International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 11 | Nov 2020 www.irjet.net p-ISSN: 2395-0072

A REVIEW ON PLANT DISEASE DETECTION USING IMAGE PROCESSING TOOLS AND MACHINE LEARNING TECHNIQUES

BINZY NAZAR[1], SHAYINI R[2]

[1]*M-Tech Student, Department of Electronics and Communication Engineering, College of Engineering Thalassery, Kannur, Kerala, India*
[2]*Assistant Professor, Department of Electronics and Communication Engineering, College of Engineering Thalassery, Kannur, Kerala, India*

Abstract *Plants has vital role in our life. Cultivation of plants which offer healthy and nutritious food have a significant role in agricultural aspects. Different plant parts are infected by various diseases. Disease detection is incredibly difficult manually because of the regular changeability of leaf/fruit skin and also due to the presence of various sorts of diseases. This problem will be solved by developing a computer-based system, which may recognize the disease affected on surface of leaf/fruit from its image. The disease affected fruits/leaf has spots and cracks on its surface. These diseases will be detect using various styles of image processing techniques. By reviewing up to 10 papers, it can understand that image processing techniques like morphological analysis by feature combinations (color, texture and shape of pixels), k-means based segmentation algorithm, etc. are used to analyze the affected region. These diseases will be identified by classifying the extracted features of the image with the assistance of various machine language algorithms like SVM, MSVM. KNN etc. In a similar way various Deep Learning networks like CNN, DenseNet etc. are used for the identical purpose.*

Keywords- Morphological analysis, K-means based segmentation algorithm, Machine learning, CNN, MSVM, DenseNet, SVM classifier

1. INTRODUCTION

Agriculture plays a vital role in every human's lifestyle. In India, 70 percentage of people depends on the field of agriculture for their earning. But diseases affected on crops damages the crop yield. Plant diseases are the major problem faced by farmers. Leaves and fruits of plant are mainly affected by various bacteria and viruses which causes diseases. Technical experts all around the world are thinking about using which method they can help the farmers to detect the diseases. They found that using image processing as well as machine learning techniques easily solve the problem of disease detection. They use large number of images of healthy and infected plant parts like leaves and fruit for processing and then to identify the disease, because the symptoms of almost all diseases first appears on the surface of leaf or fruit of plants. By reviewing various journals associated with this subject it will be understand that 3 approaches are mainly wont to detect plant diseases. They are

1) State of the art approaches in image processing
2) Supervised learning
3) Un supervised learning

The stat of the art approaches are image acquisition, preprocessing, segmentation etc. Then disease affected region on the surface of natural object is detected, by masking all other unwanted areas in image. Then extract all feature values of the region. These are called feature data. From these feature data, we are able to classify the disease.

Supervised learning approaches are used to classify the features extracted on the image and hence to detect the disease. These are ML technique within which we have to gather data or produce data or produce data output from the previous experience. Here almost all diseases have same range of feature values. We will train a Machine learning algorithm like Support Vector Machine, K- Nearest Neighbor Classification with a large number of feature data extracted from the images. After training process an image given to the algorithm for testing. The algorithm automatically maps the given image to its corresponding class of images by comparing its feature values.

Un supervised learning could be a ML technique, where we don't have to supervise the model. It helps to seek out all quite unknown pattern or feature in data.

Different architectures od deep learning networks based on unsupervised learning are used in this era of research. In this method Deep Convolutional Neural Network is used for classification of disease affected and healthy Here various convolutional layers used to reduce image size without losing major features in the image.

Then it works on the basis of biological neural network as our brain. We can train the neural network with a large number of images in different class of diseases. After training procedure neural network can automatically map a new input test image to their corresponding class. The algorithm has the capability of learn the features of given training images. By comparing features of trained images with the given test image it can map the test image to its corresponding classes.

In this paper the comparative study of these technique carried out as a literature review. Up to 10 papers in which one among these techniques are used for the classification and detection of plant diseases, on the basis of the images of their fruit or leaves. Data set used, accuracy and methods are compared here.

2. REVIEW

1.1 STATE OF THE ART METHODS IN IMAGE PROCESSING TOGETHER WITH SUPERVISED LEARNING APPROACH

S.R Dubey et. al [1] uses fusion of color& texture cues for detection of plant diseases affected on Apple fruit, on the premise of their images. For image segmentation process K-means based segmentation is used. In this method K number of clusters identified from data by algorithm. In this algorithm, the data points are assigned to a cluster in such a manner that sum of squared distance between the data points and centroid becomes minimum. The disease affected area on the surface of fruit image will retain there and every other region are masked. From these image color & texture features are extracted and fused. them.

Extracted features are- Global Color Histogram (GCH), Color Coherence Vector (CCV), Color Difference Histogram (CDH), Structure Element Histogram (SEH), Local Binary Pattern (LBP), Completed Local Binary Pattern (CLBP).

They use combination of those features for detection and classification of diseases. Among these GCH+LBP & CDH+SEH+CLBP SHOWS more accuracy. Multiclass Support Vector Machine used for disease classification.

S.R Dubey et. al [2] also introduced Improved Sum and Difference Histogram from which texture features of images can be extract by comparing intensity values of corresponding red, blue & green channels. Gradient filters also are used with ISADH to improve the performance of this method. MSVM & KNN used for classification. Both of those methods show an accuracy of 99%.

1.2 UN-SUPERVISED LEARNING APPROACH

PENG JIANG et. al [3] introduces a Real time detection of apple leaf diseases using DL approach supported improved CNN with INAR-SSD. They use 26,377 images of Apple leaves infected by 5 forms of diseases. GoogLeNet inception module along with Rainbow concatenation is used for the designing of INAR-SSD. It helps to increase multiscale disease feature detection and also tiny diseased region detection capability. 2 GoogLeNet inception modules are accustomed improve multiscale feature fusion capability. Feature extraction & fusion structure is done by applying Rainbow concatenation method to enhance feature fusion performance. It improves inter connections among layers of CNN. Rainbow SSD helps to improve the performance of SSD algorithm.

QIU FENG WU et. al [4] introduced DCGAN-based data augmentation for Tomato plant disease identification. They use 1500 images of tomato leaves littered with 5 sorts of diseases. To reduce possibility of overfitting data augmentation is used commonly. In this method increase the number of images in dataset is carried out. For that purpose, DCGAN is using in this paper. It is a type of network I which G indicates generator and D indicates discriminator. Here an input noise is converted to image which is similar to real images used for classification. All pooling layers in Convolutional layers are replaced by discriminator and generator convolutional layers. By combining DCGAN with GoogLeNet we can improve performance of network with data augmentation. But generated images will not provide same images as real ones in some rare cases. This is the major disadvantage of this method. But it reduces cost of data collection, improves data diversity, and generalization capability of deep network model.

JUN SUN et.al [5] introduced Northern Maize leaf detection under complex field environment based on deep learning. Complex field of the field and different light intensity in the field makes the detection of disease

More difficult. In preprocessing the filter function of single scale retinex is used to solve the problem of high intensity light. Transmission module is used here to fuse features of low level and high level to improve detection accuracy of small target diseases. Generalized Intersection of Union (GIoU) is used here for optimization

UDAY PRATAP SINGH et. al [6] uses multilayer CNN for the classification of Mango leaves infected by Anthracnose diseases. Real time, efficient and law cost disease detection and monitoring system is proposed in this paper. Here Anthracnose disease infected and healthy Mango leaves of various species are classified using this system. For preprocessing of all images in data set 2 methods are used. One is contrast enhancement by histogram equalization and second one is rescaling using central square crop method. Then assign class labels to all collected images and train the model. After that, the system automatically maps given test image to its corresponding class. That is whether it is healthy or not. Here we can use another function rather than soft max activation function to reinforce the performance of CNN and build IOT enabled real time disease monitoring system as an extension of the work.

WENYAN PAN et. al [7] proposed a Smart Mobile Diagnostic System for Citrus diseases based on simplified Densely Connected CNN. They use 2000 images of fruit and leaves of Citrus plant infected by 6 forms of diseases. It's an intelligent diagnostic system supported by Mobile services computing. Here small number of images are there in data set, over fitting occurs. To avoid this problem data augmentation is employed. Here data augmentation of coaching set and test set by horizontal flip, vertical flip, increasing brightness and contrast of images in real time data set. For training and

testing DenseNet is using here. DenseNet means, each neuron in an exceedingly layer receives an input from all the neurons present within the previous layer. Hence they are called as densely connected, i.e. fully connected neural network. There are 4 Dense blocks in an exceedingly DenseNet. To cut back possibility of over fitting remove 5 bottleneck layers and added batch normalization activation function, global average pooling & softmax layers to make last dense block. Hence called simplified DenseNet. Upload the trained data set model within the cloud model. Users can upload image to the present system through We chat applet. System realized using this applet in mobile device. Users can upload images to it. Then name of the disease and its control measures are come as output. This method has the advantages as, reduce the vanishing gradient problem, strengthen feature propagation, encourage feature reuse and reduce number of parameters. They used a pretrained model, which reduces the time interval. We are able to expand dataset used for training and testing and further optimize method to attain acquires

3. COMPARISON

This paper discusses different methods of disease detection. While comparing these papers we will understand that State of the art approach gives more accuracy. Images of fruits are used for these experiments. During this method we've got to extract features of all images in data set, which could be a time consuming and difficult task. MSVM is usually used as a classifier for this method. Comparatively small number of images need for its proper execution

On the opposite hand different CNN models are used for disease detection and classification, like GoogLeNet, MCNN, DenseNet etc. So as to extend accuracy they use new and kind of data augmentation methods like DCGAN, INAR-SSD etc. In order to improve the performance of network they made modifications in layers of CNN networks. The experiments are mainly targeting leaves of the plants. This can be the foremost advanced and commonly used method for disease detection. This method needs large number of data for its proper and accurate working.

The comparative study of the mentioned models i.e. plant name, data set, method used & accuracy is given in the table1

4. CONCLUSIONS

To bridge the gap between farmers and plant diagnostic experts, various methodologies are used for plant diseases detection supported their images. This work needs sizable number of images as data set. Different methods are used for feature extraction and classification of diseases, like state of the art approaches and un- supervised learning approaches. Both of those methods give good results for disease diagnosis. It provides a technical assistance to farmers for fast detection of diseases and hence to enhance the productivity. Among these un- supervised learning approaches are most advanced and commonly used method in recent researches

TABLE 1

COMPARISON OF DIFFERENT MODELS

Author	Method used	Plant name	Disease	Dataset	Accuracy
S.R Dubey et al	Morphological features +MSVM	Apple fruit	Multiple	392 images	99%
S.R Dubey et al	ISADH+MSVM	Apple fruit	Multiple	392 images	99%
Peng Jiang et al	GoogLeNet+ INAR-SSD	Apple leaves	Multiple	26,377 images	94.33%
Qiufeng Wu et al	GoogLeNet+ DCGAN	Tomato leaf	Multiple	1500 images	91.3%
Jun Sun et al	SSD+TM+RPN +GIoU512+MCNN	Maize leaves	Northern Maize leaf blight	1000 images	97.14%
Uday Pratap sing et al	MCNN	Mango leaves	Anthracnose	2500 images	97.13%
Wenyan Pan et al	Simplified DenseNet	Citrus plant	Multiple	2000 images	88%

REFERENCES

1) Dubey, S.R., Jalal. S, "Fruit disease recognition using improved sum and difference histogram from images," Int.J. Appl. Pattern Recognition. 1(2), 199–220 (2014)

2) S. R. Dubey, S. Jalal, "Principle of disease classification using color, texture and shape features from images", Signal, Image and Video Processing, Springer (2015).

3) Pen Jian, Yuehan Chen, Bin Liu, Dongjian He, Chunquan Liang, "Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach supported Improved Convolutional Neural Networks", IEEE

4) Transaction, Volume 7, May 17,2019, Digital Object identifier10.1109/ACCESS.2019.2914929

5) Qiufeng Wu, Yiping Chun, Jun Meng," DCGAN- Based Data Augmentation for Tomato plant disease Identification" IEEE Transaction, Volume 8, June 5, 2020. Digital Object Identifier 10.1109/ACCESS.2020.299700

6) Jun sun, Yu Yang, Xiaofei He, Xiaohang Wu, "Northern Maize blight Detection Under Complex Field Environment supported Deep Learning" IEEE transaction, Volume 8, February 26, 2020. Digital

7) Uday Pratap Singh, Siddharth Singh Chouhan, Sukirty Jain, Sanjeev Jain, ," Multilayer Convolution Neural Network for the Classification of Mango Leaves Infected by Anthracnose Disease", IEEE transaction,2017,doi; 10.1109/ACCESS.2017

8) Wenyan Pan, Jiaohua Qin, Xuyu Xiang, Yan Wu, Yun Tan, Lingyun Xiang' "A Smart Mobile Diagnosis System for Citrus Diseases supported Densely Connected Convolutional Networks", IEEE Transaction, volume 7, July 18,2019, d o i; 10.1109/ACCESS.2019.2924973