IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 45-51

Deep Learner Based Earlier Plant Leaf Disease Prediction and Classification Using Machine Learning Algorithms

Mr.M.Vengateshwaran M.E., Dr. E.V.R.M. Kalaimani Ph.D
Assistant Professor, Department of CSE Arasu Engineering College, Kumbakonam Tamilnadu, India
Professor & Head, Department of CSE Arasu Engineering College, Kumbakonam Tamilnadu, India

Abstract: *Machine Learning makes changes in the world in recent days. Machine Learning is now well developing and reduces a human effort. Agriculture is the art and science of growing plant and other crops for rising economic gain. However, the diseases that affect the Plant leaf have make an impact of agriculture production. The Plant leaf usually gets infected by pathogens such as bacteria, fungus and virus. In order to address the above issue a novel prediction system for earlier detection of plant leaf diseases. In that case of situation, Machine Learning plays a major role to provide an effective solution through monitoring the whether the leaf is healthy or not. Deep Learning algorithm are used here to acquire the data from various plant health parameters and these data's are stored into the cloud. Stored data can be viewed through website from anywhere. In this paper our proposed a deep learning algorithms automatically to detect and classify the plant leaf disease using image datasets. Plant leaf disease detection and identification includes the stages like image acquisition, image pre-processing, image segmentation, feature extraction and classification. This paper discusses techniques for image pre-processing, image segmentation algorithm used for automatic recognition and research on various plant leaf disease classification algorithms that may be used for leaves disease classification. The outcome of the proposed project is represented through graphical visualization, thereby providing better prediction system projecting the high similarity symptoms.*

Keywords: *Machine Learning, Deep Learning, Image*

I. INTRODUCTION

Success of image processing and its expansion to numerous fields of applications like medical, engineering and remote sensing has paved its way to application in agriculture. Agriculture is a most important thing in the world. In farmers to cultivate the various food items like paddy, maize, sugarcane, paddy etc., Input image taken in real time is processed and transformed into useful information as an output to support farmers.

Image processing is used to processing a image and analyzing a image from various fields to generate a higher performance result. Different applications of image processing in agriculture sector are discussed following Fig 1.

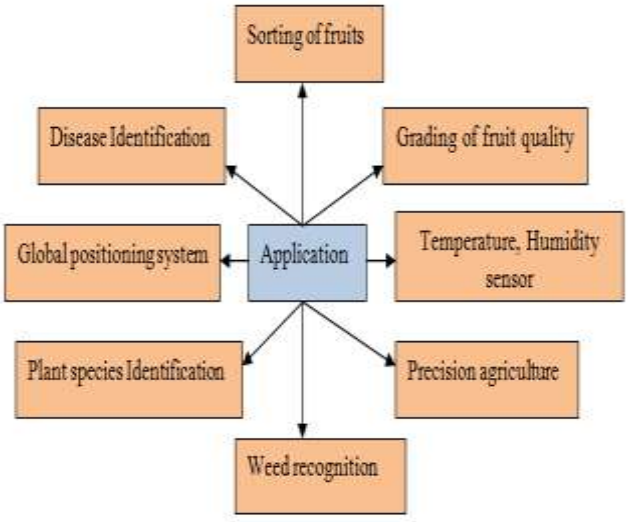

Fig.1 Applications of image processing in agriculture

In farmers are cultivate the wide range of vegetables and fruits. Now days so many Research work doing in the field of agriculture. It automatically to detect and classify the plant leaf disease.
Initially different images are captured by camera or using image database. Noisy or unwanted images are removed and then split the images for training and testing.

Image Analysis can be applied for the following purposes:
1. To detect diseased leaf, stem, fruit.
2. To quantify affected area by disease.
3. To find the range of the affected area.
4. To identify the infected area using color
5. To determine size & shape of leaf.
6. To identify the Object correctly.

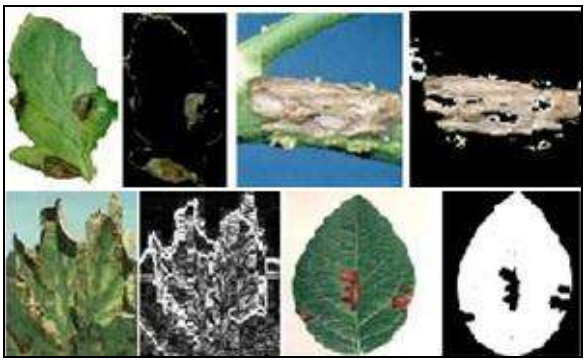

Fig. 2 Image analysis

II. DEEP LEARNING

Deep Learning is an advanced class of machine learning techniques that utilizes several layers of non-linear information processing for both supervised and unsupervised methods of feature extraction and transformation. It is extensively used for applications such as pattern analysis and classification.
Deep learning is a sub part of Machine Learning, For individual definitions:

- Artificial Intelligence is providing the intelligent systems.
- Machine Learning is automatically through experience and make predictions on data
- Deep Learning is a multiple layer of transformation.

Input values, or in other words our underlying data, get passed through this "network" of hidden layers until they eventually converge to the output layer. The output layer is used to predicting the result.

To figure out what these small weights should be, we typically use an algorithm called back propagation.

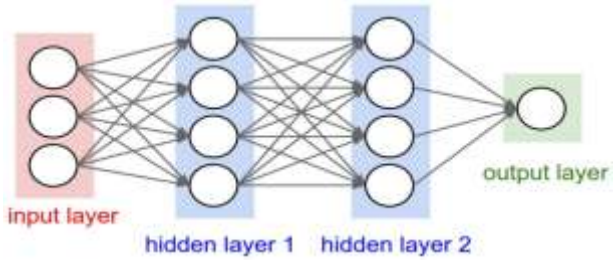

Fig.3 Neural Network Layer

III. LITERATURE SURVEY

A. Kamilaris, F.X. Prenafeta-Boldu., (2018), Deep learning constitutes a recent, modern technique for image processing and data analysis, with promising results and large potential. As deep learning implemented in all domains like agriculture, healthcare etc., Perform a survey of 40 research efforts that employ deep learning techniques, applied to various agricultural and food production challenges. To identify the agricultural problems based up on the previous study and models. It provide preprocessing of a data and achieve the better performance according to the metrics used at each work under study. Moreover, study comparisons of deep learning with other existing popular techniques, in respect to differences in classification or regression

performance.

J. Lu et al., (2017) Crop diseases are responsible for the major production reduction and economic losses in agricultural industry worldwide. Monitoring the plant leaf health condition status. This paper presents an in-field automatic wheat disease diagnosis system based on a weekly supervised deep learning framework, Furthermore, a new in-field image dataset for wheat disease, Wheat Disease Database 2017 (WDD2017), is collected to verify the effectiveness of the system. Under two different architectures, i.e. VGG-FCN-VD16 and VGG-FCN-S, system achieves the mean recognition accuracies of 97.95% and 95.12% respectively over 5-fold cross validation on WDD2017, exceeding the results of 93.27% and 73.00% by two conventional CNN frameworks, i.e. VGG-CNN-VD16 and VGG-CNN-S. Experimental results demonstrate that the proposed system outperforms conventional CNN architectures on recognition accuracy under the same amount of parameters, meanwhile maintaining accurate localization for corresponding disease areas. Moreover, this system has been packed into a real-time mobile app to provide support for agricultural disease diagnosis.

IV. MACHINE LEARNING

Machine Learning is the study of Computer algorithms that improve automatically through experience and make predictions on data. Machine Learning is a computer program that is said to learn from experience E with respect to any class of tasks T and performance measure P, if its performance at the task improves with the experiences (Mitchell 1997). Machine learning algorithms are either classified as supervised or un-supervised depending upon the labeling of data. Supervised learning is based on the training data with pre-defined label where the class label of each instance is known in advance. Classification is an example for supervised systems that learns the given example with the class label and assigns a correct class label for unknown instances. On the other hand, unsupervised learning is based on unlabeled training data. The patterns are grouped based on similarity and termed as clustering.

4.1Machine Learning Algorithms

Many supervised machine learning techniques were developed for document classification (Sebastiani 2002), some of the renowned classifiers are Regression model, k-Nearest Neighbor (k-NN), Decision Tree classifier, Naive Bayes (NB), Support Vector Machines (SVM) and Neural Networks. The un-supervised machine learning algorithm is a Clustering that generates groups or clusters of related documents and the similarity among them.

4.1.1 k-NearestNeighbor (k-NN)

K-NN is employed to perform tests on the degree of likeliness between test documents and k training data. Here certain classification data are stored to determine the category of test documents(Ko & Seo 2000). This method is designated as an instant-based learning algorithm which categorizes the data objects based on the proximity of the feature space in the training set. Here training data are represented in a multi-dimensional feature space. This feature space in turn divided into regions based on the class of the training data. A data point in the feature space is mapped to a particular class if it is the most frequent classes among the k nearest training data. The distance between the data points are often calculated using the Euclidean Distance measure.

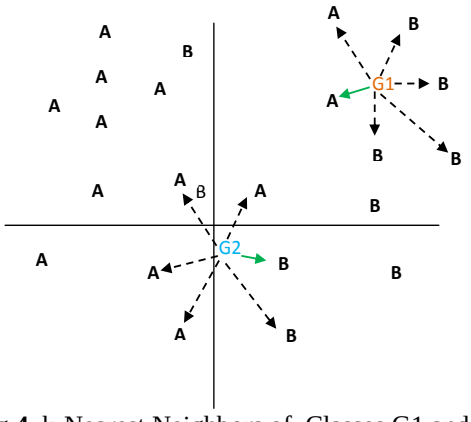

Fig.4 k-Nearest Neighbors of Classes G1 and G2

4.1.2 Naive Bayes (NB) Classifier

Naive Bayes classification is a probabilistic approach that does not require more instances for all possible combinations of attributes. Here, every attribute of interest assumed to be independent of each other. One of the important assumptions behind the technique is that the influence of an attribute is always independent of other attributes for a given class and this assumption is called as class conditional independence (McCallum& Nigam 1998). The joint probability of document features is calculated based on the probability that a new document fit in a specific class is defined in given below Equation

$$P(c_i|d') = \frac{P(d'|C_i).P(C_i)}{\sum_{cj \in C} P(d'|C_i).P(C_j)}$$

4.1.3 Genetic Algorithm (GA)

Genetic algorithm proposed by Holland (1992) performs better on hybrid problems dealing with both discrete as well as continuous data and on combinatorial problems. Here, three important processing elements such as selection operation, crossover operation and mutation operations are used to generate various learning models. GA is best suited for the problem domain where optimal solution is needed, but the possibilities of algorithm are restricted to stick at local optimal solutions. The computational requirements of learning system are not fulfilled by GA, hence it may not be concerned when there is strong computing power is demanded.

4.1.4 Neural Network Based Classier

A Neural Network (NN) text classifier is a network of processing units called neurons. Each input units represent words of the document, the output unit(s) represent the class or classes of interest, and the weights on the link that connects the processing units denotes dependence relations (Miguel Ruiz & Padmini Srinivasan 1998). To classify a given testing document d_j, the word weights wk_j are assigned to the visible units, where wk_j denotes the k^{th} term in document d_j. The activations of these input units are propagated forward through the network and the value of the output unit(s) determines the prediction decision.

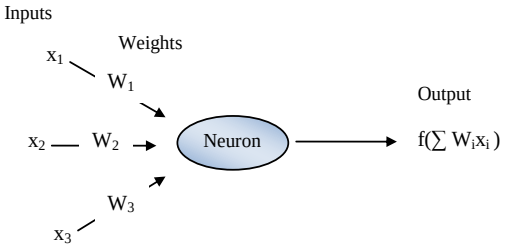

Fig.5 Structures of Artificial Neural Networks

4.1.5 Deep Learning Approach

According to Hinton (2007), Deep Learning is an advanced class of machine learning techniques that utilizes several layers of non-linear information processing for both supervised and unsupervised methods of feature extraction and transformation. It is extensively used for applications such as pattern analysis and classification.

Table 1.Comparison of Different Learning Algorithms

Approaches	Advantages	Limitations
Naive Bayes	Simple and quick classification	low accuracy
SVM	Find the best classification function.	Input space is high
k- Nearest Neighbor	Cost of learning process is zero. Complex task into simple procedure	Increase in the training data causes a decrease in efficiency
Genetic Algorithm	Provides an optimal solution	Provides a local optimal solution.
Neural Networks	Approaches an Expert's classification results	It require sufficient training data, Learning is to slow across multiple hidden layers
Deep Learner	Highly flexible	Involves multiple layers with complex structures

V. SYSTEM ARCHITECTURE

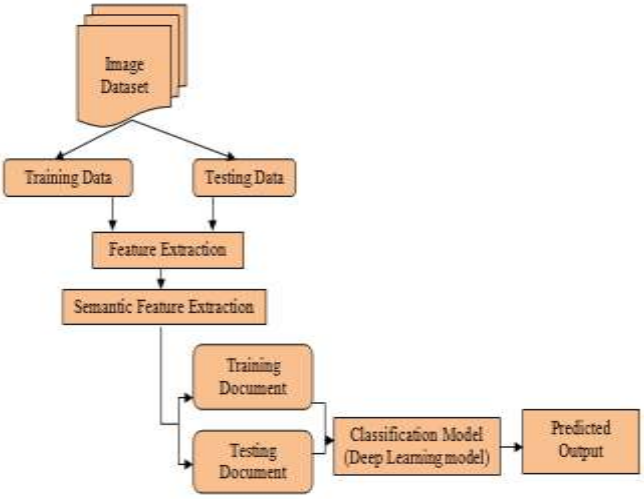

Fig. 6 System Architecture

VI. VPROPOSED SYSTEM

Plants are susceptible to several disorders and attacks caused by diseases. There are several reasons that can be characterizable to the effects on the plants, disorders due to the environmental conditions, such as temperature, humidity, nutritional excess or losses, light and the most common diseases that include bacterial, virus, and fungal diseases. Those diseases along with the plants may shows different physical characteristics on the leaves, such as a changes in shapes, colors etc. Due to similar patterns, those above changes are difficult to be distinguished, which makes their recognition a challenge, and an earlier detection and treatment can avoid several losses in the whole plant. In this project, we are discussed to use one of the recent object detectors such as Faster Region-Based Convolutional Neural Network (Faster R-CNN) for detection and classification of plant leaf diseases that affect in various plants. We consider some of the commercial/cash crops, cereal crops, and vegetable crops and fruit plants such as sugarcane, cotton, potato, carrot, chilly, brinjal, rice, wheat, banana and guava, these leaves images are selected for our purpose. Leaf images are collected in various resolutions from various sources. After that to perform first and second order gradient process on selected leaf image in pre-processing stage. Then Region of Interest will be finding by applying Adaptive mean and Gaussian thresholding on pre-processed image. Canny edge detector technique will be used for image segmentation. And then finally to find diseases affected area by applying contour function method. The challenging part of our approach is not only deal with disease detection, and also known the infection status of the disease in leaves and tries to give solution (i.e., name of the suitable organic fertilizers) for those concern diseases. The proposed approach comprises of four fundamental stages: Image acquisition of plant leaf images, pre-processing of images, image segmentation, feature extraction and classification of images in different disease classes.

6.1 Image Acquisition

The RGB color images of most frequently encountered pathological problems affecting leaves were captured using camera. Images were stored in any image format. In this research work the help of the research centers is taking which is working in agriculture department.

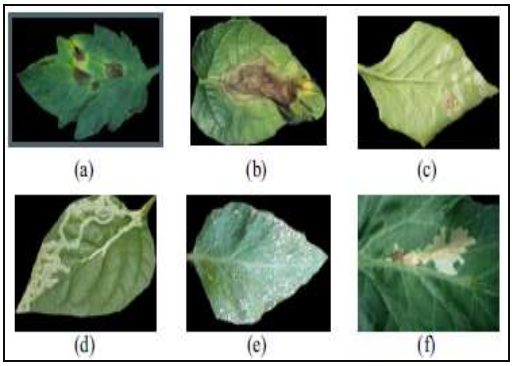

Fig.7 some images of database

6.2 Image Pre-Processing

In this phase to remove unwanted images or noisy images. The pre-processing involved the procedures to prepare the images for subsequent analysis. The affected leaf images were converted from RGB color format to gray scale images. Different pre-processing methods should be chosen for different image applications.

6.3 Image Analysis

In these phase, used to analyze the images from huge amount of image dataset. Object-Based Image Analysis (OBIA) contains classification and segmentation. Image segmentation is based up on the pixel. However, OBIA groups pixels into homogeneous objects. These objects can have different shapes and scale. It determines the color, text, size and shape of an image.

6.4 Feature Extraction

Feature extraction is the process of defining a set of features (characteristics) such as position, shape, size and texture which will more efficiently or meaningfully represent the information that is important for analysis and classification. These phases are used to extracting relevant features from image dataset.

6.4.1canny Edge Detector

These algorithms are used to extracting a procedural information from multiple objects. It also implemented in computer vision systems.

Process flow of a canny Edge Detector:

1. *Smoothing:* Blurring of the image to remove noise.
2. *Finding gradients:* The edges should be marked where the gradients of the image has large magnitudes.
3. *Non-maximum suppression:* Only local maxima should be marked as edges.
4. *Double thresholding:* Potential edges are determined by thresholding.
5. *Edge tracking by hysteresis:* Final edges are determined by suppressing all edges that are not connected to a very certain (strong) edge.

VII. CONCLUSION

Agriculture is not a simple thing. But most important thing in our world. There are several reasons that can be characterizable to the effects on the plants, disorders due to the environmental conditions, such as temperature, humidity, nutritional excess or losses, light and the most common diseases that include bacterial, virus, and fungal diseases. In our system specialized deep learning models were developed, based on the detection of plant diseases through leaves images of healthy or diseased plants. Our detector applied images captured in-place by various camera devices and also collected from various resources. Our experimental results and comparisons between various deep-architectures with feature extractors demonstrated how our deep-learning-based detector is able to successfully recognize different categories of diseases in various plants and also give solution for concern diseases. Healthy plants are living based on the soil contaminant with nutrition's are efficiently to resist pests attack. We hope our proposed system will make a suggestive contribution to the agriculture research.

FUTURE WORK

In future, the progress of identifying the plant diseases has been further modified to analyze various parts of plants such as stem, flower, leaf etc., and improve the performance speed.

REFERENCES

[1]. Alvaro Fuentes , Sook Yoon , Sang Cheol Kim and Dong Sun Park A Robust Deep Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition, Sensors 2017, 17, 2022; doi:10.3390/s17092022.

[2]. Andreas Kamilaris, Francesc X. Prenafeta-Boldu Deep learning in agriculture: A survey, Computers and Electronics in Agriculture 147 (2018) 70–90.

[3]. Bashir Sabah, Sharma Navdeep. Remote area plant disease detection using image processing. IOSR J Electron Commun Eng 2012;2(6):31–4. ISSN: 2278-2834.

[4]. C. C. Stearns and K. Kannappan, "Method for 2-D affine transformation of images," US Patent No. 5,475,803, 1995.

[5]. D. M. Hawkins, "The problem of over-fitting," Journal of Chemical information and Computer Sciences, vol. 44, no. 1, pp. 1–12, 2004.

[6]. Dai, J., Li, Y., He, K., Sun, J. R-FCN: Object Detection via Region-based Fully Convolutional Networks. arXiv 2016, arXiv:1605.06409v2.

[7]. Huang, J., Rathod, V., Sun, C.; Zhu, M., Korattikara, A., Fathi, A., Fischer, I., Wojna, Z.,Song, Y., Guadarrama, S., et al. Speed/accuracy trade-offs for modern convolutional object detectors. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 22–25 July 2017.

[8]. Jiang Lu, Jie Hu, Guannan Zhao, Fenghua Mei, Changshui Zhang, An in-field automatic wheat disease diagnosis system, Computers and Electronics in Agriculture 142 (2017) 369–379.

[9]. Konstantinos P. Ferentinos, Deep learning models for plant disease detection and diagnosis Computers and Electronics in Agriculture 145 (2018) 311–318.
[10]. Kulkarni Anand H, Ashwin Patil RK. Applying image processing technique to detect plant diseases. Int J Mod Eng Res 2012;2(5):3661–4.
[11]. Rakesh Kaundal, Amar S Kapoor and Gajendra PS Raghava "Machine learning technique in disease forecasting: a case study on rice blast prediction," BMC Bioinformatics, 2006.
[12]. Sankaran, S. Mishra, A.; Ehsani, R. A review of advanced techniques for detecting plant diseases. Comput. Electron. Agric. 2010, 72, 1–13.

Author's Profile:

M.Vengateshwaran M.E., Assistant Professor in CSE Arasu Engineering College Kumbakonam, India Specialization: Big Data, Machine Learning, Data mining

Dr.E.V.R.M Kalaimani Ph.D., Professor & HOD in CSE Arasu Engineering College Kumbakonam, India Specialization: Data Analytics , Machine Learning, IR