IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Detection And Prediction Of Crop Diseases Using Machine Learning Models

P. Venkat Rao[1], Sk. Sharooq[2], G. Prameela[3], V. Guna Venkatesh[4], K. S. G. Gowtham[5]

[1,2,3,4,5]Electronics and Communication Engineering

[1,2,3,4,5]Godavari Institute of Engineering and Technology (A)

Abstract: -- Agriculture, a fundamental human necessity from ancient times, remains crucial for our sustenance. Throughout history, plants have been our primary food source. Today, agriculture continues to be a vital aspect of our lives, serving as the backbone of various economies, irrespective of their developmental stages. Agriculture faces a significant challenge in the form of plant diseases and their impact on trade. Detecting and managing these diseases in a timely manner is crucial for crop health. Plant diseases are defined as natural issues that afflict plants, impeding their growth and potentially causing plant death in severe cases. In modern farming, new technology like IoT and automation can be super helpful. It's really important to keep plants healthy and check their surroundings to catch diseases early and get a good crop. We use smart tools like artificial intelligence (AI) and deep learning to look at plant pictures and find diseases. AI makes the job faster and helps us spot sick plants and control the farm's conditions. Scientists did some research and found out that these technologies are great at finding plant diseases by looking at the leaves. In this project we propose a deep neural network that would be trained to classify a variety of plant diseases. We intend to use PLANT VILLAGE dataset to train the model for the purpose of classification. Plant village consists of almost 16,000 images of leaves(some sick, some healthy) covering 19 different plant diseases.

Keywords: --PLANT VILLAGE; machine learning; disease detection; convolutional networks; modern farming

1. INTRODUCTION

The demand for agricultural products has surged due to recent rapid population growth, which has in turn caused a significant expansion of farming. Crop yield output must increase to fulfill the growing demand from the population for food, biofuels, and animal goods. One of the biggest issues facing agriculture and having an impact on commerce is the detection and treatment of plant diseases in a timely manner to enhance crop health. Plant diseases are, by definition, a class of natural disorders that impact plants' overall growth and, in severe situations, may result in plant mortality. Plant diseases can arise at any time during the various stages of plant life, including the formation of seeds, seedlings, and seedling growth. Traditionally, manual plant inspection was the method used by farmers, scientists, and even breeders to find and identify plant diseases. Naturally, in order to make the right detection, this procedure calls for experience and understanding. Over time, physical inspection become tedious, time-consuming, and less effective—especially when a lot of plants needed to be examined. comparable conditions that could be brought on by various infections that might have comparable effects on the plant are another issue that demonstrates the futility of manual inspection. In a number of industries, most notably medical, technology and the development of deep and machine learning proved advantageous. However, cutting-edge technical methods can also be used to identify plant illnesses. Because deep learning and machine learning are based on image processing techniques rather than destructive molecular or serological procedures, they can therefore be regarded as non-destructive disease detection tools. Nevertheless, the plant must to have had a discernible alteration due to the illness, mainly in the leaf region or stem, for these methods to function. But in order for these methods to be effective, the plant has to have already seen a noticeable alteration due to the illness, especially in the leaf region or stem. Plant disease detection can be substantially improved by the use of artificial intelligence, computer vision, and machine learning, which has already been done in a number of research studies. These technologies have the ability to identify the precise type of disease that is present in a given plant sample, as well as to detect its existence and assess its severity. Many applications of machine learning (ML) have shown promise, including the identification of diseases from medical images, the classification of images on big datasets, self-driving cars, and physics research. Though they are still in their infancy, machine learning applications for agriculture clearly show promise. For instance, well-known Convolutional Neural Network (CNN) architectures can be used to classify diseases from photos for various plants with various ailments. The overall process of using machine learning techniques to identify plant diseases. The data is first gathered and categorized into classes based on the types of disease or health. After that, a special dataset is built for the model so that the input photos can be pre-processed before feature extraction. Machine learning algorithms can distinguish between healthy and unhealthy outputs by comparing attributes and identifying changes in them.

2. LITERATURE SURVEY

Jeyalakshmi et al. [1] developed an approach based on machine learning for classification of diseases of potato and grape crops depending on their leaf images. The dataset was acquired from Pant Village dataset, consisting of 1000 healthy leaf images, as well as 2000 images of diseased potato leaves, and 3270 of diseased grape leaves. In the potato subset, there exist three different classes, while there are four classes in the grape leaves subset. From each RGB leaf images, the background was removed through using "enhanced Grab Cut algorithm".

Lamba et al. [2] published a paper discussing the implementation of several machine learning algorithms as well as deep learning techniques to properly detect diseases in crops. The images comprising the dataset are from Kaggle and rice dataset. In fact, four different datasets are created which are: the rice dataset containing 120 images, the pepper dataset containing 1,994 images, potato dataset containing 2,152 images, and tomato dataset containing 16,072 images. As usual, the images are pre-processed accordingly, and then autocolor correlogram filter is applied as an additional preprocessing step. After that, classification took place through machine learning techniques, or through deep learning methods with various activation functions to assess the overall performances.

Xian et al. [3] relied on a supervised machine learning termed Extreme Learning Machine "ELM" to determine the presence or absence of tomato diseases. The developed system was based on analyzing tomato leaves; the whole dataset was acquired from Kaggle, specifically the Plant Village dataset. The dataset comprises 10 classes, which makes a total of 1,000 images. The model is a feed forward neural network made up of several hidden nodes with weights that connect and the complete neural network has no iterations. Image preprocessing is done on the images in the dataset including image resizing, segmentation, color space conversion, and HSV.

CNN-based approach, namely, Few-Shot Learning (FSL) was proposed by Argüeso et al. [4] to identify and categorize the plant leaf affected portions. After performing the preprocessing step, the Inception V3 model was applied for computing the deep features from the input image. After the feature computation, a multiclass SVM classifier was used to train it over the key points to perform the classification task. The study by Argüeso et al. (2020) performs well to categorize the several diseases of plant leaves with an accuracy of 91.4%; however, evaluation results are discussed over a database of small size.

Deep CNN (Shoaib et al.[5]) is a type of feedforward AI model that is consisting of several hidden layers of convolutional and pooling layers, the CNN model are the best of the DL model for achieving higher detection accuracy using imaging data The CNN model consist of two blocks, the features learning and classification blocks. The features learning block extract various kind of features using the convolutional layer where the features learning is performed at the fully connected layers. The higher accuracy of the CNN model for plant disease classification has proofed to be the best then all other kinds of ML and DL methods. Studies have shown that CNNs can achieve high accuracy rates in the range of 99-99.2% in classifying images of plant leaves affected by diseases and pests.

DL technologies, such as CNNs and DBNs, have also been proposed for identifying plant abnormalities and infestations. These technologies have been showing promising outcomes in the detection and identification of lesions from digital images (Kaur and Sharma; Siddiqua et al.; Wang) [6]. DL models can automatically learn features from the images and can identify subtle symptoms of diseases that traditional image processing methods may not be able to detect. Though, Deep Learning models necessitate a significant volume of labeled training data and involve intensive computational resources, which may be a limitation for some applications.

Another AI technology that has been applied to plant pathology is computer vision (CV). CV algorithms, such as object detection and semantic segmentation, can be used to identify and localize specific regions of interest in images, such as plant leaves and symptoms of diseases (Kurmi and Gangwar, 2022; Peng and Wang, 2022) [7]. These algorithms can be used to automatically transforming the images into recognizable patterns or characteristics can be integrated with ML or DL algorithms for disease detection and classification. However, CV algorithms need a huge number of labeled image data for model training and may not be suitable for diseases that have not been seen before.

Another CNN-based approach was presented by Agarwal et al [8]. To detect and classify tomato crop disease. The approach comprises 3 convolution layers together with max-pooling layers to compute the deep features from the suspected images and categorize them. This framework shows robust tomato disease recognition performance with the classification score of 91.2%, however, suffers from the model over-fitting problem.

3. METHODOLOGY

Obtaining a varied collection of crop photos, comprising both healthy and diseased samples, is a necessary step in the process of employing deep learning models to detect and forecast crop diseases. To improve the resilience of the model, preprocessing methods including image augmentation and normalization are used. Dense layers are used for classification, and a convolutional neural network (CNN) architecture is selected for feature extraction. For better performance, pre-trained models can be used with transfer learning. A different test set is used to validate the model's performance once it has been trained on labelled data. The trained model can be deployed on edge devices or cloud platforms to provide real-time illness detection. In order to ensure adaptation to changing disease patterns, the model is continuously monitored and updated with new data. This helps to create a more robust and productive agricultural system.

Training and Testing:

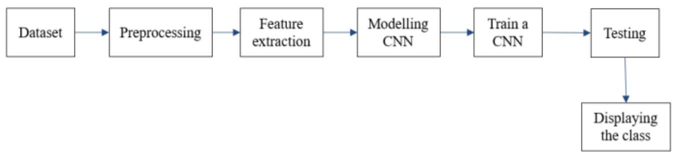

Fig.1. Block Diagram of Training and Testing

Validation:

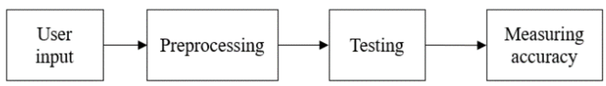

Fig.2. Block diagram of Validation

Flowchart:

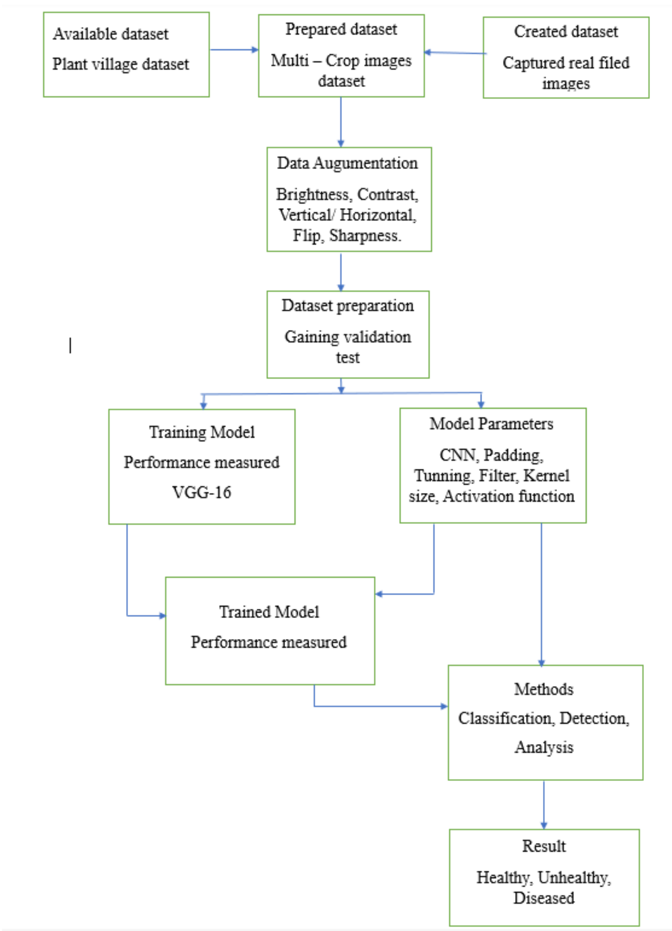

Fig: 3 Flowchart

4. CONVOLUTIONAL NEURAL NETWORK

The Convolutional network algorithm is mainly used for the feature extraction of images and objects. This algorithm will work on the pixels of the given images. CNN is one of the best network architectures for identifying and recognizing objects and makes it highly suitable for computer vision tasks and for applications of recognition techniques. CNN consists of three layers. In these layers, only extraction of an image will take place through various functions.

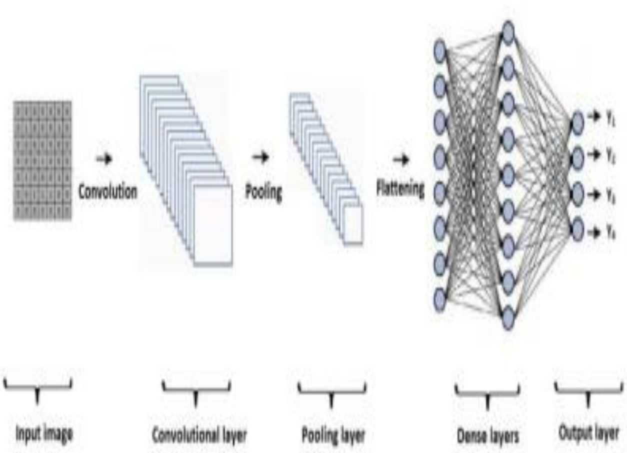

Fig:4 CNN Architecture

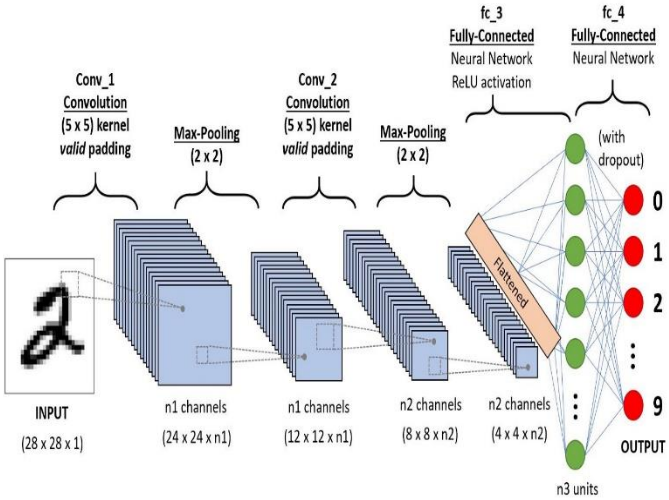

Fig 5: Workflow of CNN

4.1 Convolutional layer: The convolutional layer is used to filter the matrices in an image and for detecting the patterns in the image. These matrices will blur the images, sharpen and the edges of the images. Filtering of the image can be done in 3*3 matrices. If the image array is greater than the filter, size, then we need to slide the filter matrix over image and compute dot product to detect the patterns.

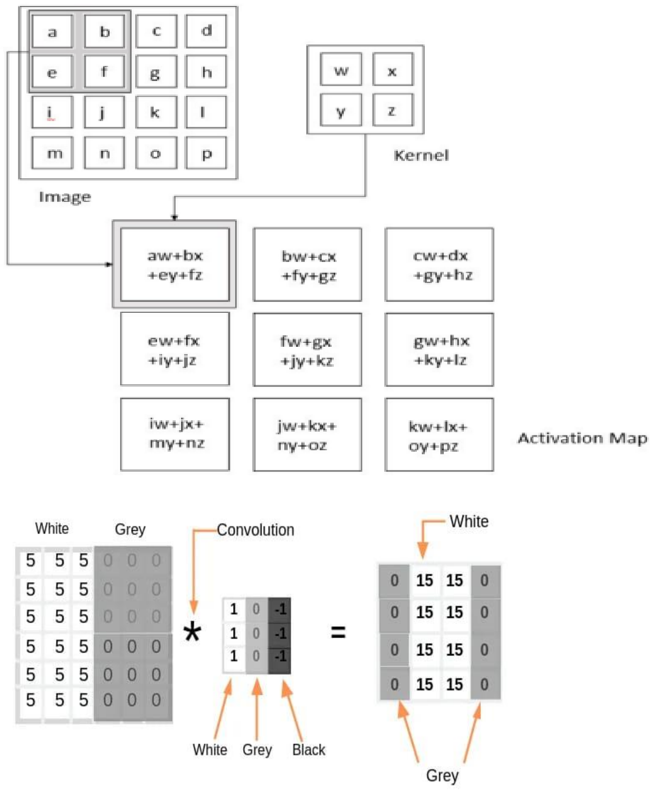

Fig:6 Convolutional layer

4.2 Pooling layer: This layer is responsible for rectifying features by using various filters for identification of different parts of images like eyes, corners and edges etc., the pool layer consists of two types: max pooling and average pooling. Max pooling will give the maximum value of the array, and it will reduce the size of the matrix and in average pooling we need to average all the matrices to reduce the size of the matrix.

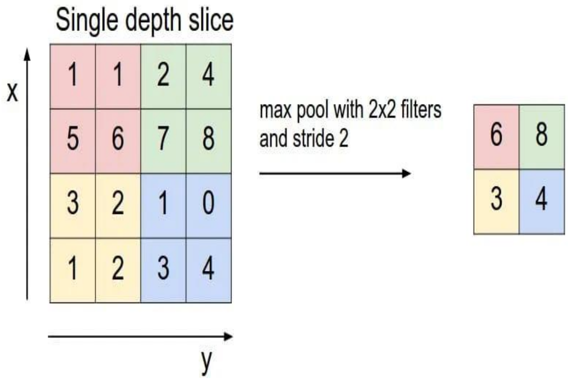

Fig:7 Pooling layer

4.3 Fully connected layer: This layer will occur after completion of convolutional and pooling layers. A fully connected layer is mainly used for the connection between all neurons from the previous layers to neurons of the current layer for predictions or classification. The main role of this layer is to transform the extracted features from layer to layer.

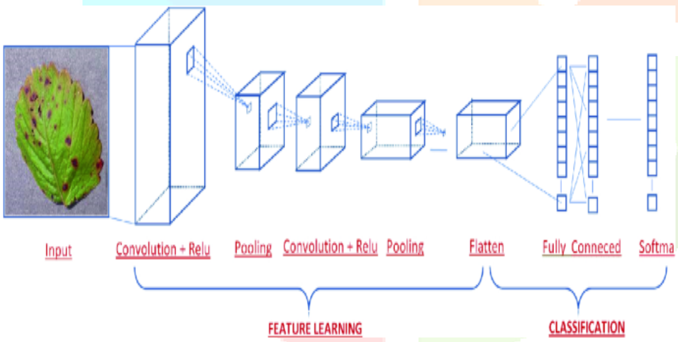

Fig:8 Fully connected layer

5.Results and Discussions

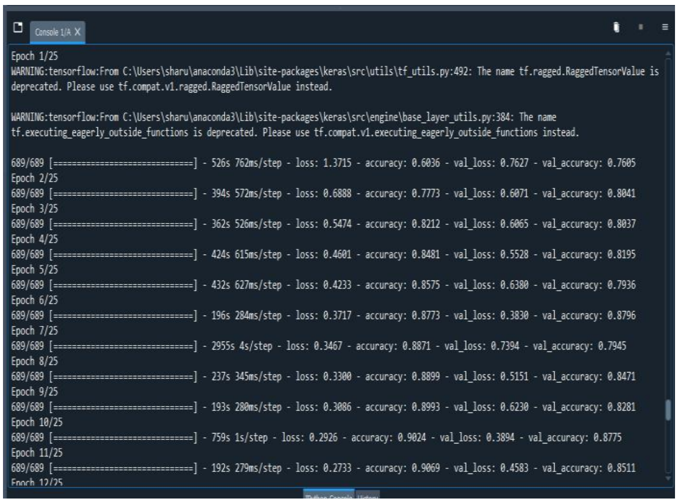

Fig 9: Training and validation of dataset

Fig 10: Training and validation of dataset

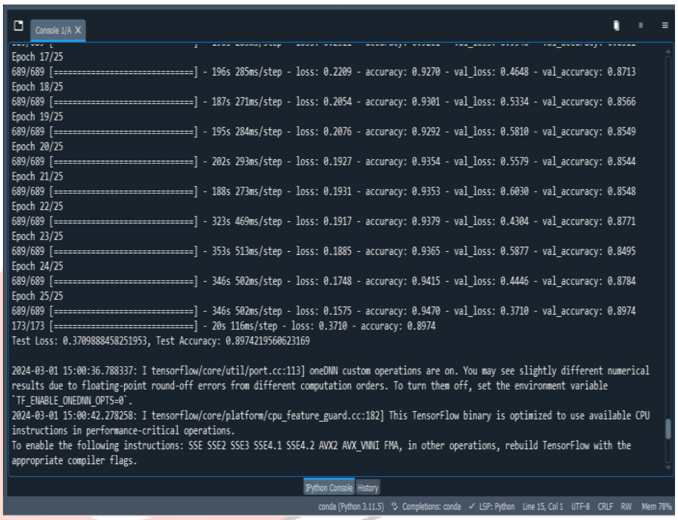

Fig 11: Training and validation of dataset

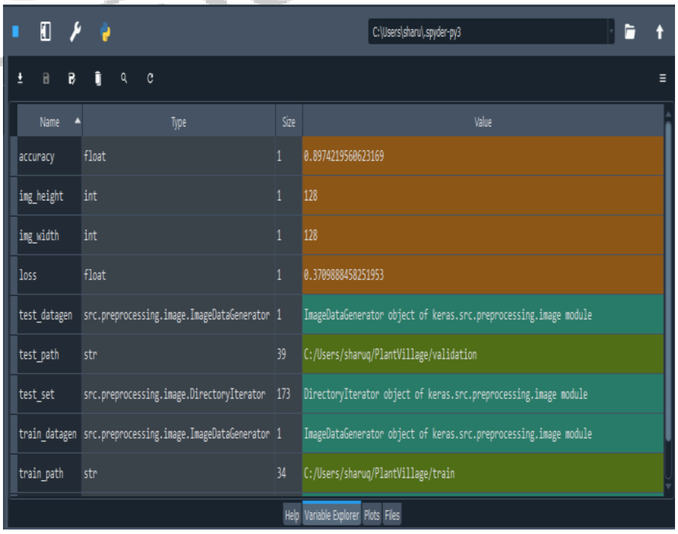

Fig 12: Training and validation of dataset

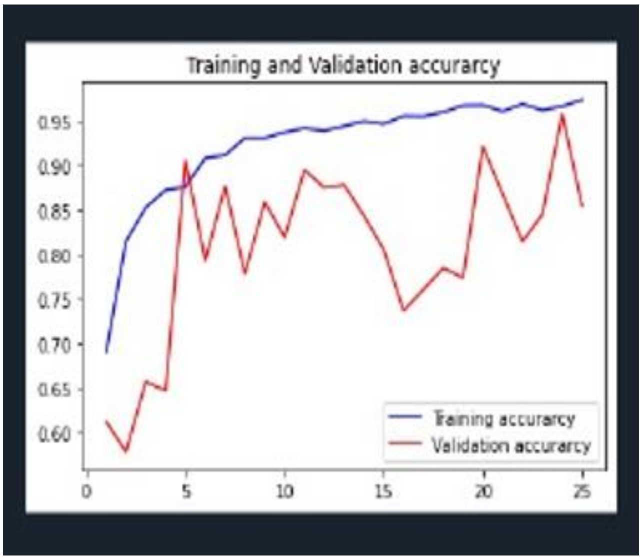

Fig 13: Training and Validation Accurarcy

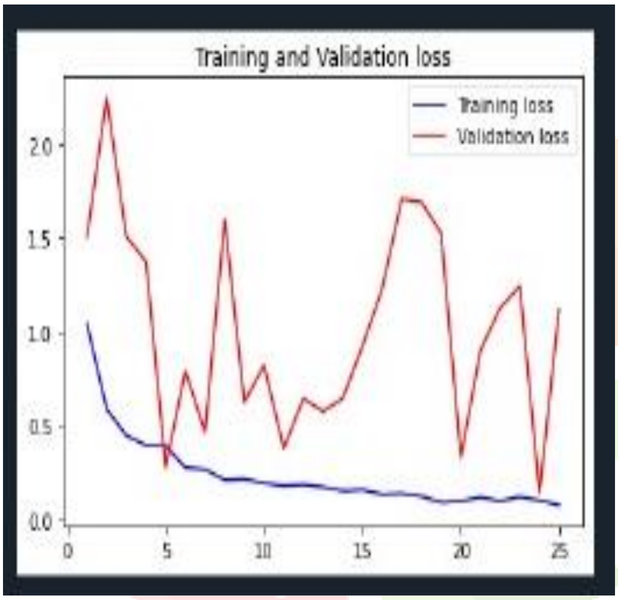

Fig 14: Training and Validation Loss

6.Conclusion & Future Scope

The proposed system uses deep learning models to find the missing individuals, which helps to solve the complex problems of locating individuals in various situations. When compared to earlier picture classification techniques based on "manual" feature extraction, computer vision and deep learning algorithms based on CNN models perform better at pest and illness detection and classification. Deep learning models, however, need a lot of data, which can be challenging to collect. Few-shot learning techniques or transfer learning may be helpful in addressing this problem. However, even though deep learning-based techniques perform well when applied to photos obtained in controlled environments, further study is needed to analyze photos captured in the field under actual circumstances.

Computer vision and deep learning techniques based on CNN models can be used for pest and disease detection and classification. These methods outperform traditional approaches to picture classification that rely on "manual" feature extraction. But deep learning models need a lot of data, and that can be hard to come by. Transfer learning and few-shot learning techniques can be helpful in addressing this problem. Still, more investigation is needed into the analysis of photos taken in the field, in actual circumstances, even though the performance of deep learning-based techniques is good for images obtained under controlled circumstances. The future scope is the potential for employing machine learning models to identify and forecast agricultural diseases is bright. Integration with drones, IoT devices, and satellite imagery will improve real-time monitoring as technology develops. With the development of AI algorithms, diseases will be accurately identified through the use of a variety of data sources. This will optimize agricultural productivity and minimize losses by providing farmers with timely interventions. Precision agricultural innovation will be fueled by partnerships between agri-tech firms, academic institutions, and government programs, resulting in the development of a robust and sustainable farming ecosystem. The ongoing advancement of machine learning in agriculture holds great potential for revolutionizing both agricultural sustainability and global food security.

It looks potential for the future to use machine learning algorithms for agricultural disease detection and prediction. Real-time monitoring will be improved as technology develops through integration with drones, IoT devices, and satellite imaging. As AI algorithms get more complex, they will be able to accurately identify diseases using a variety of data sources. By enabling prompt responses, farmers will be better able to maximize crop yield and minimize losses. Precision agriculture will become more innovative through partnerships with agri-tech firms, academic institutions, and government programs, which will build a robust and sustainable farming environment. Global food security and agricultural sustainability are expected to undergo radical changes as a result of machine learning's ongoing advancements in the field.

10.REFERENCES

1. Agarwal, M., Singh, A., Arjaria, S., Sinha, A., and Gupta, S. (2020). ToLeD: Tomato leaf disease detection using convolution neural network. Proc. Comput. Sci. 167 (2019), 293–301.
2. S. Jeyalakshmi and R. Radha, "An effective approach to feature extraction for classification of plant diseases using machine learning," Indian Journal of Science and Technology, vol. 13, pp. 3295-3314, 2020.
3. M. Lamba, Y. Gigras, and A. Dhull, "Classification of plant diseases using machine and deep learning," Open Computer Science, vol. 11, pp. 491-508, 2021.
4. T. S. Xian and R. Ngadiran, "Plant diseases classification using machine learning," Journal of Physics: Conference Series, vol. 1962, p. 012024, 2021.
5. Few-Shot Learning approach for plant disease classification using images taken in the field Author links open overlay panel
   David Argüeso, Artzai Picon, Unai Irusta, Alfonso Medela, Miguel G San Emeterio, Arantza Bereciartua, Aitor Alv arez-Gila.
6. Deep learning for plant bioinformatics: an explainable gradient-based approach for disease detection Muhammud Shoaib Front. Plant Sci., 13 October 2023 Sec. Plant Bioinformatics, Volume 14 – 2023.
7. Identification of plant diseases and distinct approaches for their management. October 2021. Bulletin of the National Research Centre 45(1):169.
8. An advanced deep learning models-based plant disease detection: A review of recent research. (Kurmi and Gangwar, 2022; Peng and Wang, 2022) Published online 2023 Mar 21.
9. Abdulridha, J.; Ampatzidis, Y.; Kakarla, S.C.; Roberts, P. Detection of target spot and bacterial spot diseases in tomato using UAV-based and benchtop-based hyperspectral imaging techniques. Precis. Agric. 2020.
10. I. Ahmad, M. Hamid and S. Yousaf, "Optimizing Pretrained Convolutional Neural Networks for Tomato Leaf Disease Detection," Complexity, 2020.
11. K. Akshai and J. Anitha, "Plant disease classification using deep learning," in International Conference on Signal Processing and Communication (ICPSC), 2021.
12. B. V. Gokulnath and G. Usha Devi, "A survey on plant disease prediction using machine learning and deep learning techniques," Inteligencia Artificial, vol. 23, pp. 136-154, 2020.