93 | |
| 9 | | | | | | | | | | 0.88 | | | | | | | | | | 0.91 |
| | Training Loss | | | | | | | | | | Validation Loss | | | | | | | | | |
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 0 | 2.99 | | | | | | | | | | 1.74 | | | | | | | | | |
| 1 | | 2.10 | | | | | | | | | | 1.84 | | | | | | | | |
| 2 | | | 2.08 | | | | | | | | | | 1.84 | | | | | | | |
| 3 | | | | 1.90 | | | | | | | | | | 1.86 | | | | | | |
| 4 | | | | | 1.89 | | | | | | | | | | 1.63 | | | | | |
| 5 | | | | | | 1.72 | | | | | | | | | | 1.37 | | | | |
| 6 | | | | | | | 1.84 | | | | | | | | | | 2.72 | | | |
| 7 | | | | | | | | 1.80 | | | | | | | | | | 1.04 | | |
| 8 | | | | | | | | | 1.60 | | | | | | | | | | 0.84 | |
| 9 | | | | | | | | | | 1.62 | | | | | | | | | | 1.13 |

Table1. Table that contains Training Accuracy, TestingAccuracy, Training Loss and Validation Loss

### *D. CNN*

CNN is a neural network method that is frequently applied nowadays to train or analyze image data. The convolution neural network's matrix format is intended to filter images. For data preparation, CNN uses the input layer, convolutional layer, fully associated layer, pooling layer, where CNN creates a drop-out layer, and ultimately the densely integrated dataset validation set. The input test agreed in each layer can be mapped to a number of calculations. Convolutional neural network use 2- Dimensional Pooling to limit the size of feature maps after receiving an input image. Additionally, a linear activation function called RELU (Rectified Linear Activation Function), a typical activation function used in deep learningtechniques, was applied in the input layer. Where RELU can be defined as.

$$f(x) = \max(0, x)$$

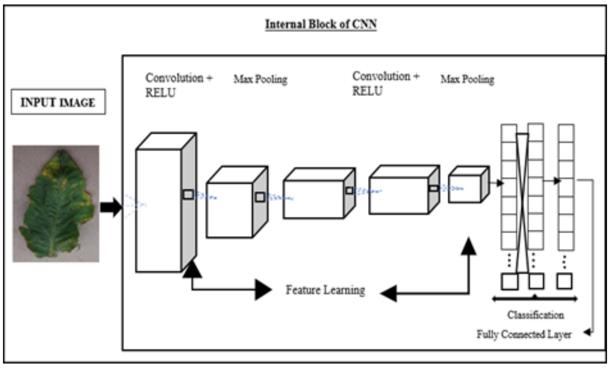


Fig. 6. Block Diagram of CNN

Sigmoid activation function, a non-linear function that generates values between 0 and 1, is present in the output layer of the fully connected system that the sequential model creates, and this function ensures that the output generated isnon-linear.

$$f(x)=\frac{1}{1-e^{-x}}$$

## V. RESULTS

The proposed approach performs better than previous classification models and classifies the diseased leaves into distinct categories. The training and testing data graphs can be used to study training and testing accuracy. The CNN model for classification generates a result of 96%, which is able to classify and detect the disease affected by the tomato leaves.

| **Class** | **Precision** | **Recall** | **F1-Score** |
|---|---|---|---|
| 0 | 0.82 | 0.68 | 0.68 |
| 1 | 0.84 | 0.78 | 0.78 |
| 2 | 0.55 | 0.77 | 0.77 |
| 3 | 0.5 | 0.79 | 0.79 |
| 4 | 0.88 | 0.68 | 0.68 |
| 5 | 0.7 | 0.78 | 0.78 |
| 6 | 0.68 | 0.77 | 0.77 |
| 7 | 0.67 | 0.79 | 0.79 |
| 8 | 0.5 | 0.6 | 0.68 |
| 9 | 0.8 | 0.4 | 0.78 |

Table.2: Evaluation Matrix

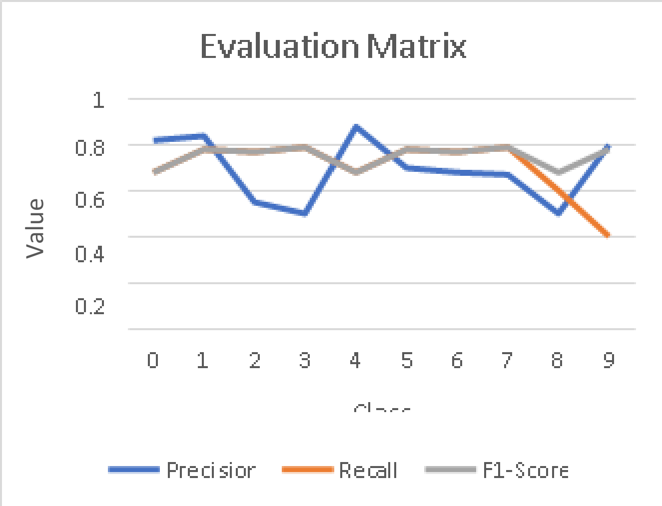


Fig 7. Evaluation Matrix

## VI. CONCLUSION AND FUTURE SCOPE

Developing a network model called CNN that takes plant's morphological properties, such as the margins of the leaves, as well as its color and texture, to identify and categorize tomato leaf diseases. This paper presents the conventional in-depth study model with several variations. The antibiotic diseases caused by bacterial and fungal pathogens that are covered in this study include blight, blisters, and browning of tomato leaves. When the results were evaluated, the proposed model outperformed the other Classification models. The targeted method to detect tomato disease is a novel concept. We intend to extend the model by including a number of abiotic diseases caused by nutrient deficiencies in crop leaves. Our long-term objectives include expanding the collection of distinctive data and gathering a sizable quantity of information on diverse plant diseases. Future technological advancements will improve detection in difficult circumstances.

## REFERENCES

[1] Al-Hiary, Heba, et al Fast and reliable plant disease diagnosis and categorization." Global Journal of computer Applications 17.1 (2011): 31-38.

[2] Sulieman Bani-Ahmad, Malik Braik, Al Bashish, and Dheeb. "K-means-based segmentation the identification and categorization of leaf diseases." The Computer Technology Journal 10.2 (2011): 267-275.

[3] Anand H Ashwin Patil., and Kulkarni. " Plant disease detection using image processing techniques." Sciences of Modern Nanoscience 2.5 (2012): 3661-3664

[4] De Luna, Robert G., et alutilising an artificial neural network to identify the leaves of philippine herbal medicinal plants Symposium on Humanoid, Biotechnology, Information Technology and Related to Network in the Environment, 2017IEEE 9th International Conference and Management (HNICEM). IEEE, 2017.

[5] J. Singh and H. Kaur, International Conferences on Innovative Systems and Control (ICISC) 2018, "A study of: various strategies of plant leaf diseasedetection,"

[6] Shi, Weinan, et al. "Deep learning and multi-view vision for plant component

segmentation." 2019;Biosystems Engineering 187: 81–95.
[7] Giménez-Gallego, Jaime, et al. "Deep learning integration of several images of tree leaves with natural backgrounds for image-based agricultural applications." 2019's Applied Sciences 10.1: 202.
[8] Ashok, Surampalli, et al. "Detection of tomato leaf disease using deep learning methods." The 5th International Conference on Communication and

Electronic Systems will take place in 2020.(ICCES). IEEE, 2020.
[9] Yong, Kunlung, Weizhen Zhong, and Fenguo Li. "Leaf segmentation and classification with a complex contextual using deep learning." Agronomy 10.11 (2020): 1721.
[10] Nigam, Sapna & Jain, Rajni. (2020). Plant disease identification using Deep Learning: A review. Indian Journal of Agricultural Sciences.
[11] Chohan, Murk, et al. "Plant disease detection using deep learning." International Journal of Recent Technology and Engineering 9.1 (2020)
[12] P. Sharma, P. Hans and S. C. Gupta, "Classification of Plant Leaf Diseases Using Machine Learning and Image Preprocessing Techniques," 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence).
[13] Sharma, Paral, Yash Paul Singh Bherwal, and Wiqas Ghai. "Enactment study of deep learning CNN models for syndrome finding in plants using image segmentation." Information Processing in Agriculture 7.4 (2020): 566-574.
[14] Joshi, R. C., Kaushik, M., Dutta, M. K., Srivastava, A., & Choudhary, N. (2020). VirLeafNet: Automatic analysis and viral disease diagnosis by deep learning in Vigna mungo plant. Biological Informatics.
[15] Pushpa, B. R., N. Megha, and K. B. Amaljith. "Comparison and Cataloguing of Medicinal Plant Leaf Based on Texture Feature." 2020 International Consultation for Emerging Technology (INCET). IEEE, 2020.
[16] Manohar, N., and Karuna J. Gowda. "Image Processing System based Identification and According to the 2020 International Conference on Electronics and Sustainable Communication Systems, "Categorisation of Leaf Disease: A Case Study on Paddy Leaf" (ICESC). IEEE, 2020.
[17] Das, Rahul, V. Pooja, and V. Kanchana. "Detection of diseases on visible part of plant—A appraisal." *2017 IEEE Technological Innovations in ICT for Agricultural science and Rural Development (TIAR)* (2017): 42-45.
[18] Shobha Rani, N., and B. R. Pushpa. "A simple and efficient technique for leaf extraction in complex upbringings of low resolve mobile photographed images." *Journal of Intelligent & FuzzySystems* Preprint (2022): 1-17.
[19] Chowdhury, Muhammad EH, et al. "Automatic and reliable leaf disease detection using deep learning techniques." AgriEngineering 3.2 (2021): 294-312.
[20] Kan, Junling, et al. "Leaf Segmentation Algorithm 2021 IEEE 4th Detailed Info Management, Transmits, Electronics and Automation Control Conference "Based on Enhanced U-shaped Networks under Complex Background" (IMCEC).Vol. 4. IEEE, 2021
[21] Akshay, S., and Deepika Shetty. "Categorization of Fruit images using Artificial Bee Colony Algorithmbased on GLCM features." IEEE, 2022