ISSN No:-2456-2165

Tomato Plant Diseases Detection Via Image Processing Using ML and DL

Rashmi Yadav
Dept. of CSE, Indira Gandhi Delhi Technical University for Women,
New Delhi, India.

Abstract:- As of now decade, many researchers have worked in the tomato plant disease detection field. It's not easy for farmers to identify tomato leaf diseases detect is difficult, for farmers challenging for them to discover other plant illnesses, such as tomato plant disease. So, the ongoing development with the help of machine-learning and deep-learning has greatly helped in identifying tomato plant disease detection by operating various methods along with tools. Precise the outcome but the accuracy of models depends on the volume as well as the quality of labeled data for training. In this article, for the detection of the disease convert the image into RBG and then identify the region-based image along with the help of segmentation by using the k-mean method. Then extract the image together with gray level co-occurrence matrix (GLCM) features used to identify a diseased infected part. Performance-based, classify images with respect to improving the efficiency of the overall model. The final output indicates that the proposed method achieved an accuracy of 95% through resnet50 for ten classes, nine disease classes, and one class that is healthy.

Keywords:- *Grey Level Co-Occurrence Matrix (GLCM), Convolution Neural Network (CNN)*

I. INTRODUCTION

The continuous development of economic growth toward agriculture depends on the environment and Climate. But temperature affects plants to occur diseases on leaves. The tomato plant is a food-rich, edible vegetable that is commonly farmed. An estimated 160 million tonnes of tomatoes are consumed annually worldwide. The tomato is regarded as a source of revenue for farm households and makes a substantial contribution to the reduction of poverty [3]. According to the received data from the food and agriculture organization of the "United Nations", the year 2014 saw the production of about 170.750 kilotons of tomatoes worldwide. In the financial year 2021, the tomato production volume in India is estimated to be 21 million metric tonnes. Diseases affect the leaves, roots, fruits, and stems of tomato plants. Phonetic variations on leaves and leaflets on tomato plants can cause irregular growth [1], staining, spots, damage, drying, mortification, etc.

Identifying the plant disease from the naked eye is a difficult task for Framer and time-consuming for farmers [2]. And farmers don't have time to contact agricultural specialists and do pest inspections because it is time-consuming for farmers. But with the help of Artificial Intelligence for Framer, it is relevant in agriculture. Farmer can easily detect tomato plant disease with help of image processing. Easley monitors large crop fields. Precision farming uses operators, sensor networks, drones, stationary stations, and mobile robots for detection. The major drawback of mechanisms so that they are not able to explore the area as a specialist. To be efficient to perform precision farming, these tools must have the potential to process and estimate the gathered data like a professional on the farm or in the greenhouse. About 70 percent population depends on the agriculture sector.

Plant disease is a kind of infection that is naturally caused by the normal growth of plants. It is common for all crops. It is caused by fungi, bacteria, some of the plants damaged by temperature, etc. Such as diseases distinguish abnormal functioning from environmental factors. The bacteria fill with water-conducting tissue of the plant, multiplying rapidly inside it[14]. These results will affect the vascular system of the plant, leaves may stay green. The disease may affect yield growth and also the integrity of the plant which will damage crops and also kill whole plants. There are several biotic diseases caused by fungal, bacterial pathogens, browns of tomato leaves, specifically with blight and blast [3].

Plant disease detection by visual is a more laborious task. Whereas plant disease detection automatically is more time-consuming and less effort and gets accurate results. But with machine vision, it is easy to detect damage to tomato leaves

As compare, manually detection of disease is not possible and gives low efficiency and high cost [9]. Among the advanced technology, its image based disease identification has seen huge adoption in computer vision applications. So, now can apply the latest technology for the detection of plant disease [13] which reduces cost, is easily detected, is easily recognized, and improve efficiency.

ISSN No:-2456-2165

There are several techniques used to detect Tomato plant disease detection by image-processing whereas cnn is generally used for better performance and gives better accuracy. Deep-learning-based plant illness detection techniques can be perceived as the application of classic networks relevant in the realm of farming. The significant improvement in deep learning models with their latest architectures like VGGNet, LeNet, ResNet, etc., are trained for the detection of tomato leaf disease to provide the highest classification performance.

In this research article, the goal have proposed for identifying tomato leaf disease detection. This work will be helpful for farmers to solve their problem of tomato disease leaf identity without running plant scientists. It will restore to health of the tomato leaf and increase the quality and quantity of crop food so, it will increase profits for farmers. For this experiment purpose, download the plantvillage dataset. After this, 1.structured the dataset images by image color conversion, and image resizing. 2. segmentation process which is broke the images into various subgroups with a k-mean algorithm shows the no. of the infected area selected for segmentation. 3. Feature extract will extract the image by using GLCM. 4. Classification is done by Resnet the main challenge for classification is to improve accuracy with the dataset and to test results on a dataset that is different from what it is being trained on.

II. SURVEY ON TOMATO PLANT DISEASE DETECTION

Tomato Plant disease detection was invented to overcome the problems. The main reason for using tomato plant disease detection is to quickly identify tomato disease detection. There are different-different researchers have focused on their work to enhance their model performance for automatic tomato plant disease detection. This section represents a discussion and review of recent work on identifying plant diseases based on different algorithms. As in given fig1 can see the algorithm performs on tomato disease detection.

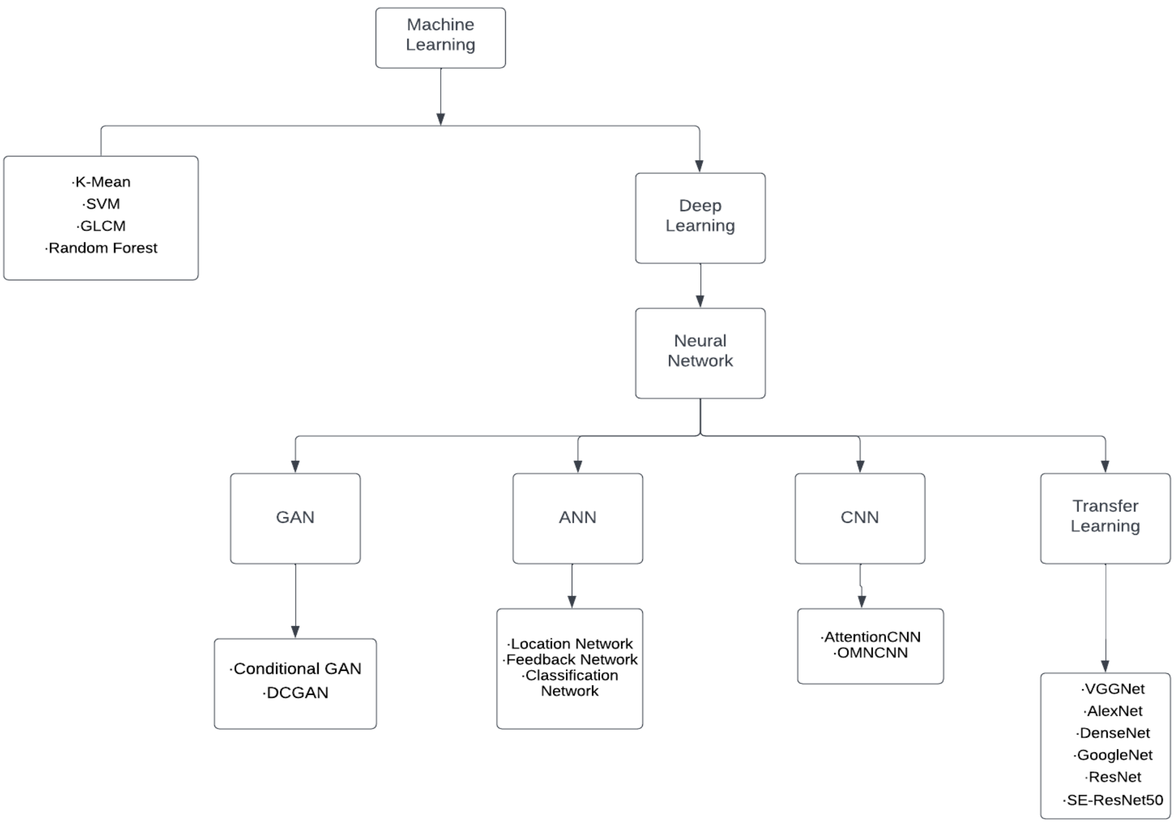

Fig 1:- Algorithm used in Tomato Plant disease detection

A. Convolution Neural Network and Transfer Learning

In this [3] they are trying to innovate a new way to identify and classify tomato leaf disease by using CNN. In this plantvillage dataset is used whereas they consider 3000 images of healthy and unhealthy leaves. CNN model is divided into parts firstly preprocessing the data, and segmenting the images with the original image. Then process with CNN whereas CNN plays a role like a feature extract and also compare with VGG and ResNet but CNN gives the best accuracy 98.48%

In this [4] paper they are improved accuracy. In this work, they proposed a new architecture SE-Resnet50. SE-ResNet50 was suitable for the early diagnosis of tomato leaf disease detection. Whereas some of the models have done previously but researchers compared a new model from previous work. After the coming result, SE-Resnet50 gives better accuracy as compared to other models like GoogleNet, Xception, Vgg19, and Resnet-101. The archive 96.81 has % accuracy in the Tomato plant disease dataset.

In [8] proposed automatically tomato disease detection using residual CNN. The proposed attention-based Residual CNN is subjected to a 5-fold cross-validation process. In this 95999 images were used for training the model and 24001 images were used for validation purposes. The researcher improves the performance of the residual CNN. The achieved accuracy is 98%.

In [9] proposed work was to improve the accuracy of previous work by adding a deep convolutional generative adversarial network (DCGAN) to identify tomato plant infected images. By adding some adjustments in hyper-parameter, modified in CNN architecture, and improving the trained model by distributing the dataset into 5 classes of tomato plant diseases images. For better quality and better approach -Distributed Stochastic Neighbor Embedding (t-SNE) for visual tuning and more convincing obtain improved accuracy is 94.33%.

In [10] proposed disease detection model that learns to detect instances of disease based on generic features. The algorithm utilizes the ResNet deep learning architecture. investigated a novel computer vision system for recognizing a number of diseases in an automatic manner, and detecting invisible diseases. The results depicted the effectiveness of the investigated system when various models which is integrated into an automated surveying system to diminish the costs and measurement bias and maximize the precision and greenhouse coverage.

In [13] the main approach for this work is to identify miner-pest (Tuta absoluta) invasion. The CNN Architecture was used on tomato leaf which is captured by the tomato crop. In this paper, the researcher used their own dataset which contains healthy and unhealthy images. Classified on the different model but for experiment purpose 2125 images were used and also classified by transfer-learning. VGG16 was a gives high-performing model, which obtain 91.9% accuracy.

Detecting disease in tomato crop [14]. The images present in class were not balanced, this issue was resolved using data augmentation methods. The experimental results validated that the intended algorithm offered an average accuracy of up to 91.2% while diagnosing the diseases in plants.

In [17] the proposed method is using gan in the agriculture field. Whereas DATFFAN performs state-of-the-art or visualize quality and classification accuracy. It is the first Gan in the agriculture field for crop identification and disease detection. DATAFFGAN to get better quality output and better performance. This work is only for comparison purposes.

In [18] they generate a synthetic image of a tomato plant by using conditional generative adversarial networks for augmentation purposes. Pretrained the densenet121 model for finetuning through the original image and also image synthesis. Plantvillage dataset is used. The suggested technique offered the accuracy of 99.51% to classify the image into 5 classes, 98.65% for 7 classes, and 97.11% for 10 classes. The suggested technique was more adaptable as compared to other techniques.

In this[19] paper they propose a restructuring of the dense residual network to adjust structure and parameter. Improve performance of crop plants and also less impact of crop plant disease. Images were used for super-resolution. This work is a combination of residual and densenet and restructured RDN To RRDN. Whereas reducing the output layer, upscale layer, and also adding a dense layer, and adjusting the input features and hyperparameters. The achieved accuracy is 95%..

In this[20] paper the proposed work is to identify the disease and classify the dataset of images with the help of an OMNCNN (optimal mobile network-based on CNN. The experimental results indicated to formulated mechanism offered precision up to 98.5%, recall of 98.9%, accuracy up to 98.7%, and an F-score around 0.98.

The number of metrics was minimized by means of Depth wise Separable Convolution [24]. Diverse datasets were applied to train and test the established framework. This framework had an accuracy of 99.39% on the Plant Village dataset, 99.66% on rice disease, and 76.59% on the cassava dataset.

B. ANN

In this[5] paper, the proposed work is trying to identify tomato diseases identify whereas they work on various features and also various techniques for recognizing diseases like backpropagation, BPNN, and neural network. Compared results based on neural networks by using different features.

The [10] approach is to develop a framework that performed more effectively on the tomato dataset, ImageNet, and attained 99.7% accuracy in comparison with the traditional technique. The proposed work has LFC-Net which consists of three networks:- the location network, feedback network, and classification network which is trained the dataset from end to end. And also compared other models for better performance

In this [22] it has two approaches first identify infected tomato plant images by using a smartphone and furthermore overcome non-standardized environmental illumination. Extracting images, and reshaping images also used ANN. They detect rice disease with the help of the plant village. Which capture from a crop field. The accuracy of the model is 100%.

ISSN No:-2456-2165

C. GLCM

In [2] the main purpose is to detect the infected images and the proposed work depends on image segmentation and classification method. For image smoothing used kurtosis and skewness filter. SVM model is used for classification training purposes. Also include contrast, homogeneity, energy, and correlation algorithm for feature extraction. The SVM classification obtains accuracy is 93.75%.

The main[12] approach is to identify and classify tomato diseases automatically for which superpixel-based optimized segmentation was developed in natural images to remove the unwanted background from the image. Whereas for the segmentation part using k-mean clustering for finding infected textures o images. To extend the feature extraction part using glcm and phog were used in this paper for finding better results. The most important is they used two datasets one is plantvillage and another is they collect the dataset and then perform on both datasets. The leaves were split from the background using a new technique planned on the basis of color variations in a successful manner. The proposed model accuracy achieved 98%.

Detect [25] the disease of the plant on the basis of 4 operations where the image was pre-processed, segment extracted the infected images and classified models. The classified symptoms analyze from the glcm algorithm and classification is done by model comparison this is also key to this paper. Results proved that the recommended method yielded higher performance with regard to accuracy while detecting the diseases of plants at the primary phase.

There are some limitations finding that, the Few of leaf image which was captured from the natural environment noted they have shadow and uneven illumination which get difficult in the segmentation part. The main drawback of agriculture is that different diseases that affect plants. The existing limitations are taking images from the environment that affect on approach also and not getting proper segment images. And this technique will improve the productivity of crops

III. PROPOSED METHODOLOGY

The proposed methodology has the following steps (a) import dataset for pre-processing. (b) segment the image (c) feature- extraction (d) classifying (e) result.

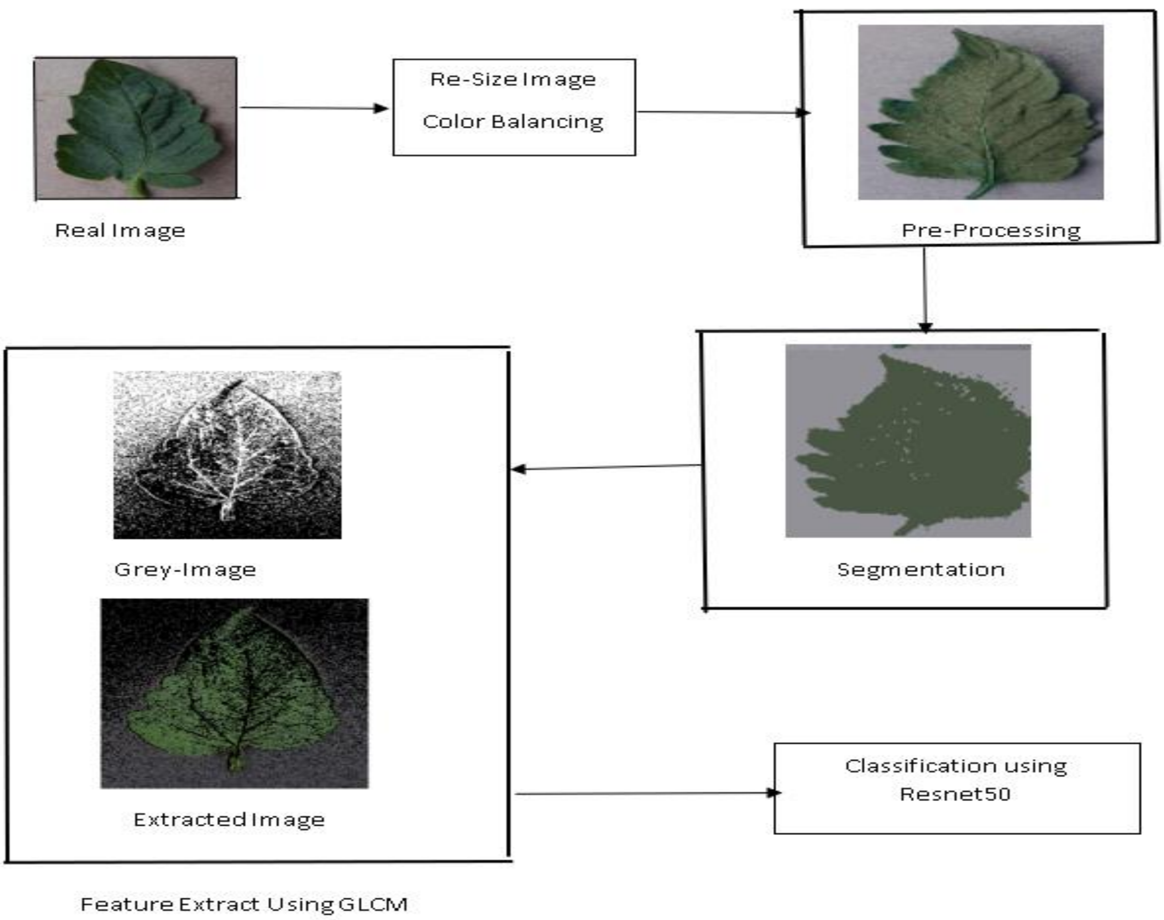

Fig 2:- Proposed Model

A. Pre-processing

For pre-processing, In the beginning, resize the image into 32x32 and then convert BGR TO RBG. Whereas the range of the RBG is [0, 255]. Reshape the original dataset images.

B. Segmentation

Segmentation of images is the most important part of recognizing images. In this work, recognize the infected diseases in the image from the dataset and remove the unwanted background of the image only segment on infected images. But in recognition, it deals only with infected diseases. Basically, image segmentation is used to identify objects and boundaries. Whereas it uses K-means segmentation clustering an unsupervised method that is optimized for segmented images and partitions the images into a number of clusters where each image represents different aspects of inputs images. This partial enhancement to the image improves its quality of the image.

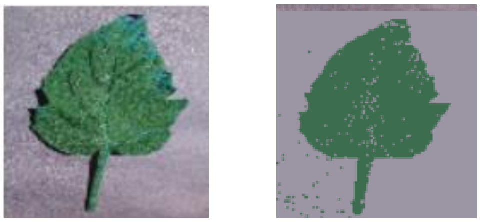

Fig 3:- (a) Original Image (b)Segmented Image

C. Feature Extraction

For feature extraction, images are extracted that are useful to classify the images. It also creates new features from the original ones. It is used for medical images, optical characters, and many more fields. In this section, So, feature extraction re-sizes the image, converting RBG to LAB space color. Basically, GLCM is used to characterize the texture of the image by calculating by pair of pixels and to measure gray level co-occurrence model intensity levels. GLCM has four properties contrast, homogeneity, energy, and correlation as shown in fig 4.

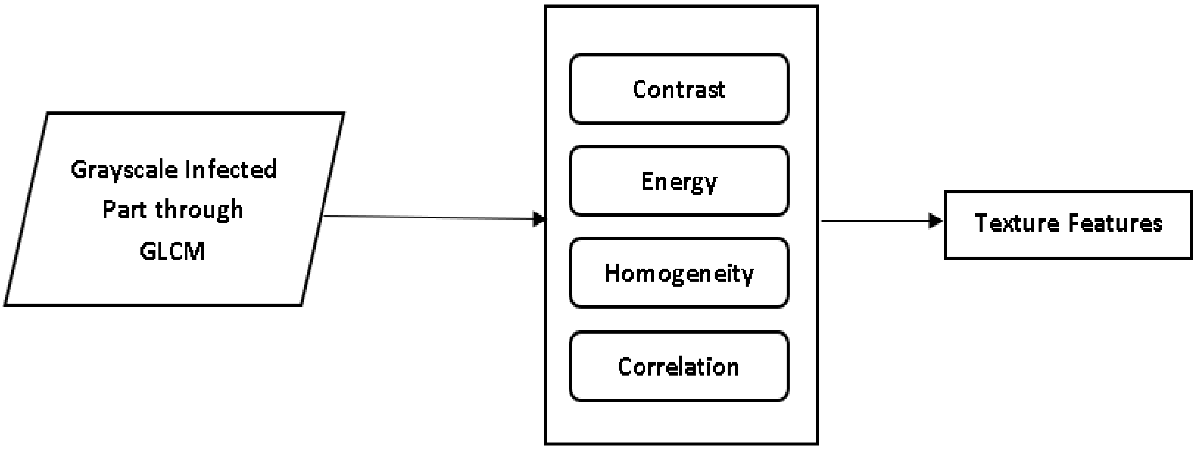

Fig 4:- GLCM architecture

D. Classification – Transfer Learning

Classification is a final step this technique is used to classify the image with transfer-learning. So transfer learning is a very popular algorithm in deep learning models for training a large amount of dataset into a small amount of data. The network used for the information purpose was learned from the previous activity to increase prediction for the new task in the transfer-learning. This proposed work on resnet50 is based on classification work. After carefully analyzing, Resnet50 gives better results as compared to another two transfer learning models. Firstly pre-trained architectures Resnet-50, Vgg16, and Densnet201.

IV. IMPLEMENTATION AND RESULTS

A. Dataset

For tomato plant diseases detection dataset can be created on its own by taking leaf images from yield, network collection, or public dataset. Whereas Plant Village is a public dataset and also a well-known dataset. But has its own capture image dataset of leaves and is also more practical. Researchers also use this dataset. There are so many open datasets that anyone can collect dataset. But for research 1083 images are taken and divided into 10 classes all the images are divided into 10 classes as shown in fig 5.

ISSN No:-2456-2165

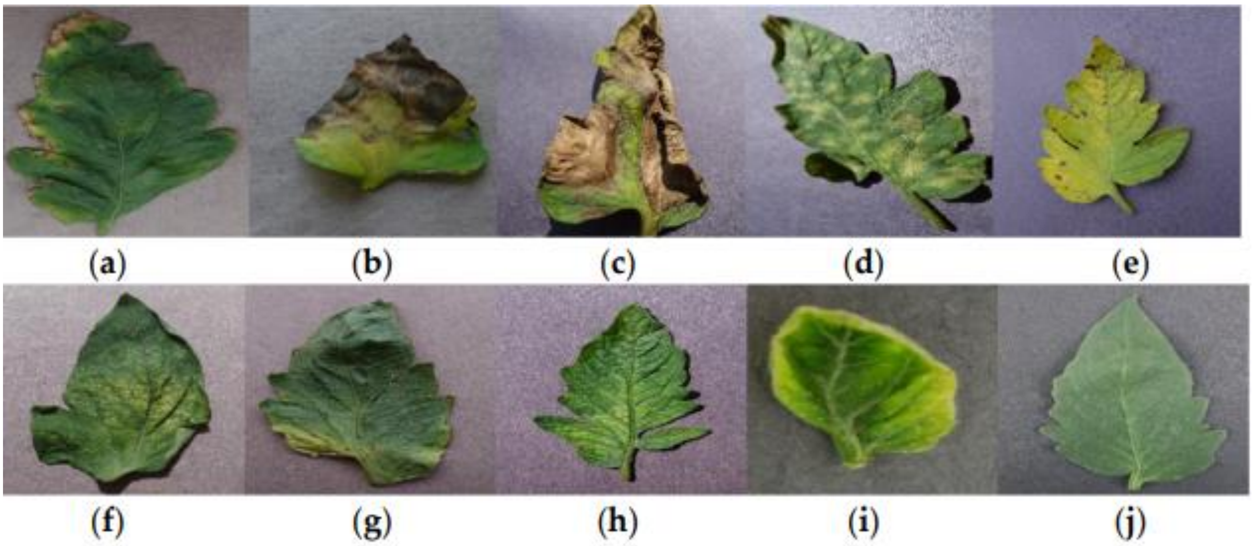

Fig 5:- Sample of tomato leafs (a) Bacterial_Spot(Xanthomonas vesicatoria), (b) Early_Blight(fungus Alternaria solani), (c) Late_Blight(Phytophthora infestans), (d) Leaf_Mold(Cladosporium fulvum), (e) Septoria_Leaf_Spot(fungus Septoria lycopersici), (f) Spider_Mites(floridana), (g) Target_Spot(fungus Corynespora), (h) Tomato_Mosaic_Virus(Tobamovirus), (i) Tomato_Yellow_Leaf_Curl_Virus(genus Begomovirus), (j) Healthy_Leaf.[3]

B. Tools/Libraries

- Anaconda-Navigator

  Anaconda-navigator is a desktop graphical user interface (GUI) embedded in the Anaconda distribution that allows us to launch applications and easily manage conda packages, locations, and channels without using command-line commands.
- Keras

  Keras is an in-depth Python learning API, running on the TensorFlow machine learning platform. Keras was built with a focus on compliance and rapid testing.
- Programming Language used: Python

C. Parameter

Each of the experiments runs a total of 15 epochs. Whereas it completely trained the networks. The choice of 15 epochs is based on empirical observation. Hyperparameter of deep neural networks used is a number of layer – 6, number. of Max Pooling layers – 8, Loss - categorical cross-entropy, Epoch – 15, Batch Size – 20, Activation Function – Relu, Weight – ImageNet, Leaning Rate – 0.0001. These layers are the same for the three models which were compared with each other performance.

D. Results

From the experimental work, choosing a better efficiency accuracy model takes the better performance of the model. Classifiers including VGG16, Resnet50, and Densnet201, all are examined for best performance. Ran the proposed model for15 epochs and validation and training accuracy. The proposed work also compares the models by running transfer-learning models. Build the model (VGG16, Densnet201, Resnet50) with a connected layer with one-fully of 256 neurons with a flattened layer and three Dense layers. Relu were used for activation function and classify with 10 neurons because of 10 classes. The matrix is used to evaluate the matrix performance that is accuracy, precision, recall, and F1 score for comparison purposes.

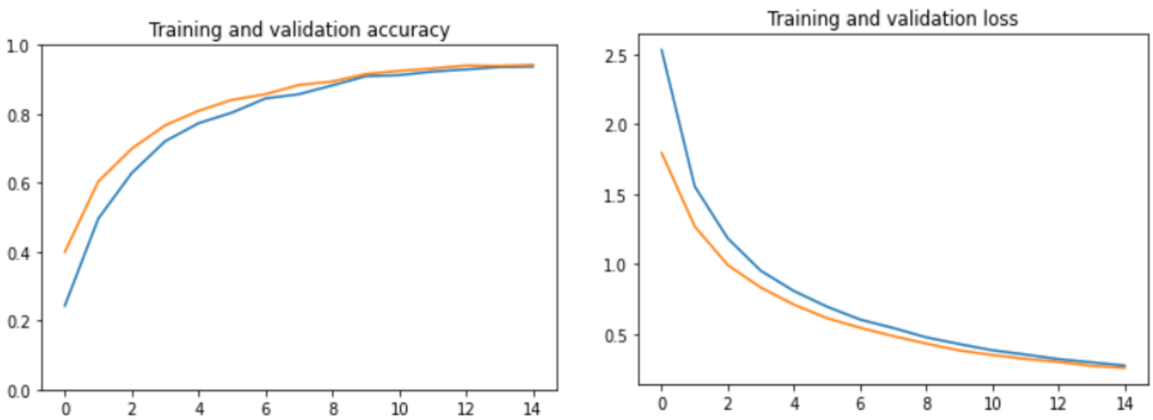

Fig 6:- Vgg16 Validation-accuracy and validation-loss

ISSN No:-2456-2165

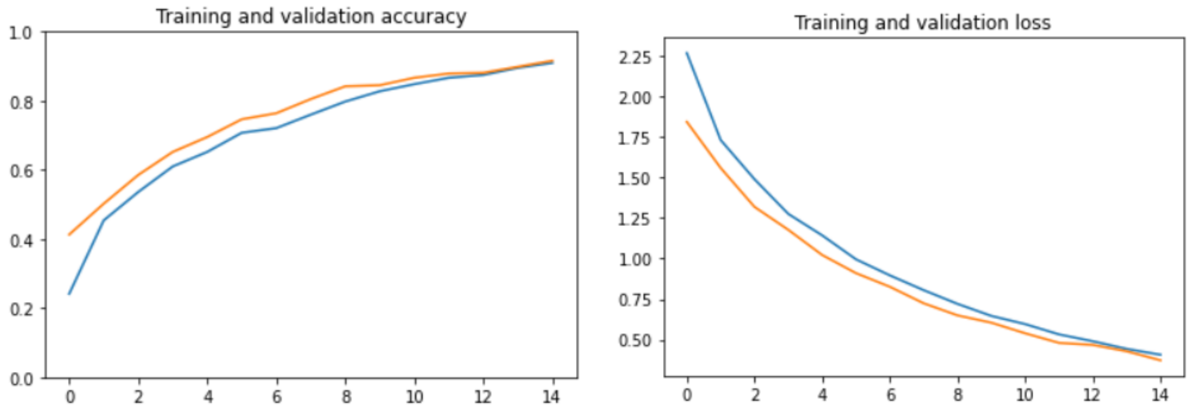

Fig 7:- Densnet201 Validation-accuracy and validation-loss

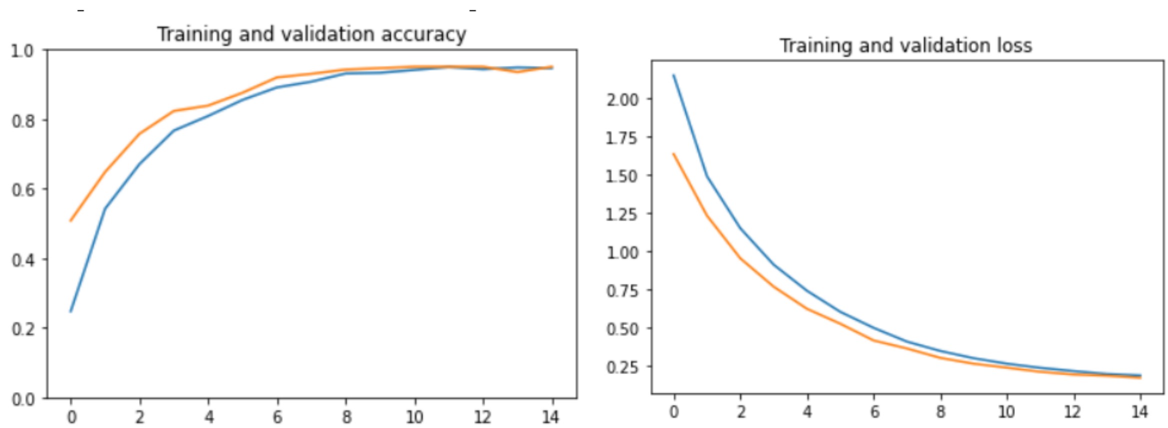

Fig 8:- Resnet50 Validation-accuracy and validation-loss

Classifier	**Accuracy**	**Precision**	**Recall**	**F1-Score**
Densenet201	91.58%	93.29%	91.48%	91.69%
Vgg16	94.25%	95.81%	94.19%	94.45%
Resnet50	**95.07%**	**96.68%**	**95.02%**	**95.29%**

Table 1:- Comparison of Classification

Now can conclude from the above table that the best classification accuracy received by Resnet50 had categorical accuracy of 95.07% when the model was trained on 15 epochs, followed by Vgg16 with a categorical accuracy of 94.25% and Densnet201 model accuracy of 95.07. The comparison matrix is shown in table 1 and validation accuracy and loss are shown in fig 6, fig 7, and fig 8.

V. CONCLUSION AND FUTURE WORK

This paper demonstrates automatic disease detection using GLCM and the Transfer Learning model for classifying images. The proposed performance is tested on Plant Village Dataset which is downloaded from the internet. In the experiment, 1083 images are taken from the plant village dataset. In the dataset, there are 9 disease classes, and one class which is having healthy images. The pre-processing step consists of color balancing and image resizing for segmentation.

Segmentation through k-mean can remove easily identify the defective part and remove unwanted background. After removing unwanted background GLCM can easily extract the image. GLCM texture feature proved to be efficient in providing the uniqueness of specific disease patterns. The implemented classification is classified by three transfer learning models but the resnet50 gives better performance and improves the proposed work.

Furthermore, it has been noticed that there is still more chance to improve the classification method and feature extraction method. And can-do modifications to dataset method like by using a large number of datasets and making own dataset for research work. Various segmentation techniques can be used like thresholding and can use fracture extraction like the PCA algorithm and compare the model with the current proposed model.

REFERENCES

[1]. Qimei Wang, Feng Qi, Minghe Sun, Jianhua Qu, and Jie Xue, “Identification of Tomato Disease Types and Detection of Infected Areas Based on Deep Convolutional Neural Networks and Object Detection Techniques”, 2019, Hindawi, Computational Intelligence and Neuroscience, Article ID 9142753, 15 pages

[2]. Sagar Vetal, R.S. Khule, “Tomato Plant Disease Detection using Image Processing”, 2017, International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), Vol. 6, Issue 6

[3]. Naresh K. Trivedi, Vinay Gautam, Abhineet Anand, Hani Moaiteq Aljahdali, Santos Gracia Villar, Divya Anand, Nitin Goyal, and Seifedine Kadry, “Early Detection and Classification of Tomato Leaf Disease Using High-Performance Deep Neural Network”, 2021, Sensors, Vol. 21

[4]. Shengyi Zhao, Yun Peng, Jizhan Liu and Shuo Wu, “Tomato Leaf Disease Diagnosis Based on Improved Convolution Neural Network by Attention Module”, 2021, Agriculture, vol. 11, no. 651

[5]. Roopali Gupta, ToranVerma, “Tomato Leaf Disease Detection using Back Propagation Neural Network”, 2020, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Vol. 9, no. 8

[6]. C. Zhou, Z. Zhang, S. Zhou, J. Xing, Q. Wu and J. Song, "Grape Leaf Spot Identification Under Limited Samples by Fine Grained-GAN," in IEEE Access, vol. 9, pp. 100480-100489, 2021

[7]. Siddharth Singh Chouhan, Ajay Kaul, Uday Pratap Singh, Sanjeev 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”, 2018, IEEE Access

[8]. Karthik R., Hariharan M., Menaka R., “Attention embedded residual CNN for disease detection in tomato leaves”, 2019, Applied Soft Computing

[9]. Qiufeng Wu, Yiping Chen, Jun Meng, “DCGAN-Based Data Augmentation for Tomato Leaf Disease Identification”, 2020, IEEE Access

[10]. Guofeng Yang, Guipeng Chen, Yong He, Zhiyan Yan, Yang Guo, Jian Ding, “Self-Supervised Collaborative Multi-Network for Fine-Grained Visual Categorization of Tomato Diseases”, 2020, IEEE Access

[11]. Patrick Wspanialy, Medhat Moussa, “A detection and severity estimation system for generic diseases of tomato greenhouse plants”, 2020, Computers and Electronics in Agriculture

[12]. Saiqa Khan, Meera Narvekar, “Novel fusion of color balancing and superpixel based approach for detection of tomato plant diseases in natural complex environment”, 2020, Journal of King Saud University - Computer and Information Sciences

[13]. Lilian Mkonyi, Denis Rubanga, Dina Machuve, “Early identification of Tuta absoluta in tomato plants using deep learning”, 2020, Scientific African

[14]. Mohit Agarwal, Abhishek Singh, Suneet Gupta, “ToLeD: Tomato Leaf Disease Detection using Convolution Neural Network”, 2020, Procedia Computer Science

[15]. Xinda Liu, Weiqing Min, Shuhuan Mei, Lili Wang, Shuqiang Jiang, “Plant Disease Recognition: A Large-Scale Benchmark Dataset and a Visual Region and Loss Reweighting Approach”, 2021, IEEE Transactions on Image Processing

[16]. Wei-Jian Hu, Jie Fan, Yong-Xing Du, Bao-Shan Li, Naixue Xiong, Ernst Bekkering, “MDFC–ResNet: An Agricultural IoT System to Accurately Recognize Crop Diseases”, 2020, IEEE Access

[17]. Qiang Dai, Xi Cheng, Yan Qiao, Youhua Zhang, “Crop Leaf Disease Image Super-Resolution and Identification with Dual Attention and Topology Fusion Generative Adversarial Network”, 2020, IEEE Access

[18]. Amreen Abbas, Sweta Jain, Swetha Vankudothu, “Tomato plant disease detection using transfer learning with C-GAN synthetic images”, 2021, Computers and Electronics in Agriculture

[19]. Changjian Zhou, Sihan Zhou, Jinge Xing, Jia Song, “Tomato Leaf Disease Identification by Restructured Deep Residual Dense Network”, 2021, IEEE Access

[20]. S. Ashwinkumar, S. Rajagopal, B. Jegajothi, “Automated plant leaf disease detection and classification using optimal MobileNet based convolutional neural networks”, 2021, Materials Today: Proceedings

[21]. Penghui Gui, Wenjie Dang, Qijun Zhao, “Towards automatic field plant disease recognition”, 2021, Computers and Electronics in Agriculture

[22]. Rahim Azadnia, Kamran Kheiralipour, “Recognition of leaves of different medicinal plant species using a robust image processing algorithm and artificial neural networks classifier”, 2021, Journal of Applied Research on Medicinal and Aromatic Plants

[23]. Jiangtao Qi, Xiangnan Liu, Yang Li, “An improved YOLOv5 model based on visual attention mechanism: Application to recognition of tomato virus disease”, 2022, Computers and Electronics in Agriculture

[24]. Sk Mahmudul Hassan, Arnab Kumar Maji, “Plant Disease Identification Using a Novel Convolutional Neural Network”, 2022, IEEE Access

[25]. Amal Mathew, Anson Antony, Apeksha Kulkarni, “Plant disease detection using GLCM feature extractor and voting classification approach”, 2022, Materials Today: Proceedings