**ENHANCING CROP PRODUCTIVITY**

**THROUGH EARLY PLANT DISEASE IDENTIFICATION**

*Dr. Meeta Chaudhary*

*Deepanshu Goel*

*Department of, CSIT*

*Kiet group of institutions*

*Delhi, NCR*

*deepanshu.2024csit1010@kiet.edu*

*Associate Professor, Department of CSIT*

*KIET Group of Institutions*

*Delhi NCR*

*meeta.chaudhary@kiet.edu*

*Aditya Singh*

*Department of, CSIT*

*Kiet group of institutions*

*Delhi, NCR*

*aditya.2024csit1076@kiet.edu*

*Garvit Chauhan*

*Department of, CSIT*

*Kiet group of institutions*

*Delhi, NCR*

*garvit.2024csit1010@kiet.edu*

*Anshul*

*Department of, CSIT*

*Kiet group of institutions*

*Delhi, NCR*

*anshul.2024csit1151@kiet.edu*

**Abstract**:India is an agricultural country, with agriculture contributing the most to GDP (about 64%), followed by other industries [6]. Agriculture is one of the pillars of our economy. Therefore, agriculture is the main source of income. However, due to many factors such as pests and diseases, and climate change, agricultural production has decreased and productivity has also decreased. Identification of plant diseases is important in preventing production and yield losses. Control of plant diseases is not clear, and diagnosis of the disease is very important. It requires a lot of effort combined with knowledge of plant diseases and a lot of study time. Therefore, imaging techniques are used to identify plant diseases by capturing full images and comparing them with recorded data. This document has many sheets in image format. This project is designed to create a standalone application that will provide necessary information about the disease. The aim of the project is to help monoculture farmers and provide good support. In this paper, we propose a method that will use different imaging techniques to observe and diagnose plant diseases. Application results show that the developed system can identify and classify plant diseases and achieve good results. In this article, we focus on three types of plant diseases: bacterial blight, Cercospora leaf spot, and Alternaria.

***Keywords***: Image Processing; K-Means; GLCM; SVM; Plant Disease Detection; Standalone Application,RCNN.

*Keras, TensorFlow*

1. **INTRODUCTION**

Local agronomists may find it difficult to distinguish between diseases that may be present in their crops. It is not normal for them to go to an agricultural enterprise and find out what kind of infection it is. Our main goal is to recognize diseases occurring in plants by observing their morphology using image processing and machine learning. Agriculture is the backbone of India's economy. Most of our country's economy relies primarily on agriculture. Farmers have a big range of options to select the crops and choose the convenient chemicals and pesticides to use. Thus, crop damage could result in generating unexpected losses that will impact the productivity of the farming industries which will directly affect the economy. Therefore, taking care of the plants is necessary to maintain an excellent quality of agriculture and guarantee the efficient productivity along with the high profit.

Plants are sensitive to diseases especially the plant leaves as symptoms of the disease appear first on the leaves. Due to the bad impacts of plant diseases on the both the economy and the environment, the farmers should consider monitoring the crops in such a way that they may mitigate losses. Pests and diseases destroy crops or plant parts, reducing food production and causing food insecurity. Additionally, in some less developed countries there is less knowledge about pest and disease management. Toxic pathogens, poor disease control, rapid climate change, etc. are some of the major factors causing decline in food production. Recently, automatic detection of plant diseases has been actively researched. Identification of plant diseases requires accurate and precise information on disease quantification [8]. In [9,10], the authors studied diseases of potatoes and tomatoes and grapes and showed how these crops are affected by viruses. In [11], the authors analyzed several articles on rice disease classification and also considered various criteria such as the dataset used, disease classification, preprocessing and segmentation methods, and classifier used. Prajapati et al. [12] conducted a study to classify diseases in cotton plants using machine learning techniques. Iqbal et al. [13] investigated the classification of citrus plant diseases using image processing. Kaur et al. [14] conducted a survey on methods for identifying and classifying plant diseases from leaf images. These studies discussed in [11,12,13,14] are based on handicraft elements. Classifying diseases using manually generated features requires preprocessing, segmentation, and feature extraction from images, which is a time-consuming and labor-intensive process.

In recent times, server based and mobile based approaches for disease identification have been employed for disease identification and gave what preventive measure we take to save the crops. Several factors of these technologies being using high resolution camera, high performance processing and extensive built in accessories are the added advantages resulting in automatic disease recognition. The automated detection of plant diseases has been studied largely in recent times. The identification of diseases in plants requires accurate and precise information regarding the quantitative measurement of diseases [8]

The histogram of oriented gradients (HOG) is an element descriptor utilized as a part of PC vision and image processing for the sake of object detection. Here we are making utilization of three component descriptors:

1. Hu moments

2. Caral texture

3. Color histogram

Hu moment is mainly used to determine the shape of leaves. We use the Haralick texture to obtain the leaf texture, and use the color histogram to represent the color distribution of the image.

**LITERATURE REVIEW**

[1] S.S. Sanakki, V.S. Rajpurohit proposed “backpropagation neural network based pomegranate disease classification”, which mainly features color and texture based on defect area segmentation method. Here they used a neural network classifier for classification. The biggest advantage is that the image saturation layer is extracted by converting to L\*a\*b, and the classification accuracy is 97.30%. The biggest drawback is that it is only used for limited yields.

[2] PR Roth and R. V. Kshirsagar presented “Identification of cotton leaf diseases using pattern recognition techniques” using snake segmentation. Here the Hu moment is used as a distinguishing feature. The BPNN classifier, an active circuit model used to constrain the feasibility of hotspots, solves the multi-class problem. The average classification is 85.52%.

[3] Aakanksha Rastogi, Ritika Arora and Shanu Sharma,” Leaf Disease Detection and Grading using Computer Vision Technology &amp;Fuzzy Logic”. K-means clustering used to segment the defected area; GLCM is used for the extraction of texture features, Fuzzy logic is used for disease grading. They used artificial neural network (ANN) as a classifier which mainly helps to check the severity of the diseased leaf.

[4] Godliver Owomugisha, John A. Quinn, Ernest Mwebaze and James Lwasa, proposed” Automated Vision-Based Diagnosis of Banana Bacterial Wilt Disease and Black Sigatoka Disease “Color histograms are extracted and transformed from RGB to HSV, RGB to L\*a\*b.Peak components are used to create max size tree, five shape attributes are used and area under the curve analysis is used for classification. They used nearest neighbor, decision tree, random forest, superrandom tree, Naive Bayes and SV classifiers. A random tree of seven classifiers generates very high scores, provides real-time information, and provides application flexibility. [5] Yuan Tian, ​​​​​Chunjiang Zhao, Shenglian Lu, and Xinyu Guo, “SVM-based multi-classification system for wheat leaf disease recognition.” The color properties of HIS are expressed in RGB and GLCM is used to take the seven invariant moments as shapes. parameter. They used SVM classifier with MCS used for unsupervised detection of wheat plant diseases.

**PROPOSED METHODOLOGY**

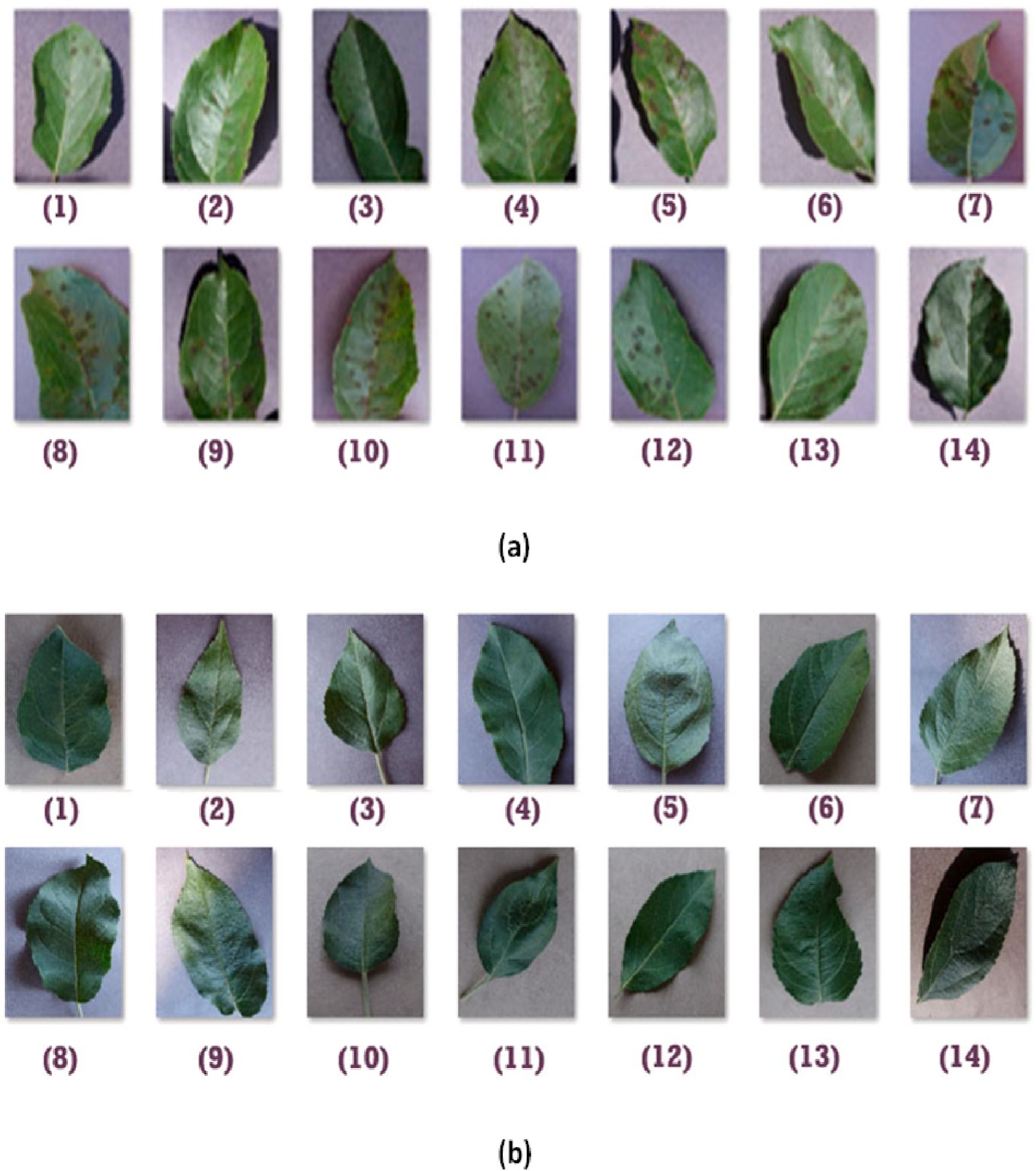
Retrieved from Maharashtra Agriculture website. The data were categorized into nine agricultural regions [14]. MLP neural networks were developed by activating new features and randomly reestimating yield, bias, and weight values ​​using several methods. An updated MLP neural network is developed using a new activation function and modified weight and bias values ​​for random yield estimation based on diverse weather datasets. The MLP model was used to test the validated activation function including bias and weights. In this study, we analyze the results of different activation potentials and propose various basic functions such as DharaSig, DharaSigm, and SHBSig to improve the efficiency and accuracy of neural networks. Additionally, three additional activation functions with slight differences were developed using the DharaSig functions DharaSig 1, DHaraSig2 and DharaSig3. We proposed the ET0 method to evaluate data studied in semi-arid China using limited climate data. The k-Nearest Nearest Neighbor algorithm was used in China [11]. We also tested the PM-56 equation using the KNN-dependent ET0 prediction model. Traditional agricultural methods involve collecting data manually.

Dealing with adverse weather conditions, spraying pesticides against diseases and other practices that put farmers' lives at risk, especially during periods of drought.

Areas prone to occurrence. Concerning the current situation in conventional farming, there has been an urgent need for predicated data in farming that can assist farmers in identifying and responding to real-time problems.

To help them solve their problems, we’d like to propose a method that uses a RCNN,Keraz to predict apple,tomato and grapes crop diseases based on

temperature, soil moisture, and other variables [33,41]. Focuses on using plant images for supervised machine learning detection of tomato plant diseases, such as Naive Bayes (NB), Decision Tree (DT), Nearest Neighbor (KNN), Supported Vector Machine (SVM), and Random Forest (RF).



**Fig. 1.** (a) Diseased leaf image samples and (b) healthy leaf imag

IV .ALGORITHM DESCRIPTION

Here, the random forests classifier is used to implement the algorithm. They are adaptable and suitable for both regression and classification methods. Random forests yielded higher accuracy with less picture data sets when compared to other machine learning approaches such as SVM, Gaussian Naïve Bayes, logistic regression, and linear discriminant analysis. The architecture of our suggested method is depicted in the following image.

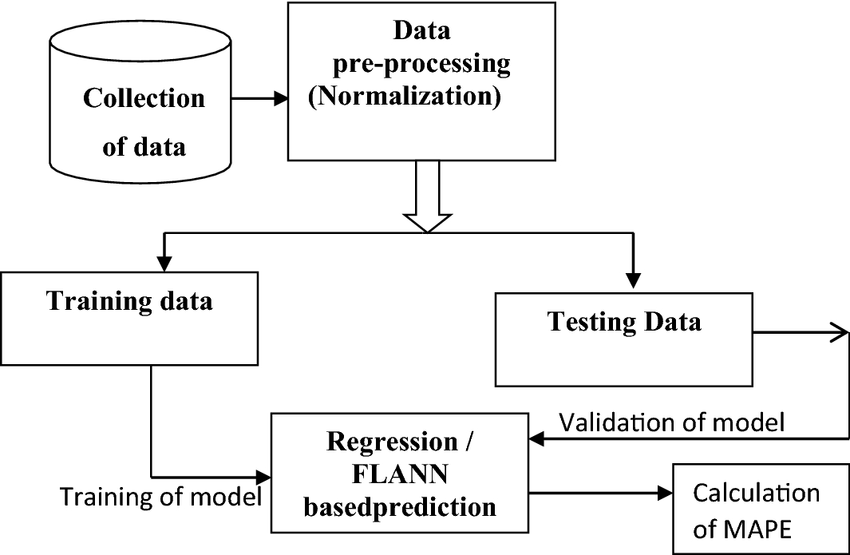


Fig.4. Architecture of the proposed model

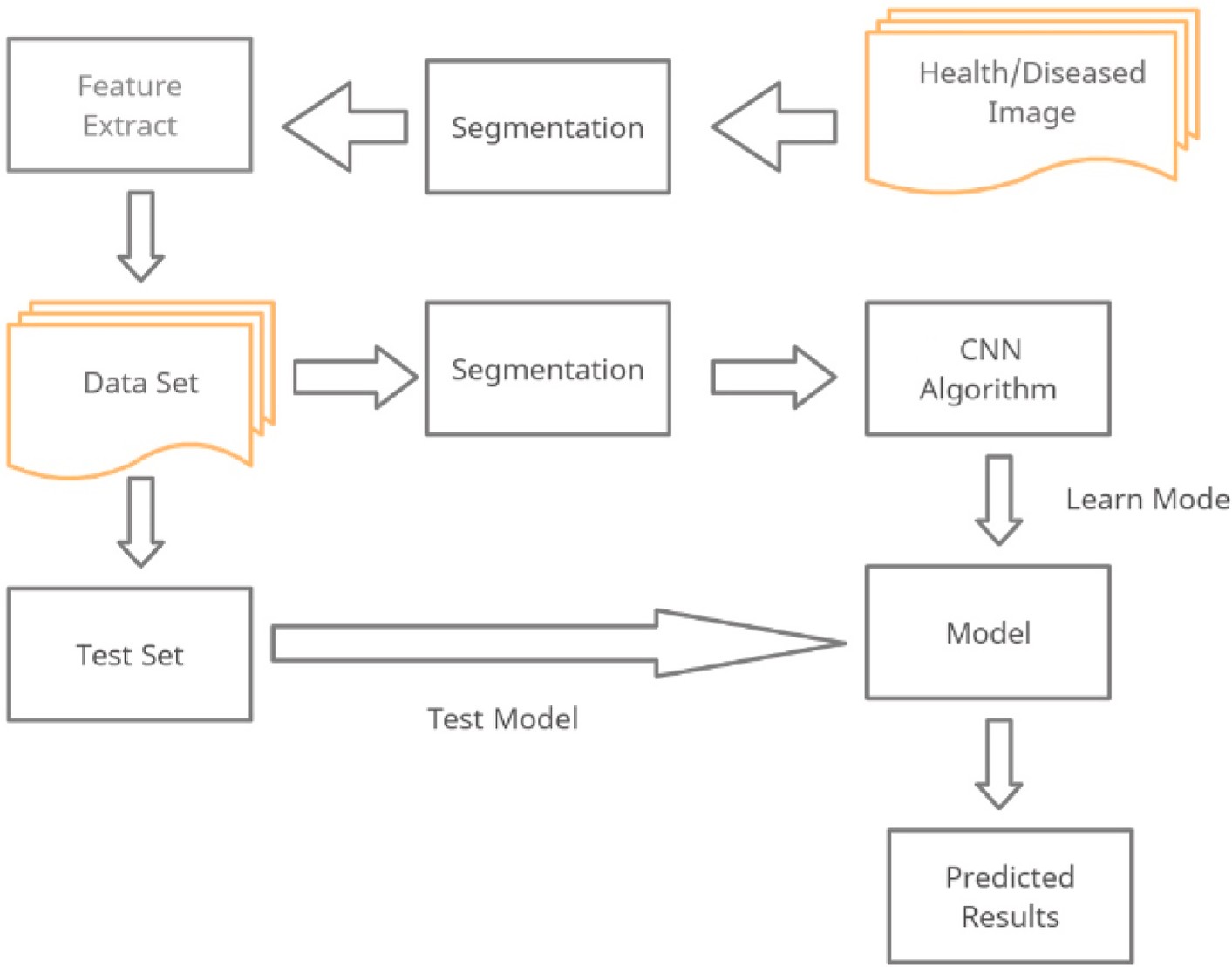


Fig.5. Flow chart for training.

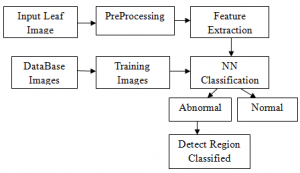
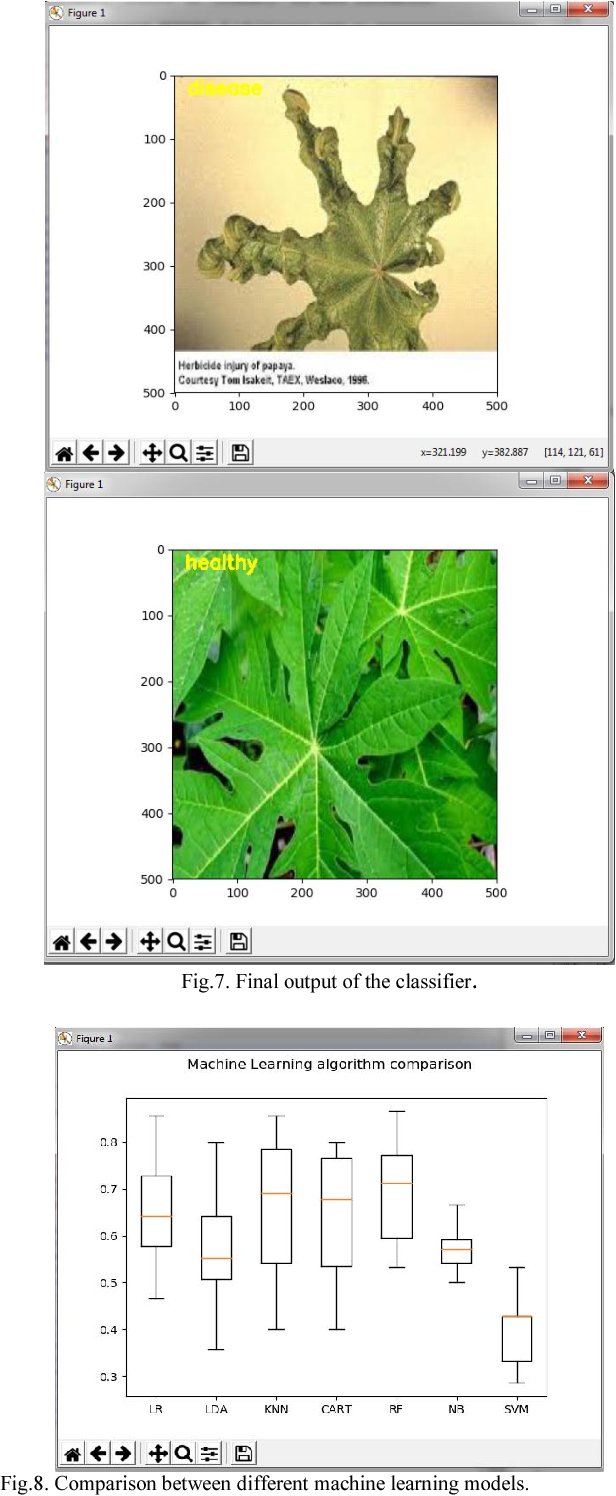


Fig.6. Flow chart for classification

Training and testing data are separated from the labeled dataset. HoG feature extraction is used to generate feature vectors for the training dataset. A random forest classifier is used to train the resulting feature vector. Additionally, the trained classifier obtains feature vectors for the test data obtained through HoG feature extraction for prediction, as shown in “Figure 4”. The labeled training dataset is converted into the corresponding feature vector through HoG feature extraction, as shown in Figure 2. 5. The training data set contains the stored extracted feature vectors. Additionally, a Random Forest classifier is used to train the trained feature vectors [9, 10].



## **Conclusions**

The most practical way to maintain an effective yield is through image processing for the detection of plant diseases. This paper's major objective was to demonstrate how an image processing tool can help farmers increase yields by helping to detect plant diseases accurately and providing correct results. By the project's conclusion, we had succeeded in achieving the goal, which was to use image processing to the detection of plant diseases. Additionally, developing the stand-alone application will increase farmers' access to and utility of this technology. Consequently, a stand-alone application for distinguishing between healthy and diseased plants has been created. In addition, future study aims to develop a mobile application to facilitate the procedure for farmers and utilise drones to increase the amount of training image dataset and boost the precision of our suggested approach.

**RESULT**

First for any image we need to convert RGB image into gray scale image. This is done just because Hu moments shape descriptor and Haralick features can be calculated over single channel only. Therefore, it is necessary to convert RGB to gray scale before computing Hu moments and Haralick features. As depicted in the figure 4.

To calculate histogram the image first must be converted to HSV (hue, saturation and value), so we are converting RGB image to an HSV image as shown the figure5.

Finally, the main aim of our project is to detect whether it is diseased or healthy leaf with the help of a RCNN,,Keraz which is as depicted in the “Fig.7.”

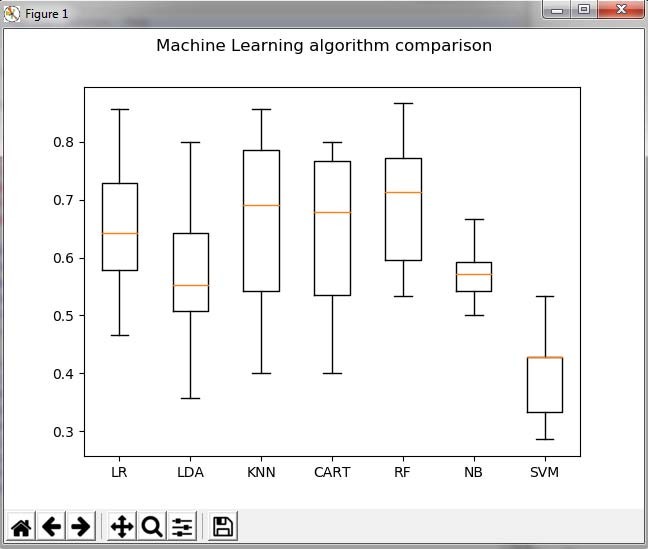


Fig.8. Comparison between different machine learning models.

TABLE I.

|  |  |
| --- | --- |
| Various Machine learning  models | Accuracy(percent) |
| Logistic regression | 65.33 |
| Support vector machine | 40.33 |
| k- nearest neighbor | 66.76 |
| CART | 64.66 |
| Random Forests | 70.14 |
| RCNN | 95.40 |

Fig .9. Table showing the comparison

**VI Challenges**

Identifying plant diseases from leaf images presents several challenges. Addressing these issues and challenges is key to developing practical systems for identifying plant diseases from real-time images under diverse field conditions. In this section, we discuss some unresolved issues in plant disease detection.

1.1. ***Data sets lacking in size and diversity***

In many papers and papers, the main limitation is the dataset used to train the CNN network, resulting in poor disease detection accuracy. DL requires large datasets containing a variety of images. The PlantVillage dataset [23] and the Plant Disease Symptom Image Database (PDDB) [137] are the only major disease datasets currently freely available. The available images are from a laboratory setting and were taken against a uniform background. However, collecting field images is expensive and requires agricultural knowledge to accurately identify diseases.

1.2. **Image segmentation**

Segmentation involves finding regions of interest in an image. There are two approaches to segmentation: traditional and soft computing approaches. K-means clustering and color thresholding are traditional and fuzzy logic, while artificial neural networks and region growing are soft computing-based segmentation methods. Segmenting leaf images against complex backgrounds is a challenging task for disease identification. Leaf region partitioning can improve performance accuracy. Images with many illegal elements are often difficult to identify.

1.3. **Identify diseases with visually similar symptoms**

Some diseases have similar symptoms that even experts cannot distinguish with the naked eye. Sometimes one of the symptoms of the disease may vary depending on the geographical location, stage of crop development and weather conditions. To date, no studies have been found in the literature that include these issues in the identification of plant diseases.

1.4. **Multiple diseases occur simultaneously**

Most plant disease identification models assume that only one type of disease is present in an image. However, several diseases can occur simultaneously, as well as some nutritional disorders. This may affect disease detection. From the survey, we can see that very little work has been done in this field to identify multiple diseases. Fuentes et al. [73] only considered the detection of several diseases in tomato leaves.

1.5. **Disease identification through real-time imaging ,**

Through the literature, we observed that most of the work is based on disease identification using laboratory images. For real-time disease detection, the model performs poorly. In [72], the authors achieved an accuracy of 99.35% on the PlantVillage dataset, and when the model was tested on another dataset, the model performance dropped to 31%. In [2], the authors recorded an accuracy of 99.53% on a variety of datasets. When the model was trained only on laboratory images and identification images obtained in the field, the success rate dropped to 66%. Therefore, efficient disease identification through real-time field imaging is an important challenge. 1.1. Creating a lightweight deep learning model

Most deep learning architectures implemented in the literature are based on AlexNet, VGG, GoogleNet, ResNet, DenseNet, and InceptionV3. Deep learning requires high-performance computing devices, expensive GPUs, and hundreds of machines. This increases costs for users. Small CNN models would be highly desirable. especially

used in embedded, robotics, and mobile applications that require real-time performance and low computational costs. It requires very large amounts of data to perform better than other methods. Training costs are very high due to complex data models.

REFERENCES

\ [1] S. S. Sannakki and V. S. Rajpurohit,” Classification of Pomegranate Diseases Based on Back Propagation Neural Network,” International Research Journal of Engineering and Technology (IRJET), Vol2 Issue: 02 | May-2015

* 1. P. R. Rothe and R. V. Kshirsagar,” Cotton Leaf Disease Identification using Pattern Recognition Techniques”, International Conference on Pervasive Computing (ICPC),2015.
  2. Aakanksha Rastogi, Ritika Arora and Shanu Sharma,” Leaf Disease Detection and Grading using Computer Vision Technology &Fuzzy Logic” 2nd International Conference on Signal Processing and Integrated Networks (SPIN)2015.
  3. Godliver Owomugisha, John A. Quinn, Ernest Mwebaze and James Lwasa,” Automated Vision-Based Diagnosis of Banana Bacterial Wilt Disease and Black Sigatoka Disease “, Preceding of the 1’st international conference on the use of mobile ICT in Africa ,2014.
  4. uan Tian, Chunjiang Zhao, Shenglian Lu and Xinyu Guo,” SVM-based Multiple Classifier System for Recognition of Wheat Leaf Diseases,” Proceedings of 2010 Conference on Dependable Computing (CDC’2010), November 20-22, 2010.
  5. S. Yun, W. Xianfeng, Z. Shanwen, and Z. Chuanlei, “Pnn based crop disease recognition with leaf image features and meteorological data,” International Journal of Agricultural and Biological Engineering, vol. 8, no. 4, p. 60, 2015.
  6. J. G. A. Barbedo, “Digital image processing techniques for detecting, quantifying and classifying plant diseases,” Springer Plus, vol. 2, no.660, pp. 1–12, 2013.
  7. Caglayan, A., Guclu, O., & Can, A. B. (2013, September). “A plant recognition approach using shape and color features in leaf images.” In International Conference on Image Analysis and Processing (pp. 161-170). Springer, Berlin, Heidelberg.
  8. Zhen, X., Wang, Z., Islam, A., Chan, I., Li, S., 2014d. “Direct estimation of cardiac bi-ventricular volumes with regression forests.” In: Accepted by Medical Image Com- puting and Computer-Assisted Intervention– MICCAI 2014.
  9. Wang P., Chen K., Yao L., Hu B., Wu X., Zhang J., et al. (2016).” Multimodal classification of mild cognitive impairment based on partial least squares”.