International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

Lemon Leaf Disease Detection and Classification using SVM and CNN

Balambigai Subramanian, Jayashree.S,Kiruthika.S, Miruthula.S

Abstract*: India is a developing country and agriculture has always played a major role in bolstering the country's economic growth. Due to various factors like industrialization, mechanization and globalization, the green fields are facing complications. So, identifying the plant disease incorrectly will lead to a huge loss of both quantity and quality of the product and it will also incur loss in time and money. Hence, identifying the condition of the plant plays a major role for successful cultivation. Now a day's image processing technique is being employed as a focal technique for diagnosing the various features of the crop. The image processing techniques can be used for identification of the plant disease and hence classify the plant disease. Generally, the symptoms of the disease are observed on leaves, stems, flowers etc. Here, the leaves of the affected plant are used for the identification and classification of the disease. Leaf image is captured using a smart phone as the first step and then they are processed to determine the condition of the plant. Identification of plant disease follows the steps like loading the image of the plant leaf, histogram equalization for enhancing contrast of the image, segmentation process by using Lab color space model, extracting features of the segmented image using GLCM (Grey Level Co-occurrence Matrix) and finally classification of leaf disease by using MCSVM (Multi Class Support Vector Machine).This procedure obtained an accuracy percentage of 83.6%.Also, it takes long training time for large datasets. To improve the accuracy of the detection and the classification of the plants, Convolutional Neural Network (CNN) is used. The main advantage of CNN is that it automatically detects the main features of the input without any supervision of human. In CNN identification of disease follow the steps like loading the image as the input image, convolution of the feature map and finally max pooling the layers to calculate the features of the image in detail. The plant diseases are classified with an accuracy of 93.8 %.

Keywords: Classification, CNN, Lemon leaf, SVM

Manuscript published on November 30, 2019.
* Correspondence Author
Balambigai Subramanian, Associate Professor, Department of Electronics and Communication Engineering, Kongu Engineering College
Jayashree.S, Bachelor of Engineering, Electronics and Communication Engineering. Kongu Engineering College, Perundurai
Kiruthika.S, Bachelor of Engineering, Electronics and Communication Engineering, Kongu Engineering College, Perundurai.
Miruthula.S, Bachelor of Engineering, Electronics and Communication Engineering, Kongu Engineering College, Perundurai

I. INTRODUCTION

Plants are the multi-cellular photosynthetic eukaryotes of the kingdom Plantae. Impairment in the normal state of a plant interrupts and modifies its pivotal functions such as respiration and photosynthesis. Diseases in plants affect them by interfering with several processes such as absorbance and translocation of water, minerals and nutrients. Flower and fruit development, plant growth, cell division and cell enlargement are some of the processes that are affected because of the plant diseases. All plant species irrespective of the type of their type of sowing and rearing are subjected to various kinds of diseases. Plant diseases occur based on the presence of various pathogens and environmental conditions. Plant diseases are usually caused by different types of fungi, bacteria, phytoplasma, viruses, viroids, nematodes and other pathogens. Certain plant varieties show resistance towards these pathogens while some are subjected to outbreaks and are hence affected by the pathogens. Plant diseases are also the effect of habitat loss poor land management. Plant diseases can annihilate natural ecosystems and the biological community. Thereby it aggravates environmental problems. Identifying the cause, symptoms and knowing when and how to effectively control those diseases is a great challenge underway. The austerity of diseases caused by these pathogens varies from light and mild symptoms to dwindling of the infected plants, depending on the aggressiveness of the pathogen, resistance of the whole plant, environmental conditions, duration of infection and other such factors. Plant disease symptoms vary with the infecting pathogen and the part on which it is infected. In tropical countries like India and Malaysia where their environmental conditions are particularly favorable, incomes of the population are low, knowledge and investments in crop health management are minimal then the crop losses tend to be greatest there. Crop losses can mean that the communities become more dependent on imported foods and processed foods, often replacing a proper balanced and a healthy diet with foods that create various health issues. Agriculture also provides employment opportunities for the country's rural population on a large scale and hence it becomes an important source of livelihood.The economy of many developing countries like India, Afghanistan and Brazil rely greatly on the agriculture of that country. The quantity, quality and the yield of agricultural product are reduced as a major result of diseases. When the plant diseases are caused by micro-organism like fungi, the prediction of its lifecycle is not easy and not always accurate. Some plants do not visibly show any changes throughout the early stage of the disease.

The prediction of plant disease by eyes is employed whereas the results are subjective and the extent of the infection is not exactly measured. Plant disease is one among the necessary issues that causes important reduction in the quality and also the amount of plant production.Symptoms of the plant diseases may include a variation in the color, shape and function of the plant as it responds to the infestation of the pathogen. In olden days such identifications were done manually by people who are experienced but due to many environmental changes, such predictions have become difficult. Now a day's image processing technique has become a key technique for performing diagnosis of the various features of the crop. The image processing techniques can be used for identification of the plant disease. Generally, the symptoms of disease are observed on leaves, stems, and flowers as the leaf spots, leaf blights, root rots, fruit rots, wilt and dieback. Here, the leaves of the affected plant are used for the identification of the disease. Plant pathology is the study of plant diseases. It includes identification of the pathogen based on the symptoms, disease cycles and also the economic impact of the plant diseases. In the existing methodology, the diseases are detected and classified using the Multiclass SVM. This project presents the detection and classification of diseases which are done by using the CNN (Convolution Neural Network) technique as the proposed system. Firstly the sample image of the leaf is given in as the input. Then, their color channels are separated from the input image. This is followed by the masking of the original image from the green pixels. The remaining area that is infected is calculated by removing the green area from the original image. Finally these features are given to the CNN to classify the leaves based on their disease.

II. EXISTING METHOD

Multi-class SVMs (MCSVM) are usually implemented by combining several binary SVMs. The objective of this work is to show a comparison of different constructing methods for Multi-class SVM, such as One-Against-One and One-Against-All and finally comparing the classifiers accuracy of Multi-class SVM classifier. The following are its steps.

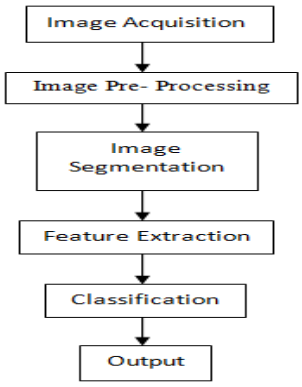

Fig. 1 Process Flow Diagram

2.1 Image Acquisition

Image acquisition is used to process the input images or the interior of an object and convert it into a digital image. It is the first step in the work flow, because processing is not possible without an image. The image acquired is completely unprocessed. Acquisition of the image is essential as the accuracy of the system depends on the samples of the image used for training the model. The quality of image samples depends on camera type and its orientation. Digital cameras are used with optical axis perpendicular to the leaf plane, but a few have used specialized techniques. Android mobile is also used to capture a leaf at some fixed distance.

2.2 Image Pre-Processing

The aim of pre-processing is an improvement of the image data that suppresses unwanted distortions or enhances some image features important for further processing. Here, the median filter is used for pre-processing of the images. Median filter, a non-linear filtering technique, uses a window that moves over a signal and at each point. The median value of the data within the window is taken as the output. The impulse response of the median filter is zero and thus makes its use attractive to suppress impulsive noise. During pre-processing, distortion removal improves images which ease further processing. Popular pre-processing techniques include color space conversion, cropping, smoothing, and enhancement. Depending upon image quality the functionality of this module varies.

After color space transformation, filters are applied for desired enhancements, like better contrast and brightness. Noise occurrence is general and so such systems use median filters. Technique like histogram equalization is also used for filtering and controlling varying lighting conditions. Cropping is essential if images are captured in an uncontrolled environment with complex backgrounds. It can be done either manually or automatically using functions.

2.3 Image Segmentation

Segmentation divides the image into regions with clear correlation along with areas of interest. The features of a perfectly segmented image will help in an easy identification of the infected samples as well as the healthy samples.

Lab color space is a three-axis system which has its dimensions as ***L*** for lightness, ***a*** and ***b*** as dimensions for color. It is a representation of color used for the process of segmentation .The color of the input image is converted from rgb to the lab color space. Dimensions will be broken down. The process of reshaping is done based on the colors. The colors are clustered and border pixels are fixed properly in a matrix. Modify the segmented function so that its minima appear at the foreground and background. It also describes color in relation to the lightness component (L) and the two chromatic components which are the red and the green components (a) and yellow and the blue components (b).

2.4 Feature Extraction

Collection of features by analyzing its texture is called as the feature extraction. The images are interpreted as color, shape and texture. Color is observed as histograms and moments. Properties like correlation, contrast, mean, entropy, homogeneity and variance can be attached to the texture.

Classical gray level co-occurrence matrix (GLCM) and its variants in space are used for computing the parameters of texture like entropy and moment of inertia of an area that is infected. The features of shape are computed along with mean, median and standard deviation.

III. STEPS IN EXISTING MECHANISM

Steps to train the images:

Step 1: Start with images of which their classes are known prior.
Step 2: Find the feature set for each of the image and label them.
Step 3: Take the successive image as an input and identify features of the one as a new input.
Step 4: Device the binary SVM to multi class SVM method.
Step 5: Train the SVM using any function.
Step 6: Find the class of the input image.
Step 7: Depending on the result, a label is given to the next image. Add these features to the database.
Step 8: Steps 3 to 7 are performed for all images that will be used in database.
Step 9: The outcome species is the class of the input image.
Step 10: To find the accuracy of the system or the SVM, in this case, random set of inputs are chosen for training and testing from the database. Two different sets for train and test are generated. The steps for training and testing are same, however, followed by the test is performed.

IV.PROPOSED METHODOLOGY

A Convolutional Neural Network (CNN) is a machine learning method from deep learning. CNNs are trained using large collections of images. From these large collections, CNNs can learn many of the feature representations for the whole of the collection of images. As an easier way to perform the classification without shedding time and effort, they are trained prior as an extractor of the features. The following are the steps involved:

Step 1: Start with the input image
Step 2: Feature map is created by applying four filters (convolution, mean and median, average).
Step 3: Apply a ReLU function for its non-linearity to increase.
Step 4: Apply pooling layer to each of the feature map.
Step 5: Flatten the pooled images into a single long vector.
Step 6: Gets the vector into a fully connected layer.
Step 7: Processes the features and the final fully connected layer provides the voting of the classes we require.
Step8: Trains for many epochs through the forward propagation and back propagation and this will repeat until a defined neural network with the trained features are obtained.

V. DESIGN OBSERVATIONS

TABLE I Confusion matrix for Multiclass SVM

Total images taken	Anthracnose	Citrus canker	Greasy spot	Healthy leaf
Anthracnose-7	6	1	0	0
Citrus canker-19	1	18	0	0
Greasy spot – 7	1	1	5	0
Healthy leaf – 16	1	5	0	10

The accuracy obtained for Multiclass SVM is 83.6%.

TABLE II Confusion matrix for CNN

Total images taken	Anthracnose	Citrus canker	Greasy spot	Healthy leaf
Anthracnose-7	7	0	0	0
Citrus canker-19	2	17	0	0
Greasy spot – 7	0	0	7	0
Healthy leaf – 16	0	1	0	15

The accuracy obtained for CNN is 93.8%.

VI. CONCLUSION

In this paper, the applications of Lab color space model and Convolutional Neural Networks (CNNs) has been used for the process of segmentation and classification of the diseases that affect the lemon plant's leaves. Recognizing the disease and classifying it is the main purpose of the introduced approach. The Multiclass SVM algorithm and the CNN Algorithm were trained and tested on a healthy lemon leaf and on three diseases which influences the lemon leaves. The model was able to detect the presence of leaf and distinguish between the healthy leaves and the leaves with different diseases, which can be visually diagnosed.

The diseases are Anthracnose, Citrus Canker and Greasy Spot. The experimental results indicate that both the approaches can support the detection and classification of the diseases on the leaf. The level of accuracy in classification of the leaves by Multiclass SVM was 83.6%, whereas it was 93.8% when the leaves were classified using the CNN algorithm. Hence the accuracy in classification of the leaves using CNN algorithm showed better results.

REFERENCES

1. Satya Priya N., Nivetha E &RashmitaKhilar "Efficient knowledge based system to detect disease in lemon leaf", Imperial Journal of Interdisciplinary Research, May, 2016.Vol-2, Issue-5, 2016 ISSN: 2454-1362.
2. Dheeb Al Bashish, Malik Braik, and Sulieman Bani (2016)-"A Framework for Detection and Classification of Plant Leaf and Stem Diseases".
3. Sukhir Kaur, Shreelekha Pandey, Shivani Goel, "Plants Disease Identification and Classification through Leaf Images", Springer, April 2019, Volume 26, Issue 2, pp 507.

4. Jahan Mohan K, Balasubramanian M, Palanivel S (2016), " Detection and recognition of diseases from Paddy Plant leaf images". IntJComput Appl 141(12):34–41.
5. Arivazhagan S., Newlin Shebiah R., Ananthi S., Vishnu Varthini, "Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features", CIGR, vol. 15(1), pp. 211-217, March 2013.
6. Bauer, S. D., F. Korc, W. Forstner. 2011. "The potential of automatic methods of classification to identify leaf diseases from multispectral images". Precision Agriculture, vol.12: 361-377.
7. Deya AK, Sharmaa M, Meshram MR (2016), " Image processing based leaf rot disease, detection of betel vine (Piper BetleL.)". In: International conference on computational modeling and security. Procedia Comput Sci 85:748–754.
8. Gui J., Hao L., Zhang Q., Bao X (2015), " A new method for soybean leaf disease detection based on modified salient regions". International Journal of Multimedia and Ubiquitous Engineering 10(6):45–52.
9. Jagadeesh Pujari D., Rajesh Yakkundimath , Abdulmunaf S. Byadgi , "Image Processing Based Detection of Fungal Diseases in Plants",Elsevier 2015.Procedia Computer Science 46 (2015) 1802 – 1808.
10. Jayme Garcia ArnalBarbedo "Digital image processing techniques for detecting, quantifying and classification plant diseases", Springerplus, 2013, 2:660.
11. Kavya RM., Gowda A., Bharathi PT., Virupakshaiah HK (2016), " Image processing techniques based plant disease detection". Int J Adv Found Res Computer: Pragmatic Rev 3(8):1–8.
12. Zhang S.,Huang W., Zhang C.(2018).Three-channel convolutional neural networks for vegetable leaf disease recognition.Cogn.Syst.Res.53,31-41.
13. Tejal Deshpande1, Sharmila Sengupta2, Raghuvanshi, "Grading & Identification Of Disease in Pomegranate Leaf and Fruit",International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 2014, 4638-4645 ISSN-0975-9646.
14. Revathi P., Hemalatha M., "Advance Computing Enrichment Evaluation of Cotton Leaf Spot Disease Detection Using Image Edge detection", ICCCNT'1226t _28th July 2012.
15. Ramakrishnan, M., and A. Nisha. "Groundnut leaf disease detection and classification by using back probagation algorithm." In Communications and Signal Processing (ICCSP), 2015 International Conference on, ISNB no. 978-1-4799-8081-9, page no. 0964-0968. IEEE, 2015.
16. Mokhtar U, Ali MA, Hassenian AE, Hefny H (2015), " Tomato leafs diseases detection approach based on support vector machines". In: Eleventh International IEEE, Computer Engineering Conference, pp.246–250
17. Lalitha K.,Muthulakshmi K.,Vinodhini S., ,"Proficient acquaintance based system for citrus leaf recognition and categorization",(IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (3) , 2015, 2519-2524.

AUTHORS PROFILE

Dr.S.Balambigai received her B.E. degree in ECE from Bharathiar University with distinction, M.E. (Applied Electronics) and Ph.D from Anna University, Chennai. She has around 35 publications in reputed national and international and National journals/conferences and one patent to her credit. She has a total of 16 years experience and is currently working as an Associate Professor in the Department of Electronics and Communication Engineering in Kongu Engineering College. Developments in biomedical engineering and computer networks keep her fascinated to carry on further research in these domains.

Ms. Jayashree S. is currently pursuing her Bachelor of Engineering in Electronics and Communication Engineering at Kongu Engineering College, Perundurai. Her areas of interests include Image Processing and Digital Electronics.

Ms. Kiruthika S. is currently pursuing her Bachelor of Engineering in Electronics and Communication Engineering at Kongu Engineering College, Perundurai. Her areas of interests include Image processing and VLSI.

Ms. Miruthula S. is currently pursuing her Bachelor of Engineering in Electronics and Communication Engineering at Kongu Engineering College, Perundurai. Her areas of interests include Image Processing techniques and Medical Electronics.