PLANT LEAF IMAGE RECOGNITION SYSTEM

[1]Gaurav Kulkarni, [2]Chandrashekhar Kumbhar

[1]Student, [2]Assistant Professor, Information Technology,

Ajeenkya DY Patil University, Pune

DOI: 10.52984/ijomrc1108

Abstract:

Plants play a vital role in our day-to-day life. Hence, a good understanding of plants is needed to help in identifying new or rare plant species. Such identification will in turn improve the drug industry, balance the ecosystem as well as the agricultural productivity and sustainability. We often come across various plants with different variety of leaves and flowers every single day. We try to recognize it, but we fail. So we need some system which can tell us about the leaf/flower instantly. So, to solve such problems, we introduce a plant recognition system (PRS) which tells you the details about a leaf by just uploading the image of the leaf. For this system, we use image processing and some identification techniques which can recognize the leaf by its structure, colour, shape etc and fetch the details about it and provide the details of it to the user. This paper gives a understanding about the different methods used under image processing and various methods and algorithm used to identify that leaf in a short and simple way. Object recognition and detection are techniques with similar end results and implementation approaches. Therefore, it requires heavy pre-processing and implements various processes to obtain the end results.

Keywords: Plant recognition system, PRS, Image processing, Object recognition, Computer vision.

INTRODUCTION:

It is important to know about the various species of plants which come across every day in our life. They play a vital role in our life and it's our responsibility to plant more trees and grow them. It is essential to know about the plants, its different species etc. So, therefore this paper presents a system which can identify the plants by just uploading its leaf images. When a user uploads an image of a leaf, the system tells the user about the plant. The main objectives of creating such a system are to identify the plants by uploading its image and getting to know about its different features, fast recognition of a plant. Also, such identification can help to improve the drug industry too. Also, it can perform a role in balancing the ecosystem as well as the agricultural productivity and sustainability. In addition to this, this paper can also be a good search for the researchers who want to work on identication systems using open CV, image processing etc.

Talking about the project implementation and different techniques used to construct the model, firstly, the language used for this project is python and implemented in Jupyter notebook. Jupyter notebooks are one the best platform for implementation of python projects. For working on any project like this, we need to collect the data and pre – process that data. Now, considering this project, after taking the leaves dataset, the images need to pre – process. Pre – processing further includes different steps such as conversion of RGB to Grayscale, smoothing of the image, closing of holes, boundary extraction using contours, etc. After pre-processing we need to move further towards feature extraction and model building. Also, after model building, model testing is one of the crucial parts of the project which needs to be implemented. So therefore, in short, this paper introduces to the different steps to construct such a plant recognition system and gives a brief idea about the system which can be further used.

LITERATURE SURVEY:

This paper explains about the different digital image processing techniques to detect, quantify and classify the plant diseases from digital images in visible spectrum. Only the methods that can the visible symptoms in were considered. Therefore the paper has the potential to be useful for the researchers working on vegetable pathology and pattern recognition [1].

In this paper, the author employ probabilistic neural network (PNN) with image and data processing techniques to implement general purpose automated leaf recognition for plant classification. 12 leaf features are extracted and orthogonal into 5 principal variables which consist the input vector of the PNN. The PNN is trained by 1800 leaves to classify 32 kinds of plants with accuracy greater than 90%. Compared with other approaches, our algorithm is an accurate artificial intelligence approach which is fast in execution and easy in implementation [2].

The identification of plant diseases is the key to preventing the losses in the yield and quantity of the agricultural product. The studies of the plant diseases mean the studies of visually observable patterns seen on the plant. Health monitoring and disease detection on plant is very critical for sustainable agriculture. This paper discussed the methods used for the detection of plant diseases using their leaves images. This paper also discussed some segmentation and feature extraction algorithm used in the plant disease detection [3].

Recognition of plants has become an active area of research as most of the plant species are at the risk of extinction. This paper uses an efficient machine learning approach for the classification purpose. This proposed approach consists of three phases such as pre-processing, feature extraction and classification. The pre-processing phase involves a typical image processing steps such as transforming to grey scale and boundary enhancement. The feature extraction phase derives the common DMF from five fundamental features. The main contribution of this approach is the Support Vector Machine (SVM) classification for efficient leaf recognition. 12 leaf features which are extracted and orthogonalized into 5 principal variables are given as input vector to the SVM. Classifier tested with Flavia dataset and a real dataset and compared with k-NN approach, the proposed approach produces very high accuracy and takes very less execution time [4].

Agricultural productivity is something on which economy highly depends. This is the one of the reasons that disease detection in plants plays an important role in agriculture field, as having disease in plants are quite natural. If proper care is not taken in this area then it causes serious effects on plants and due to which respective product quality, quantity or productivity is affected. Detection of plant disease through some automatic technique is beneficial as it reduces a large work of monitoring in big farms of crops, and at very early stage itself it detects the symptoms of diseases. This paper presents an algorithm for image segmentation technique which is used for automatic detection and classification of plant leaf diseases. It also covers survey on different diseases classification techniques that can be used for plant leaf disease detection [5].

Plant identification based on leaf is becoming one of the most interesting and a popular trend. Each leaf carries unique information that can be used in the identification of plants. In the identification of plants based on leaf, the leaf images needs to be pre-processed accordingly to extract the various critical features. In this paper, we present the identification of plants based on leaf features using Multiclass SVM (MSVM) as a classifier [6].

This paper proposes a novel methodology of characterizing and recognizing plant leaves using a combination of texture and shape features. Texture of the leaf is modelled using Gabor filter and grey level co-occurrence matrix (GLCM) while shape of the leaf is captured using a set of curve let transform coefficients together with invariant moments. Since these features are in general sensitive to the orientation and scaling of the leaf image, a pre-processing stage prior to feature extraction is applied to make corrections for varying translation, rotation and scaling factor [7].

In this paper, we propose and implement a leaf recognition system using the leaf vein and shape that can be used for plant classification. The proposed approach uses major main vein and frequency domain data by using Fast Fourier Transform (hereinafter, FFT) methods with distance between contour and centroid on the detected leaf image. Total 21 leaf features were extracted for the leaf recognition [8].

Plant recognition is an important and challenging task. Leaf recognition plays an important role in plant recognition and its key issue lies in whether selected features are stable and have good ability to discriminate different kinds of leaves. From the view of plant leaf morphology (such as shape, dent, margin, vein and so on), domain–related visual features of plant leaf are analysed and extracted first. On such a basis, an approach for recognizing plant leaf using artificial neural network is brought forward. The prototype system has been implemented. Experiment results prove the effectiveness and superiority of this method [9].

In this paper, RSC features have been proposed and extracted from leaf images and use SVM classifier to implement an automated leaf recognition system for plant leaf identication and classification. Automatic plant species identification and classification is helpful in biology, forest and agriculture to study and discover new species in plant in botanical gardens and is also used to recognize the medical plants to prepare herbal medicines [10].

DATA SET COLLECTION:

The dataset collected used for this project is of Flavia leaves dataset. The Flavia leaves dataset a standard leaves dataset which is available for everyone todownload.

label	Scientific Name	Common Name(s)	filename	URL
1	Phyllostachys edulis (Carr.) Houz.	pubescent bamboo	1001-1059	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=506646
2	Aesculus chinensis	Chinese horse chestnut	1060-1122	http://www.ars-grin.gov/cgi-bin/npgs/html/taxon.pl?1625
3	Berberis anhweiensis Ahrendt	Anhui Barberry	1552-1616	http://asaweb.huh.harvard.edu:8080/databases/specimens?id=277371
4	Cercis chinensis	Chinese redbud	1123-1194	http://www.ag.auburn.edu/hort/landscape/dbpages/306.html
5	Indigofera tinctoria L.	true indigo	1195-1267	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=26750
6	Acer Palmatum	Japanese maple	1268-1323	http://en.wikipedia.org/wiki/Acer_palmatum
7	Phoebe nanmu (Oliv.) Gamble	Nanmu	1324-1385	http://www.ars-grin.gov/cgi-bin/npgs/html/taxon.pl?28039
8	Kalopanax septemlobus (Thunb. ex A.Murr.) Koidz.	castor aralia	1386-1437	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=565257
9	Cinnamomum japonicum Sieb.	Chinese cinnamon	1497-1551	http://en.wikipedia.org/wiki/Cinnamomum_japonicum
10	Koelreuteria paniculata Laxm.	goldenrain tree	1438-1496	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=503286
11	Ilex macrocarpa Oliv.	Big-fruited Holly	2001-2050	http://asaweb.huh.harvard.edu:8080/databases/specimens?id=159529
12	Pittosporum tobira (Thunb.) Ait. f.	Japanese cheesewood	2051-2113	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=24067
14	Chimonanthus praecox L.	wintersweet	2114-2165	http://www.ars-grin.gov/cgi-bin/npgs/html/taxon.pl?10204
15	Cinnamomum camphora (L.) J. Presl	camphortree	2166-2230	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=18175
16	Viburnum awabuki K.Koch	Japan Arrowwood	2231-2290	http://www.ars-grin.gov/cgi-bin/npgs/html/taxon.pl?80375
17	Osmanthus fragrans Lour.	sweet osmanthus	2291-2346	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=505977
18	Cedrus deodara (Roxb.) G. Don	deodar	2347-2423	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=183408
19	Ginkgo biloba L.	ginkgo, maidenhair tree	2424-2485	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=183269
20	Lagerstroemia indica (L.) Pers.	Crape myrtle, Crepe myrtle	2486-2546	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=27110
21	Nerium oleander L.	oleander	2547-2612	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=30184
22	Podocarpus macrophyllus (Thunb.) Sweet	yew plum pine	2616-2675	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=183490
23	Prunus serrulata Lindl. var. lannesiana auct.	Japanese Flowering Cherry	3001-3055	http://www.ars-grin.gov/cgi-bin/npgs/html/taxon.pl?104754
24	Ligustrum lucidum Ait. f.	Glossy Privet	3056-3110	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=503450
25	Tonna sinensis M. Roem.	Chinese Toon	3111-3175	http://www.ars-grin.gov/cgi-bin/npgs/html/taxon.pl?36754
26	Prunus persica (L.) Batsch	peach	3176-3229	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=508766
27	Manglietia fordiana Oliv.	Ford Woodlotus	3230-3281	http://www.ars-grin.gov/cgi-bin/npgs/html/taxon.pl?23361
28	Acer buergerianum Miq.	trident maple	3282-3334	http://www.ars-grin.gov/cgi-bin/npgs/html/taxon.pl?1088
29	Mahonia bealei (Fortune) Carr.	Beale's barberry	3335-3389	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=18846
30	Magnolia grandiflora L.	southern magnolia	3390-3446	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=18074
31	Populus ×canadensis Moench	Canadian poplar	3447-3510	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=22457
32	Liriodendron chinense (Hemsl.) Sarg.	Chinese tulip tree	3511-3563	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=18086
33	Citrus reticulata Blanco	tangerine	3566-3621	http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=28888

Fig 1: View of Flavia Leaves dataset

PROJECT WORKFLOW:

To classify and recognize the image of leaf, there are certain steps to be followed in the project to train the system to recognize that particular image. In this project too, certain steps have been followed to train the system with recognizing power. Therefore, this project mainly follows three basic steps; Pre- processing, Feature extraction and the Model Building.

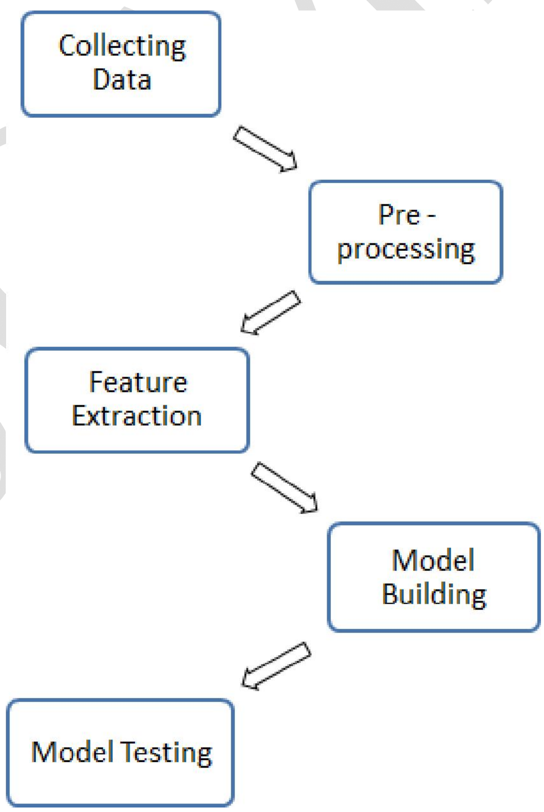

Fig 2: Flowchart of project workflow

PROPOSED SYSTEM:

For working on any project like this, we need to collect the data and pre – process that data. Now, considering this project, after taking the leaves dataset, the images need to pre – process. Pre – processing further includes different steps such as conversion of RGB to Grayscale, smoothing of the image, closing of holes, boundary extraction using contours, etc. After pre-processing we need to move further towards feature extraction and model building. Also, after model building, model testing is one of the crucial parts of the project which needs to be implemented. So therefore, in short, this paper introduces to the different steps to construct such a plant recognition system and gives a brief idea about the system which can be further used.

(i.) Pre – processing:

In this working of the system, the pre – processing includes main five steps: (a.) Converting the RGB

image to Grey scale:

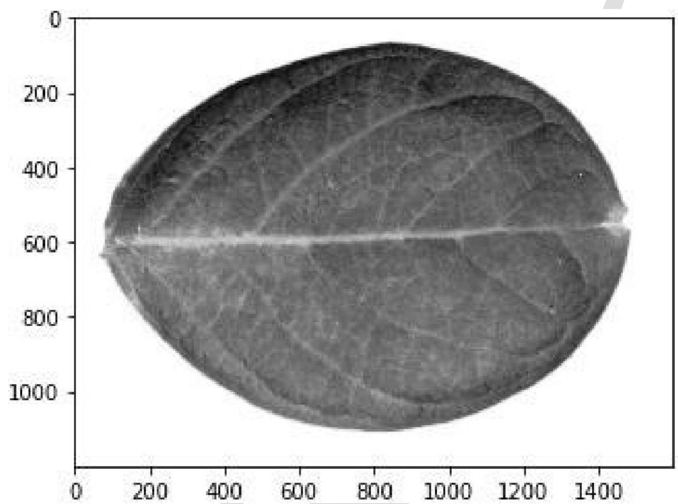

Fig 3: RGB to Grey Scale

It is necessary to convert the image from RGB to Grayscale because it is easy with the Grayscale images and perform further operations such as finding shape features, colour features, texture based features etc.

(b.) Smoothing image using Gaussian filter:

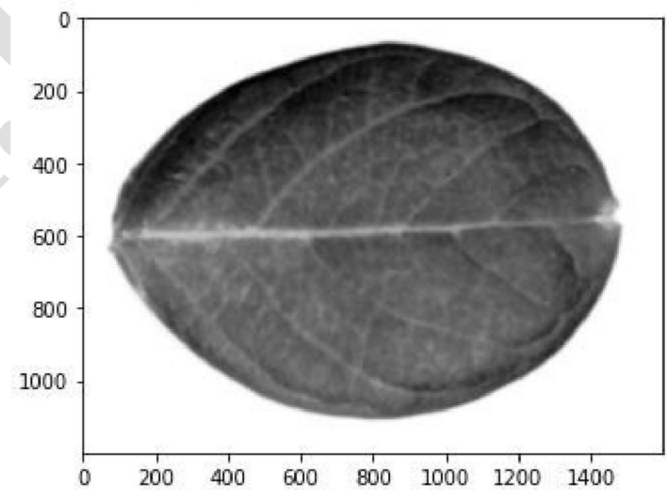

Fig 4: Smoothing image using Gaussian filter

To extract proper features from the image, we need to smooth the image and to reduce the noise from the image, we have used the gaussian filter technique. Gaussian filter basically blurs the image to reduce the noise. The gaussian filter will alone blur edges and reduces contrast.

(c.) Adaptive image thresholding using Otsu's thresholding method:

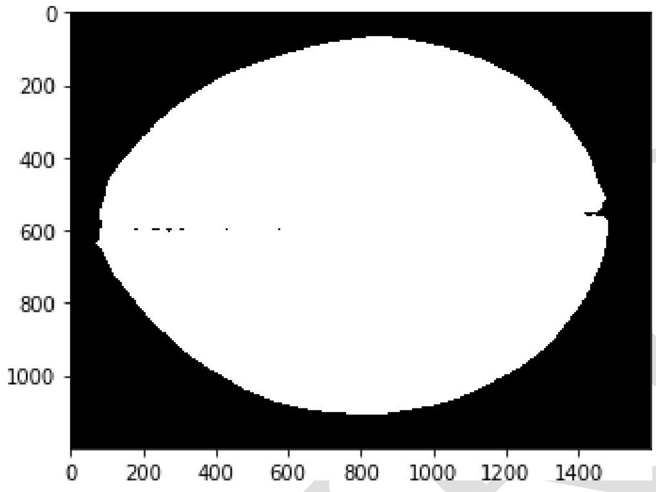

Fig 5: Adaptive thresholding using Otsu's Thresholding method

Adaptive thresholding typically takes a grey scale or colour image as input and, in the simplest implementation, outputs a binary image representing the segmentation. For each pixel in the image, a threshold has to be calculated. The threshold for each single pixel is found by interpolating the results of the sub images. Otsu's thresholding method involves iterating through all the possible threshold values and calculating a measure of spread for the pixel levels each side of the threshold, i.e. the pixels that either falls in foreground or background.

(d.) Closing of holes using morphological transformation:

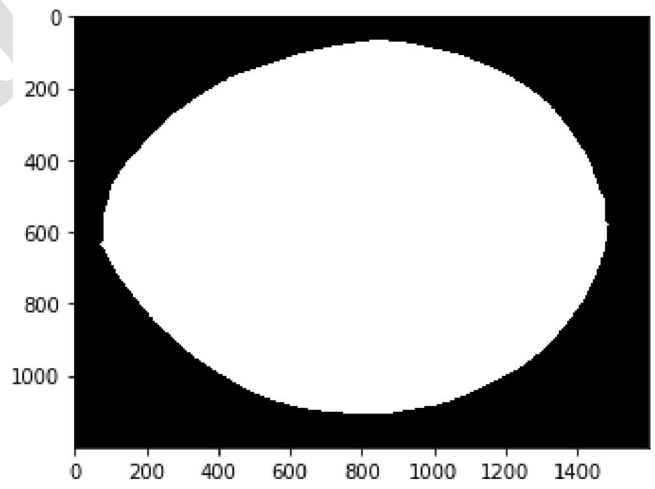

Fig 6: Closing of holes using morphological transformation

Closing of holes plays a very vital role in image processing technique. Morphological transformations are some simple operations based on the image shape. It is normally performed on binary images. It needs two inputs, one is our original image, and second one is called structuring element or kernel which decides the nature of operation. As we can see in previous image, there are some minute holes present which are closed after performing the morphological transformation technique.

(e.) Boundary extraction using contours

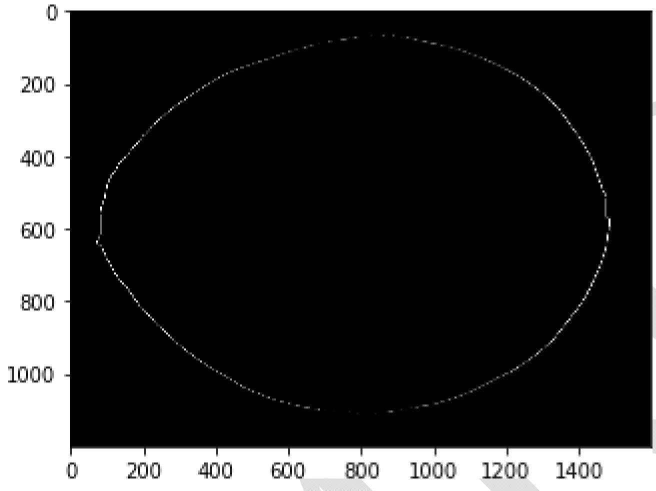

Fig 7: Boundary extraction using contours

Boundary extraction plays a significant role because to get the perfect results, the holes should be closed perfectly. But after performing the morphological transformation technique, there are chances that holes may be present in the leaf boundary (as seen in the image). So, here comes this boundary extraction where it plays a vital role. Contours are defined as the line joining all the points along the boundary of an image that are having the same intensity. Contours come handy in shape analysis, finding the size of the object of interest, and object detection.

(ii.) Feature extraction:

(a.) Shape based features:

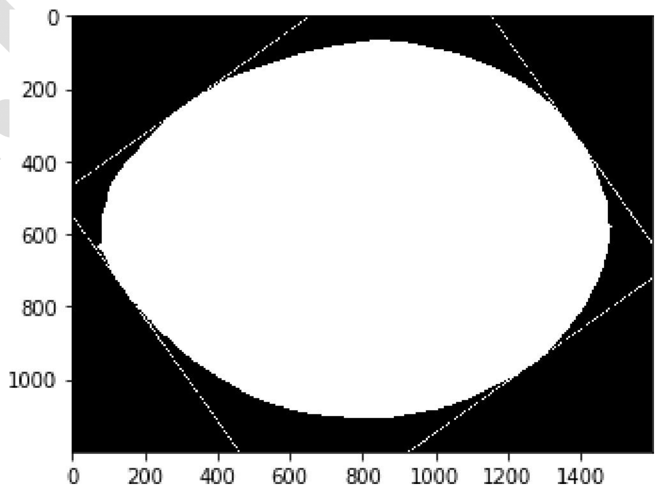

Fig 8: Shape based features

So now, after performing different pre-processing techniques, the next step comes towards extracting features based on the leaf shape. The shape features includes its physiological length, physiological width, area, perimeter, aspect ratio, rectangularity, circularity.

(b.) Colour based features:

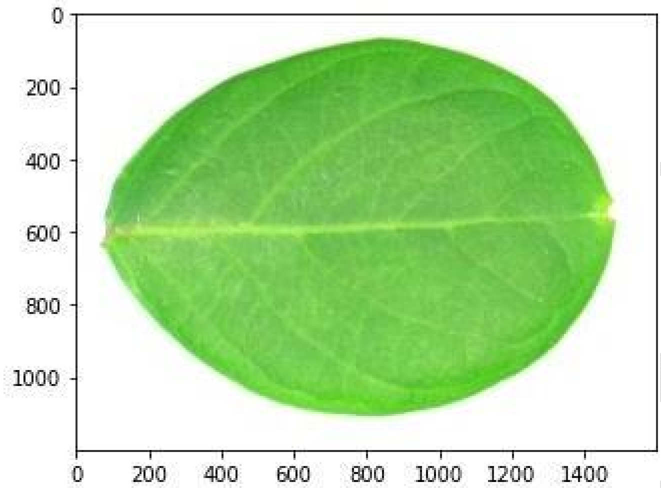

Fig 9: Colour based features

After the shape based features, it is important to extract the colour based features. The colour based features are basically calculated on basis of R, G and B values. The colour based features includes calculating the mean and standard deviations of R,G and B channels.

(c.) Texture based features:

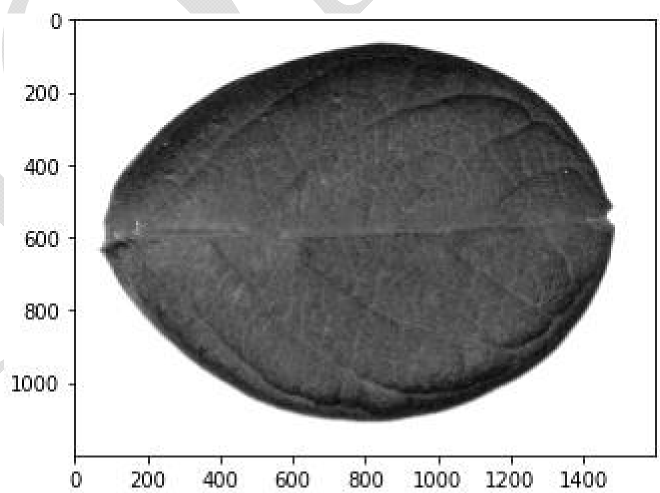

Fig 10: Colour based features

After calculating the shape based features and colour based features, the text based features includes calculating the contrast, correlation, inverse difference moments, entropy.

(iii.) Model Building and Model Testing:

For model building purpose, the algorithm used is the SVM classifier. It is a supervised machine learning algorithm which can be used for both classification and regression challenges. However, it is mostly used in classification problems. In the SVM algorithm, we plot each data item as a point in n- dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiates the two classes very well. Features were then scaled using Standard Scaler. Also parameter tuning was done to find the appropriate hyper parameters of the model using Grid Search CV.

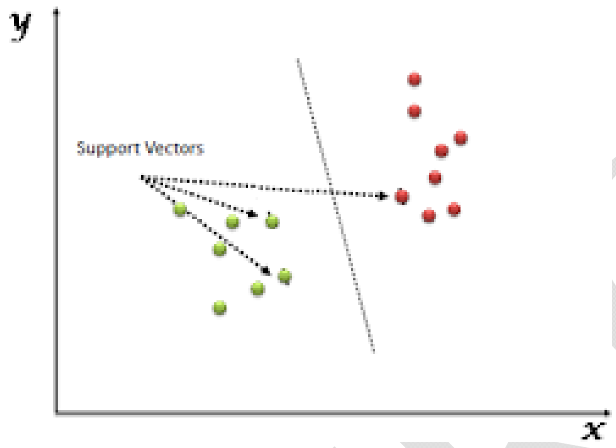

Fig 11: SVM classifier

So, after performing all these operations, the final output received is as follows:

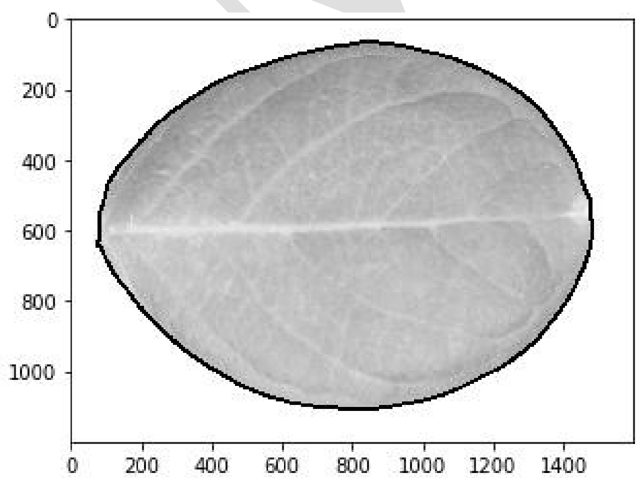

Fig 12: Final Output

After receiving this kind of output, we can easily determine that which kind of leaf it is. After performing all kind of pre-processing techniques, the feature extraction, model building and testing, the leaf obtained as above, we can create a simple GUI with some data stored with images. And as this image is obtained by performing these heavy techniques, it can easily retrieve the data stored and identify the leaf and can give the obtained results. As this project is just built at a small level, but can be a big one if it is to be applied in the industry.

REFERENCES:

1. Khirade, Sachin D., and A. B. Patil. "Plant disease detection using image processing." In *2015* International conference on computing communication control and automation, pp. 768-771. IEEE, 2015.

2. Wu, Stephen Gang, et al. "A leaf recognition algorithm for plant classification using probabilistic neural network." 2007 IEEE international symposium on signal processing and information technology. IEEE, 2007.

3. Barbedo, Jayme Garcia Arnal. "Digital image processing techniques for detecting, quantifying and classifying plant diseases." Springer Plus 2.1 (2013): 660.

4. Priya, C. Arun, T. Balasaravanan, and Antony Selvadoss Thanamani. "An efficient leaf recognition algorithm for plant classification using support vector machine." International conference on pattern recognition, informatics and medical engineering (PRIME-2012). IEEE, 2012.

5. Singh, Vijai, and Ak K. Misra. "Detection of plant leaf diseases using image segmentation and soft computing techniques." Information processing in Agriculture 4.1 (2017): 41-49.

6. Papageorgiou, Constantine, and Tomaso Poggio. "A trainable system for object detection." International journal of computer *vision* 38, no. 1 (2000): 15-33.

7. Chaki, Jyotismita, Ranjan Parekh, and Samar Bhattacharya. "Plant leaf recognition using texture and shape features with neural classifiers." Pattern Recognition Letters 58 (2015): 61-68.

8. Lee, Kue-Bum, and Kwang-Seok Hong. "An implementation of leaf recognition system using leaf vein and shape." International Journal of Bio-Science and Bio-Technology 5.2 (2013): 57-66.

9. Chaki, Jyotismita, Ranjan Parekh, and Samar Bhattacharya. "Plant leaf recognition using texture and shape features with neural classifiers." Pattern Recognition Letters 58 (2015): 61-68.

10. Prasad, Shitala, Krishna Mohan Kudiri, and R. C. Tripathi. "Relative sub-image based features for leaf recognition using support vector machine." Proceedings of the 2011 International Conference on Communication, Computing & Security. 2011.

11. Shabanzade, Maliheh, Morteza Zahedi, and Seyyed Amin Aghvami. "Combination of local descriptors and global features for leaf recognition." Signal & Image Processing 2.3 (2011): 23.

12. Pornpanomchai, Chomtip, et al. "Thai herb leaf image recognition system (THLIRS)." (2011).
13. Gu, Xiao, Ji-Xiang Du, and Xiao-Feng Wang. "Leaf recognition based on the combination of wavelet transform and gaussian interpolation." *International Conference on Intelligent Computing*. Springer, Berlin, Heidelberg, 2005.

14. Hu, R., Jia, W., Ling, H., & Huang, D. (2012). Multiscale distance matrix for fast plant leaf recognition. *IEEE transactions on image processing*, *21*(11), 4667-4672.

15. Deokar, S. R., P. H. Zope, and S. R. Suralkar. "Leaf recognition using feature point extraction and artificial neural network." *International Journal of Engineering* 2.1 (2013).

16. Wang, Zhaobin, Xiaoguang Sun, Yaonan Zhang, Zhu Ying, and Yide Ma. "Leaf recognition based on PCNN." *Neural Computing and Applications* 27, no. 4 (2016): 899-908.

17. Sari, Cihan, Ceyhun Burak Akgül, and Bülent Sankur. "Combination of gross shape features, fourier descriptors and multiscale distance matrix for leaf recognition." *Proceedings ELMAR- 2013*. IEEE, 2013.

18. Hsiao, Jou-Ken, et al. "Comparative study of leaf image recognition with a novel learning-based approach." *2014 Science and Information Conference*. IEEE, 2014.

19. Munisami, Trishen, et al. "Plant leaf recognition using shape features and colour histogram with K-nearest neighbour classifiers." *Procedia Computer Science* 58 (2015): 740-747.

20. Rastogi, Aakanksha, Ritika Arora, and Shanu Sharma. "Leaf disease detection and grading using computer vision technology & fuzzy logic." *2015 2nd international conference on signal processing and integrated networks (SPIN)*. IEEE, 2015.

21. http://www.ijfis.org/journal/view.html?doi=10.5391/IJFIS.2017.17.1.26

22. https://www.sciencedirect.com/science/article/pii/S1877050915022061

23. https://iopscience.iop.org/article/10.1088/1757-899X/288/1/012058

24. https://github.com/AayushG159/Plant-Leaf Identification/blob/master/single_image_process_file.ipynb

25. https://www.researchgate.net/publication/325625631_Plant_leaf_recognition

26. https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example- code/