Autonomous Recognition of Leaf Medicinal uses using CNN adapted by SVM (AI)

Dr.S.Suthir[1] , S.Naveen Raj[2] , V.Kumaran[2], K.Kaushik[2]

Assistant Professor, Dept. of CSE, Panimalar Engineering College, Chennai, India[1]

UG Students, Dept. of CSE, Panimalar Engineering College, Chennai, India[2]

ABSTRACT: A fully automated method for the recognition of medicinal plants using computer vision and machine learning techniques has been presented. Leaves from 24 different medicinal plant species were collected and photographed using a smart phone in a laboratory setting. A large number of features were extracted from each leaf such as its length, width, perimeter, and area, number of vertices, colour, perimeter and area of hull.
Several derived features were then computed from these attributes. The best results were obtained from a SVM classifier using a 10-fold cross-validation technique. With an accuracy of90.1%, SVM classifier performed better than other machine learning approaches such as the k-nearest neighbour, Naïve Bayes, KNN and neural networks.

KEYWORDS: Machine Learning , SVM .

I. INTRODUCTION

The agricultural land mass is more than just being a feeding sourcing in today's world. Indian economy is highly dependent of agricultural productivity. Therefore in field of agriculture, detection of disease in plants plays an important role. To detect a plant disease in very initial stage, use of automatic disease detection technique is beneficial. For instance a disease named little leaf disease is a hazardous disease found in pine trees in United States. The affected tree has a stunted growth and dies within 6 years. Its impact is found in Alabama, Georgia parts of Southern US. In such scenarios early detection could have been fruitful.

The existing method for plant disease detection is simply naked eye observation by experts through which identification and detection of plant diseases is done. For doing so, a large team of experts as well as continuous monitoring of plant is required, which costs very high when we do with large farms. At the same time, in some countries, farmers do not have proper facilities or even idea that they can contact to experts. Due to which consulting experts even cost high as well as time consuming too. In such conditions, the suggested technique proves to be beneficial in monitoring large fields of crops. Automatic detection of the diseases by just seeing the symptoms on the plant leaves makes it easier as well as cheaper. This also supports machine vision to provide image based automatic process control, inspection, and robot guidance .

Plant disease identification by visual way is more laborious task and at the same time, less accurate and can be done only in limited areas. Whereas if automatic detection technique is used it will take less efforts, less time and become more accurate. In plants, some general diseases seen are brown and yellow spots, early and late scorch, and others are fungal, viral and bacterial diseases. Image processing is used for measuring affected area of disease and to determine the difference in the color of the affected area.

Image segmentation is the process of separating or grouping an image into different parts. There are currently many different ways of performing image segmentation, ranging from the simple thresholding method to advanced color image segmentation methods. These parts normally correspond to something that humans can easily separate and view as individual objects.

Computers have no means of intelligently recognizing objects, and so many different methods have been developed in order to segment images. The segmentation process is based on various features found in the image. This might be color information, boundaries or segment of an image . We use Genetic algorithm for color image segmentation.

II. RELATED WORK

The feature extraction is done in RGB, HSV, YIQ and Dithered Images. The feature extraction from RGB image is added in the suggested system. A new automatic method for disease symptom segmentation in digital photographs of plant leaves. The diseases of different plant species has mentioned. Classification is done for few of the disease names in this system.

The disease recognition for the leaf image is performed in this work. Study and analysis of cotton leaf disease detection using image processing work is carried on. The k means Clustering algorithm is used for segmentation. The k-means concept is added to the proposed system which will divide the leaf into different clusters.

The survey of disease identification on cotton leaf is done. Comparison of different detection technique of leaf disease detection is mentioned. SVM and k-means clustering has used in this system. An identification of variety of leaf diseases using various data mining techniques is the potential research area. The diseases of different plant species has mentioned. Classification is done for few of the disease names in this system. The concept SVM for classification is used in this system

III. PROPOSED ALGORITHM

Monitoring and identifying medicinal uses and disease of a leaf as become one of the most essential activities in agriculture today. The quality of leaf is adversely affected due to various form of pollutions and mainly due to the usage of dangerous and harmful pesticides with increase in usage of harmful pesticides leaves lose their medicinal uses which results new diseases in leaves.

Here in the proposed system, which predicts the medicinal uses using leaf and also whether it is affected or not. Based on the datasets collected the images are segmented and trained .The data are tested by identifying the leaves CNN algorithm helps in classifying the data .Finally based on the prediction it shows the medicinal uses of a given leaf.

The size of each image was 256*256 pixels. Proposed an approach based on fractal dimension features based on leaf shape and vein patterns for the recognition and classification plant leaves. Using a volumetric fractal dimension approach to generate a texture signature for a leaf and the GLCM (Gray level co occurrence matrix) algorithm.

IV. SYSTEM ARCHITECTURE

System architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system Fig 3.1 represents the system architecture of the Proposed system.

The working process is as follows

- Leaf Identification Process
- Image segmentation
- Outlier detection process
- Comparing Algorithm with prediction in the form of best accuracy result.
- GUI based user interface

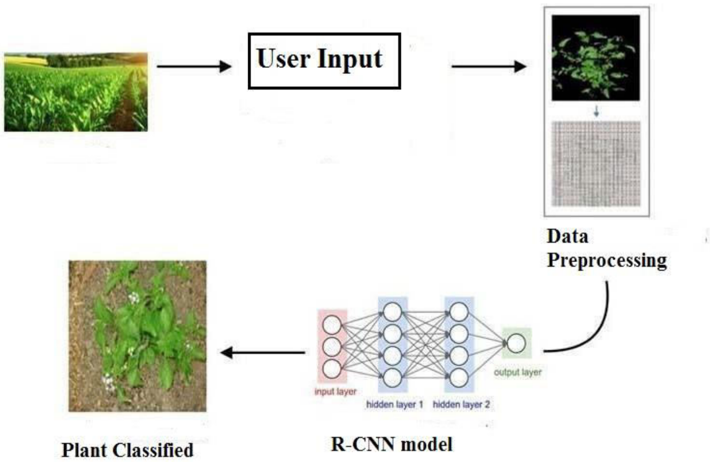

Fig 3.1 System Architecture

V. IMPLEMENTATION

This process uses the CNN Model, Convolutional neural networks have a different architecture than regular neural networks. Regular neural networks transform an input by putting it through a series of hidden layers. Every layer is made up of a set of neurons, where each layer is fully connected to all neurons in the layer before. Finally, there is a last fully-connected layer, the output layer that represent the predictions.

Fig 4.1 represents the overview of the entire system Here is the overview of what we are going to cover:

Installing the data set using python platform

1. Loading the data set
2. Summarizing the data set
3. Visualizing the data set
4. Image segmentation
5. Evaluating results

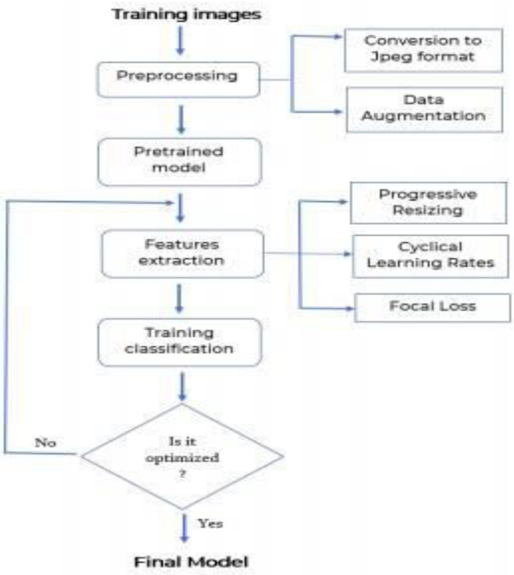

Fig 4.1 Flowchart

The stages carried out in medicinal plant recognition are as follows

Data acquisition

Data acquisition is data retrieval process of image leaf image by using a camera phone with resolution 4.1 mega pixel with distance of picture as high as 30cm.

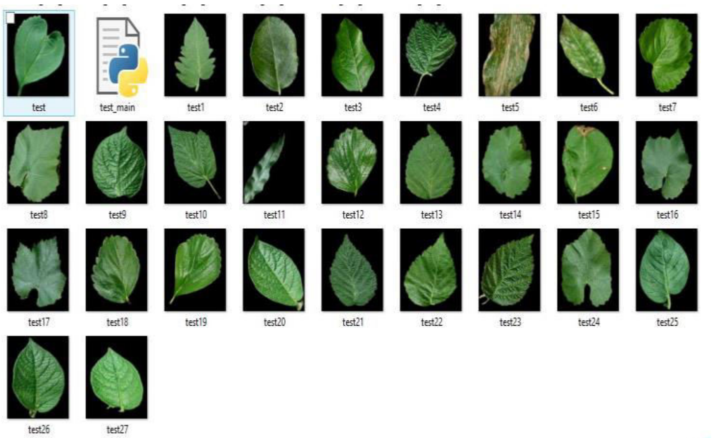

Fig 4.1 Features of various leaves collected

Automatic Preprocessing

One drawback of taking pictures using a camera, instead of using a scanner, is the presence of shadows on the image. If the shadow is not removed, this will affect all measurement. Thus, to remove the shadow, the image must first be converted to the HSV format and then split into its different colour channels. Only the second channel (saturation) is kept.

This has the effect of removing the shadow from the image. To reduce noise in the image, a median blur filter with a window size of 25 is applied to the resulting image. The next step is to perform a thresholding operation which will convert the image into a binary image with only two values: black and white pixels. This is achieved using the Otsu thresholding method. An opening operation is then performed on the images. This is an erosion operation followed by a dilation. Erosion has the effect of reducing the size of foreground (white) pixels while dilation enlarges them.

This operation is important in order to clear the image from many small noisy pixels, which are the artefacts of the thresholding operation.

Preprocessing process start by turn color input image into grayscale. Input color image will be change to grayscale by using formula :

$G(x,y)= (R+G+B)/3$

Grayscale value for each image pixel is obtained by calculating the average value from the color values of R (red), G (green), and B (blue) on each pixel. Grayscale image will be filter to eliminate obtained noise at the intake of picture by using Median Filter.

Next process is detecting edge by using Prewitt method to get leaf bone which later will be used on process feature extraction. This method takes the principle of Laplacian function known as a function for generating High Pass Filter. Kernel filters used in Prewitt method is which perform a search in horizontal (H) and vertical (V).

Feature Extraction

Derived features Using the base features which are extracted directly from the image, a number of derived features are calculated

VI. EXPERIMENTAL RESULTS

Besides the overall accuracy, the performance of the automated system was also assessed on a class-wise basis. Recall is the proportion of leaves, for each class, that was correctly picked out from the entire set. Fig 5.1 represents the identity of random plants and the true positive false positive true negative and false negative rates are calculated and evaluated using the confusion matrix in Fig 5.2.

Fig 5.1 identification of medicinal plants

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	<-- classified as
25	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	2	0	a = Antidesma
0	28	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	b = Avocado
0	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	c = Ayapana
0	0	0	29	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	d = Balloon Plant
2	0	0	0	24	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	3	0	0	0	e = Bigaignon Rouge
0	0	0	0	0	28	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	f = Bitter Gourd
0	0	1	0	0	0	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	g = Bois Carrote
0	0	0	0	0	0	0	27	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	h = Bois Cerf
0	0	0	0	0	0	0	0	26	0	0	0	0	3	0	1	0	0	0	0	0	0	0	0	i = Bois de rat
0	0	0	0	0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	j = Bramble
0	0	0	0	0	0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	k = Chinese Okra
0	0	0	0	0	6	0	0	0	1	0	21	2	0	0	0	0	0	0	0	0	0	0	0	l = Coriander
0	0	0	0	0	0	0	0	0	0	0	0	26	1	0	0	1	0	0	0	0	0	1	1	m = Curry Leaf
0	0	0	0	0	0	0	0	0	0	0	0	0	27	0	0	1	0	0	0	0	2	0	0	n = Fandamane
0	0	0	0	0	0	2	0	0	0	0	0	0	0	23	0	0	0	0	0	0	2	3	0	o = Jackfruit
0	0	0	0	0	0	1	1	3	0	0	0	0	0	0	24	0	1	0	0	0	0	0	0	p = Mango
0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	28	0	0	0	1	0	0	0	q = Moringa
0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	1	0	27	0	0	0	0	0	0	r = Neem
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30	0	0	0	0	0	s = Orange Climber
0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	29	0	0	0	0	t = Parsley
0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	28	0	0	0	u = Peppermint
1	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	27	0	0	v = Pomegranate
0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	28	0	w = Strawberry Guava
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	29	x = Tulsi

Fig 5 .2 Confusion Matrix

VII. CONCLUSION AND FUTURE WORK

Identification of the medicinal plants through plant leaves image object identification is a stage of early identification of a plant species. In this paper, a contemporary approach to identify the medicinal leaves was done using the CNN model. The performance of each model was evaluated according to the discoloration of, or damage to, leaves. The recognition rate achieved was to a greater extent even when 30% of the leaf was damaged.

In future research we will attempt to recognize leaves attached to branches, in order to develop a visual system that can replicate the method used by humans to identify plant types.

.

REFERENCES

1.FAO. “How to feed the world in 2050”. http://www.fao.org/fileadmin/ templates/wsfs/docs/expert_paper/How_to_Feed_the_World_in_2050. pdf.

2. Univeristy of Pretoria. “Important weeds in maize”. https://www.up.ac. za/sahri/article/1810372/important-weeds-in-maize.

3. Yann LeCun, YoshuaBengio, and Geoffrey Hinton. “Deep learning”. nature, 521(7553):436, 2015.

4. Thomas M. Mitchell. “Machine Learning”. McGraw-Hill, Inc., New York, NY, USA, 1 edition, 1997.

5. Saad Abouzahir, Mohamed Sadik, and Essaid Sabir. “Enhanced approach for weeds species detection using machine vision”. In 2018 International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS), pages 1-6. IEEE, 2018.

6. Jialin Yu, Shaun M Sharpe, Arnold W Schumann, and Nathan S Boyd. “Deep learning for image-based weed detection in turfgrass”. European Journal of Agronomy, 104:78-84, 2019.

7. Philipp Lottes, Jens Behley, Andres Milioto, and CyrillStachniss. “Fully convolutional networks with sequential information for robust crop and weed detection in precision farming”. IEEE Robotics and Automation Letters, 3(4):2870-2877, 2018.

8. M Dian Bah, Adel Hafiane, and Rapha¨el Canals. “Deep learning with unsupervised data labeling for weeds detection on uav images”. arXiv preprint arXiv:1805.12395, 2018.

9. Wenhao Zhang, Mark F Hansen, Timothy N Volonakis, Melvyn Smith, Lyndon Smith, Jim Wilson, Graham Ralston, Laurence Broadbent, and Glynn Wright. “Broad-leaf weed detection in pasture”. In 2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC), pages 101-105. IEEE, 2018.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

||Volume 8, Issue 5, May 2020||

10 . Barbedo, Jayme Garcia Arnal, A new automatic method for disease symptom segmentation in digital photographs of plant leaves, European Journal of Plant Pathology, 2016, 1-16.

11. Bhong, Vijay S and Pawar B.V, Study and Analysis of Cotton Leaf Disease Detection Using Image Processing, International Journal of Advanced Research in Science, Engineering and Technology, 3 (2), 2016.

12. Lumb, Manisha, and Poonam Sethi, Texture Feature Extraction of RGB, HSV, YIQ and Dithered Images using GLCM, Wavelet Decomposition Techniques, International Journal of Computer Applications, 68 (11), 2013.

12. Reena Tijare, Pawan Khade, Rashmi Jain, The Survey of Disease Identification of Cotton Leaf, International Journal of Innovative Research in Computer and Communication Engineering, 2015.

13. Sasirekha N, Swetha N, An Identification of Variety of Leaf Diseases Using Various Data Mining Techniques, International Journal of Advanced Research in Computer and Communication Engineering, 4 (10), 2015.

14. Zhang S. W, Shang Y.J, and Wang L, Plant Disease Recognition Based on Plant Leaf Image, Journal of Animal and Plant Sciences, 25 (1), 2015, 42-45.