European Journal of Molecular & Clinical Medicine
ISSN 2515-8260 Volume 7, Issue 4, 2020

Plant Disease Identifer Using K-Means and GLSM in Convolution Neural Network

S.P. Vijaya Vardan Reddy[1], T. Suresh[2], K. Naresh Kumar Thapa[3], V. Ramkumar[4], S. Mahabhoob Basha[5], Deepika.Y[6]

[1, 2,3,4,5] *R.M.K. Engineering College, Kavaraipettai, Chennai, Tamilnadu, India. 6 TATA Consultancy Services,*

Chennai, Tamil Nadu,India.

Abstract— Produces from agriculture which feeds the entire population is dependent on proper farming practices. The growth of technology must pay a way for increasing the produce per acre and also help in reducing the onset of frequently affecting plant disease. Timely help in detecting the diseases coupled with solution helps in productivity and quality of the produce. This paper aims to detect the plant leaf disease based on image detection and using machine learning to identify the disease with accuracy and suggest the solution. The product must cater to the needs of urban and rural farmer and also the person with only lay man knowledge of taking photo. This project mainly focuses on leaf disease like Anthracnose, Bacterial Blight, Cercospora, Alternaria Altermata diseases in the Pomegranate, Indian Beech, Tobacco, and Bitter Gourd leaves. This project aims to identify the disease even with lesser region of Interest and predict the leaf diseases using Convolutional Neural Network Algorithm.

Keywords—Image Processing, neural network, image segmentation, Convolutional Neural Network Algorithm

1. Introduction

Produces from the agriculture fields is plagued by pests and diseases which reduces the product quality and quantity. Even though farmers use enough amount of pesticides during the growth cycle, the problem persists. Timely identification of the diseases and its treatment is the biggest challenge faced by the novice farmers. Expert's opinion about detection of disease both native and non-native even in the remote areas will benefit the farmers. This project intends to have a strong dataset of various diseases and provide identification of the disease using image processing coupled with machine learning to provide us with basic knowledge to treat the disease. The dataset of the disease that are used in training are Alternaria Alternate Lycopersicon, Anthracnose, Bacterial blight, Cercospora. The dataset is divided into pathogen, fungal disease and bacterial disease. Alternaria Altermata is a pathogen that causes leaf spot in lot of plant species. Anthracnose is a fungal disease that attack leaves of plants during cool and wet weather. Bacterial blight is caused by the bacterial pathogen which causes pale green spots in leaves. This project has focused on lesser number of diseases with a greater number of photos to train the dataset. Going forward, we shall add more and more disease in the dataset.

In this project the major focus is on decision making algorithm. The proposed system consists of following steps: image capturing, extracting the deformed or decolored leaf part from image, which is the Image Processing cycle. We convert the image into a digital format and perform enhanced processing and extract valuable information. These digitized images are then feed into training network and the output is

checked to verify if any pattern matches with the data set. Segmentation are of many types such as Image, region and boundary based.

Image segmentation along with color for segmenting is used. As we are in need of information such as spatial arrangement of color or intensity of color in image for disease recognition, image segmentation is chosen. The image texture can be calculated from the captured scenes in an image. In our project we use Otsu method which is simple and efficient to set a threshold for the method. If the target cannot be distinguished clearly from the back ground and if we obtain bimodal condition, then we need to move to k-means algorithm. Unsupervised training in the recent times have caught up and delivers the requirement with sustainable performances. The properties which add to its fame are the processing data, neural network models and learning theory from the mistakes. The proposed system will shed the light on these properties in elaborate manner.

2. literature review

In the paper [1], the segmented Image provides input for Decision Support Systems which acts as brain to provide suggestion to farmers. The image feed in the system is converted from RGB to HSV (Hue Saturation Value) colour the input photo feed into the system is processed to find the region, known as the Region of interest (ROI) to, confirm the health of the leaf. The referred paper uses Scale Invariant Feature Transform (SIFT) algorithm to decide the necessary result.

In this manuscript [2-4], infected leaf region in legumes is proposed. It gives us details such as healthy and the area of diseased section. The data set focuses only on the legumes and broader classification of dataset is required. The various methods seen in the literature review mostly depends on image captured. The region of interest plays a vital role in all the papers. Every paper shows its variation in Image Segmentation and also varies in algorithm used in decision making. The probability of identification increases by removal of complex background from the images.

The paper [5-8] uses Generic algorithm which is used for automatic detection of little disease in pine trees. In the recent years, development of complex algorithms helps in precise prediction of the disease. In the paper [9-10] the decisions are made by rule based semi – automatic system using k-means algorithm.. In this paper Soybean plant algorithm is used to differentiate between three types of diseases (Septoria leaf, frog eye and Septoria leaf blight). Background error is removed using colour and texture parameter. Support vector Machine classifier is utilized to achieve efficient detection of the disease by training three model (each disease utilizes one model) from the dataset. The K-means clustering algorithm is used to remove the background in the input data.

In the paper [11], Subtractive clustering method is used. In this method centroid is generated based on the data points and set as initial centre point and used for segmentation of image to remove unwanted background and get the ‘Region of Interest’.

In paper [12], Gray Level Co-Occurrence Matrix is sued to extract image features. The parameter such as Angular second moment, Corelation, Entropy are extracted using GLCM. The results obtained using GLCM is good with less computational time. But this method cannot be used for multiple organ of plant. Only generic features are used to recognize the plant. Use of Hybrid Generic organ convolution neural network (HGO-CNN) is used in the paper to identify the plant based on any organ using deep learning.

In the paper [13], A new frame work is proposed where they use recurrent neural network (RNN). Here one photo of each part is used to build the database for identification of the plant. Thus disease can be accurately calculated using various plant parts of the diseased plant.

3. Proposed system

The Project focuses on removing the background and obtaining the area of interest so that the decision-making system is able to detect the diseases such as Alternaria alternate, Anthracnose, Bacterial blight, Cercospora leaf spot of Pomegranate bitter gourd and tobacco leaf based on the organs of the plants.

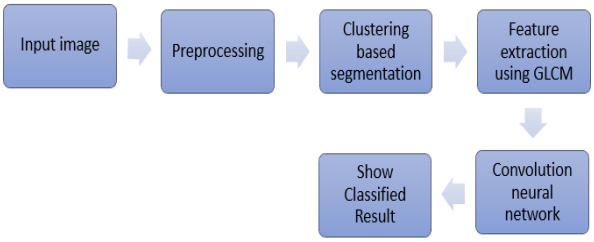

Fig-1 Proposed System

3.1 COLOR TRANFORMATION

The photo fed into the system is RGB image. For extracting the Attribute, we convert RGB to HSV colour space representation. Since the colour is related to Hue, we extract only particular information and neglect other parameters.

3.2 MASKING AND REMOVING GREEN PIXELS

Once the color information is extracted, we need to mask the green color portion of leaf. The green color states that the organs of the leaf are healthy. So, we identify the green color and set it zero if it less than threshold value.

3.3 K-MEANS CLUSTERING ALGORITHM

After the masking step, using k-means clustering the image is divided into k number of disjoint cluster. Followed by calculation of k centroid. In the subsequent step the nearest centroid from the cluster to the respective data point is calculated.

3.4 IMAGE SEGMENTATION

In our Paper, Euclidean Distance measures the nearest centroid. Centroid is the minimized distance from all objects in the cluster. K-means algorithm minimizes this distance. Let the input image be considered as (x×y) resolution divided into k clusters.

The algorithm is as follows

1. Iniate the number of cluster k and its centre values.
2. Using the formula, calculate Euclidean distance d,

$$d = \left| p(x,y) - ck \right| \quad \ldots(1)$$

where p(x,y) is the input pixel and ck is the centre of cluster.

3. Assign nearest centre value based of d.
4. Recalculate new position of the centre using the assigned pixel value using the relation given below

$$ck = y \in ck \; x \in ck \; p(x,y) \quad \ldots(2)$$

5. Iteration continues until the process satisfies the tolerant error value and reshape the cluster pixel into the image.

The output of the algorithm depends on the initial value of the centroid. If it is not chosen Properly, we won't achieve the desired output.

3.5 FEATURE EXTRACTION USING GLCM

In case of texture analysis, the features are calculated based on intensities of the image at positions relative to each other [14] using Gray Level Co-occurrence Matrix (GLCM). It consists of matrix build using grey levels of the image. The element P (i, j | Δx, Δy) is the relative frequency of two pixel in the same neighborhood separated by distance i(Δx, Δy), with different intensity 'i' and 'j'. The Elements P (i, j | d, ɵ) is the probability of changes between grey level 'i' and 'j' at distance d and angle (ɵ). These elements help in choosing the centroid value. The algorithm depends on the size of the sample value and care should be taken that gray levels are reduced.

3.6 CONVOLUTION NEURAL NETWORK

Convolution neural networks [15-16] are designed to process 2-D image. During the training session, we categorize the test images under diseases. Convolution Neural Network helps in assigning the images to correct category. In the training process the characteristic properties of the image used to form a class of unique disease. The decision is based on number of layers which are interconnected to form a network.

We use feed-forward neural network FFNN where the layers have a input, output and hidden layer. The error is reduced using back propagation method, where the errors are adjusted from output layer to the input layer. Since we have knowledge of the input and the output, this method is chosen and exploited to train using dataset. CNNs use local correlation to find patterns between adjacent layers. The layers are connected to local subset from the previous layers. The elements of GLCM are used by the layers of CNN to reduce the learning parameters. The first layer of CNN contains convolution mask, bias term and functional expression and produces the output.

4. Result

The image is feed into the system. The images are resized to 32X32, input is given as image itself. The images are converted to gray scale. The larger images are resized using pyramid reduction.

The images are fed to train the system to reduce the error. The diseases are detected with an accuracy of 97.5% using the CNN algorithm and performance is calculated using Neural Tool using the trained images.

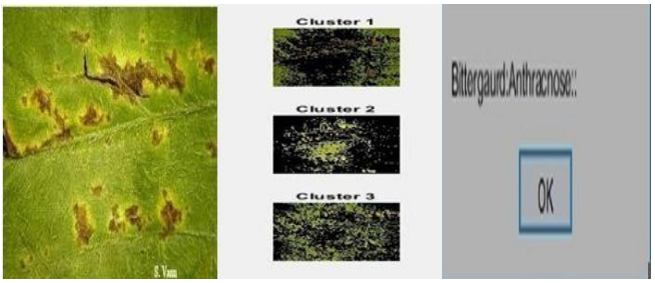

Fig-2 Output for Bitter Guard Suffering from Anthracnose

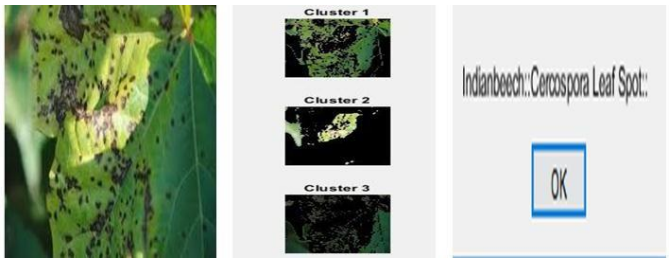

Fig-3 Output For Indian Beech Cercospora Leaf Spot

Fig-4 Output for Alternaria When Altermata image is Fed

Fig-5 Output image When Bacterial Blight is Fed

It is estimated that CNN method seems to deliver an accuracy of 97.5% for all the trained disease. In the paper, a new segmentation and classification process is introduced. K-means clustering algorithm helps in segmentation and Convolution Neural Network takes care of classification.

The statistical parameters are used as features for classification. The work works well for various part of plants and can be used to identify the disease is Pomegranate leaf, Bitter guard, tobacco leaf. The work can be extended for fruits, stems and other parts of the plants with minor modifications.

References

1. Yogesh Dandawate, Radha Kokare," An automated approach for classification of plant disease towards development of futuristic Decision Support System in Indian Perspective," vol. A247, pp. 529–551, 2015.
2. Arti N. Rathod, Bhavesh A. Tanawala, Vatsal H. Shah,—"Leaf Disease Detection Using Image Processing And Neural Network",IJAERD,2014,1(6).
3. P.Revathi, M.Hemalatha, —"Classification of Cotton Leaf Spot Diseases Using Image Processing Edge Detection Techniques", ISBN, 2012, 169-173, IEEE.
4. Chanchal Srivastava, Saurabh Kumar Mishra, Pallavi Asthana, G. R. Mishra, O.P. Singh —" Performance Comparison of Various Filters and Wavelet Transform for Image De-Noising",IOSRJCE,2013,10(1),55-63.
5. Hrushikesh Dattaray Marathe, Prerna Namdeorao Kothe, —"Leaf Disease Detection Using Image Processing Techniques", IJERT, 2013.
6. H. Al-Hiary, S. Bani-Ahmed,M. Reyalat, M.Braik and Z. AL Rahamneh,—"Fast and Accurate Detection and Classification of Plant Diseases",IJCA,2011,17(1).

7. Jayamala K. Patil, Raj Kumar, ―”Advances In Image Processing For Detection of Plant Diseases”, JABAR, 2011, 2(2), 135-141.
8. S. Arivazhagan, R. Newlin Shebiah, S. Ananthi, S. Vishnu Varthini,―”Detection of Unhealthy region of Plant Leaves and Classification of Plant Leaf Diseases using Texture Features”,CIGR,2013,15(1),211-217.
9. PawanP.Warne, Dr.S.R. Ganorkar“Detection of Diseases on Cotton Leaves Using K-Mean Clustering Method”, International Research Journal of Engineering and Technology (IRJET) Volume: 02 Issue: 04 | July2015, 425-431.
10. Daisy Shergill, AkashdeepRana, Harsimran Singh “Extraction of rice disease using image processing”,International Journal Of Engineering Sciences & Research technology,June, 2015,135-143.
11. K. Vengatesan, Rohit Ravindra Nikam, S. Yuvaraj, Ankoshe Malakappa Shankar, Punjabi Shivkumar Tanesh and Abhishek Kumar"A Random Forest-based Classification Method for Prediction of Car Price”, International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 03, 2020.
12. Dr. Mohammed Aref, Dr. Sameen Ahmed Khan, Dr. Sayyad Samee, K. Vengatesan and Abhishek Kumar,"Early Detection of Outbreaks by Monitoring the Over-the-Counter Pharmacy Sales"International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 04, 2020.
13. V.D. Ambeth Kumar, ∗, S. Malathia, R. Venkatesan, K. Ramalakshmi, K. Vengatesan, Weiping Ding and Abhishek Kumar,"Exploration of an innovative geometric parameter based on performance enhancement for foot print recognition”, Journal of Intelligent & Fuzzy Systems. Scopus (Saravana Kumar, E., Vengatesan, K. Trust based resource selection with optimization technique. Cluster Comput 22, 207–213 (2019)
14. Prabu, S., V. Balamurugan, and K. Vengatesan. "Design of cognitive image filters for suppression of noise level in medical images." Measurement 141 (2019): 296-301.
15. alvika Ranjan1, Manasi Rajiv Weginwar, NehaJoshi,Prof.A.B. Ingole, detection and classification of leaf disease using artificial neural network,International Journal of Technical Research and Applications e-ISSN: 2320-8163, Volume 3, Issue 3 (May-June 2015), PP.
16. RenukaRajendraKajale. Detection & recognition of plant leaf diseases using image processing and android o.s “International Journal of Engineering Research and General Science Volume 3, Issue 2, Part 2, March-April, 2015. ISSN 2091-2730.
17. NiketAmoda, Bharat Jadhav, SmeetaNaikwadi,” Detection and Classification of Plant Diseases By Image Processing”, International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 2, April 2014.
18. SmitaNaikwadi, NiketAmoda,” Advances in Image Processing for Detection of Plant Diseases,” International Journal of Application or Innovation in Engineering & Management (IJAIEM) Volume 2, Issue 11, November 2013. 168-175
19. Lidi Wang, Tao Yang, Youwen Tian “Plant Disease Leaf Image Segmentation Method Based on Color Features”, the International Federation for Information Processing, Volume 258, 2008, pp 713-717.
20. A. Iyswariya, Sarvepalli Chandrasekhar, Yaddala Chandrasekhar Reddy, Vunnam Sai Tathwik, V. Praveen Kumar “Identification And Detection of Plant Diseases by Convolutional Neural Network.