Recognition of Vegetable Leaf using MATLAB

Ahalya Arun[1], Bhumika U[1], Divyashree K[1], Rashmi H S[1] and Prof. Guruprasad L[2]
1-Student, Dept of ECE, Maharaja Institute of Technology, Mysuru
2-Asst. Prof, Dept of ECE, Maharaja Institute of Technology, Mysuru
VTU university
Mysuru

***Abstract*— Recognizing plants is a vital problem, especially for biologists, agricultural researchers, and environmentalists. Plant recognition can be performed by human experts manually but it is a time consuming and low-efficiency process. Automation of plant recognition is an important process for the fields working with plants. This paper presents an approach for plant recognition using leaf images. In this study, the proponents demonstrated the development of the system that gives users the ability to identify vegetables based on photographs of the leaves taken with a high definition camera. The main objective of this system is to modernize process of identification, so as to automate the way of identifying the vegetable plants through leaf image and digital image processing. The proposed system uses RGB to labspace, K-means, GLCM and Multi-SVM techniques to acquire physical parameters of the leaf using MATLAB.**

Keywords*— labspace, K-means, GLCM, Multi-SVM, MATLAB.

1. INTRODUCTION

In Agriculture, vegetables plants have become an important source of energy and source of living for farmers. Correctly identifying a vegetable leaf allows farmers to differentiate between vegetables as well as a vegetable seedling and weed in the garden. With so many varieties of leafy greens coming from our local farmers each week, it can be difficult to figure out vegetable it belongs to. Though these leaves may appear similar at a glance, they are actually quite unique in terms of Shape, Texture and Color. And with the increasing use of innovative computer technology, digitalized ways have become a possibility for plant identification. The proposed system will solve the problem of determining the vegetables just through the photograph of their leaves. Leaf Recognition is now emerging for research purposes. Leaf recognition technology plays an important role in plant classification and its key issue lies in whether selected features are stable and have good ability to discriminate different kinds of leaves. It is well known that the correct way to extract plant features involves plant recognition based on leaf images. In particular, identification process is carried out by gathering leaves detached from the plants, treated and stained prior to the imaging. Recognition of Vegetable Leaf using Matlab, is to create an Informative Vegetable's Leaf Recognition using Matlab to help the farmers, botanist and Agricultural Researchers in identifying a vegetable and its common details in a convenient and reliable way. The proposed system can be used by the government to survey the agriculture land to keep a tab on the different types of vegetables grown in different regions. It also helps researchers, botanist and farmers to identify the vegetable plant.

2. REVIEW OF RELATED WORKS

[1] "**A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network"** by Stephen Gang Wu1, Forrest Sheng Bao2, Eric You Xu3, Yu-Xuan Wang4, Yi-Fan Chang5 and Qiao-Liang Xiang4 1 Institute of Applied Chemistry, Chinese Academy of Science, P. R. China.

The referred paper employs Probabilistic Neural Network (PNN) with image and data processing techniques to implement a general purpose automated leaf recognition for plant classification. 12 leaf features are extracted and orthogonalized 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 an 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**] "A Leaf Recognition Of Vegetables Using Matlab"** by Nadine Jaan D. Caldito, Eusebelle B. Dagdagan, Mark G. Estanislao, Kim Leonard B. Jutic, Mary Regina B. Apsay, Marissa G. Chua, Jeffrey F. Calim, Florocito S. Camata.

In the above proposed scheme, the proponents demonstrated the development of the system that gives users the ability to identify vegetables based on photographs of the leaves taken with a high definition camera. At the heart of this system is a modernize process of identification, so as to automate the way of identifying the vegetable plants through leaf image and digital image processing. The system used the Gabor Filter, Edge Detection, RGB Color and Grayscale Image to acquire the physical parameter of the leaves. The output parameters are used to compute well documented metrics for the statistical and shape. Base on the study, the following conclusion are drawn: The system can extract the physical parameters from the leaf's image that will be used in identifying Vegetable`s. From the extracted leaf parameters, the system provides the statistical analysis and general information of the identified leaf. The used algorithm can organize data and information to useful resources to the future researchers.

[3] **"Plant Leaf Recognition using Shape based Features and Neural Network classifiers"** by Jyotismita Chaki, Ranjan Parekh

School of Education Technology Jadavpur University Kolkata, India

The above scheme proposes an automated system for recognizing plant species based on leaf images. Plant leaf images corresponding to three plant types, are analyzed using two different shape modeling techniques, the first based on the Moments-Invariant (M-I) model and the second on the Centroid Radii (C-R) model. For the M-I model the first four normalized central moments have been considered and studied in various combinations viz. individually, in joint 2-D and 3-D feature spaces for producing optimum results. For the C-R model an edge detector has been used to identify the boundary of the leaf shape and 36 radii at 10 degree angular separation have been used to build the feature vector. To further improve the accuracy, a hybrid set of features involving both the M-I and C-R models has been generated and explored to find whether the combination feature vector can lead to better performance. Neural networks are used as classifiers for discrimination. The data set consists of 180 images divided into three classes with 60 images each. Accuracies ranging from 90%-100% are obtained which are comparable to the best figures reported in extant literature.

[4] "**A Leaf Recognition Technique for Plant Classification Using RBPNN and Zernike Moments"** by A.H. Kulkarni, Dr. H.M.Rai, Dr. K.A.Jahagirdar, P.S.Upparamani.

Plants are among the earth's most useful and beautiful products of nature. Plants have been crucial to mankind's survival. The urgent need is that many plants are at the risk of extinction. About 50% of ayurvedic medicines are prepared using plant leaves and many of these plant species belong to the endanger group. So it is indispensable to set up a database for plant protection. We believe that the first step is to teach a computer how to classify plants. Leaf /plant identification has been a challenge for many researchers. Several researchers have proposed various techniques. In this paper we have proposed a novel framework for recognizing and identifying plants using shape, vein, color, texture features which are combined with Zernike movements. Radial basis probabilistic neural network (RBPNN) has been used as a classifier. To train RBPNN we use a dual stage training algorithm which significantly enhances the performance of the classifier. Simulation results on the Flavia leaf dataset indicates that the proposed method for leaf recognition yields an accuracy rate of 93.82%.

[5] **"Plant leaf recognition using shape featrures and colour histogram with k-nearest neighbor classifiers"** by Trishen Munisami, Mahesss Ramsurn, Somveer Kishnah, Smaeerchand Pudaruth.

Automated systems for plant recognition can be used to classify plants into appropriate taxonomies. Such information can be useful for botanists, industrialists, food engineers and physicians. In the proposed scheme, a recognition system capable of identifying plants using their images of their leaves has been developed. A mobile application was also developed to allow user to take pictures of leaves and upload them to server. The serve runs pre-processing and feature extraction techniques on the image before a pattern matcher compares the information from this image with the ones in the database in order to get potential matches. The different features that are extracted are the length and width of the leaf, the area of the leaf, the perimeter of the leaf, the hull area, the hull perimeter, a distance map along the vertical and horizontal axes, a color histogram and a centroid-based radial distance map. A k-nearest neighbor classifier was implemented and tested on 640 leaves belonging to 32 different species of plants. An accuracy of 83.5% was obtained. The system was further enhanced by using information obtained from a color histogram which increased the recognition accuracy to 87.3%.

3. SYSTEM MODEL DESCRIPTION

For easy understanding of our system, we have divided our scheme into 5 modules, namely

1. Image Acquisition
2. Image Pre-processing
3. Image Segmentation
4. Feature Extraction
5. Feature Classification

3.1 Image Acquisition

In this module, we will capture the images of the vegetable leaves to be recognized using a high definition camera having resolution 16 MP or above. In our scheme we have considered 5 different classes with 20 samples each. Images are taken keeping camera perpendicular to the leaves inside a studio type box. Studio type box is a box which is a box which is an imitation of a studio for high quality pictures.

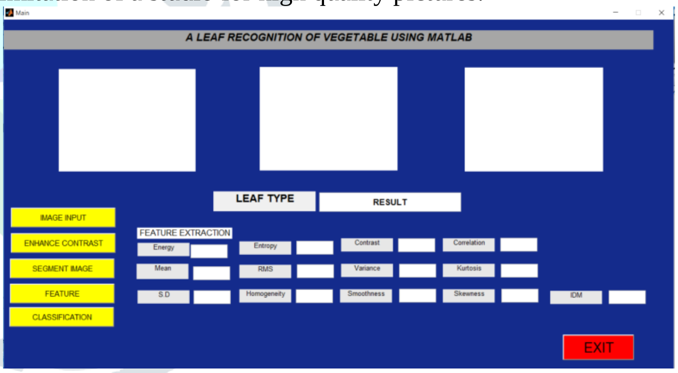


Fig.1: Vegetable leaf recognition main form

3.2 Image Pre-processing

Enhancement process is used so that it aids in adjusting the digital image hence obtained result are more suitable for further image analysis. The use of leaf image data for leaf database requires several preprocessing procedure. These procedures include; resizing of the image and image enhancement which improves the leaf image quality. These image preprocessing techniques doesn't affect the leaf image content. The goal of digital image preprocessing is to increase both the accuracy and the interpretability of the digital data during the image processing phase. In the work proposed image pre-processing methods are applied to the captured image which are stored in image database. A successful implementation of this step produces improved results and higher accuracy rates. After an image is acquired, it goes through different levels of processing as mention above before it is ready for the next step of image segmentation. The following are the reasons why image preprocessing is important:

- It is important technique for creating a level of similarity in the general features of leaf image, like the size aspect. This enhances the comparison between leaf images.
- It is use for image enhancement.
- It improves the quality of image information.
- It eases the process of feature extraction, on which the recognition depends.

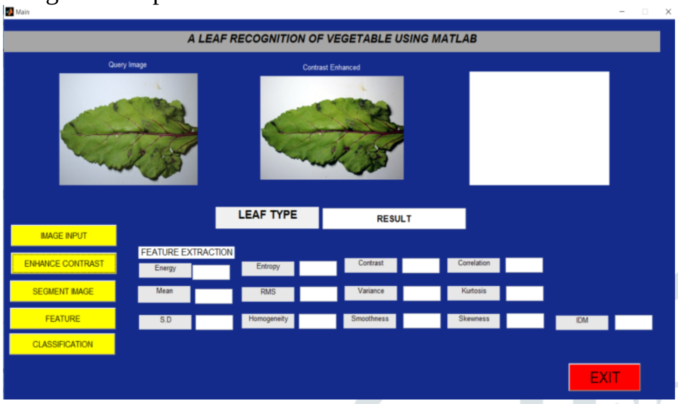


Fig.2: Enhanced image

3.3 Image Segmentation

Image segmentation is a method of partitioning a digital image into multiple segments. It is used to identify objects or other relevant information in digital images. The method used in image segmentation is K-means algorithm. It is a method of vector quantization, originally from signal processing, that is popular for cluster analysis in data mining. K-means clustering aims to partition n observations into k clusters in each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. This results in a partition of the data space into Voronoi cells. We can select the cluster based on our region of interest. The k-means algorithm works as follows:

- First we initialise k points, called means or centroids randomly.
- We categorize each item to its closest mean and we update the mean's coordinates, which are the averages of the items categorized in that mean so far.
- We repeat the process for a given number of iterations and at the end, we have our clusters.

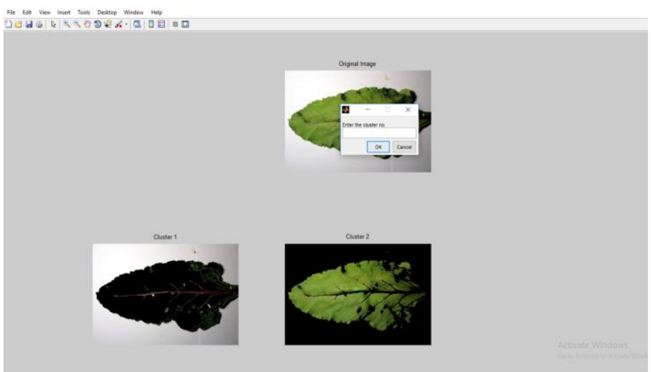


Fig.3: Selection of clusters based on region of interest

3.4 Feature Extraction

In machine learning, pattern recognition and in image processing, feature extraction starts from an initial set of measured data and builds derived values(features) intended to be informative and non-reduntant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations. Feature extraction is related to dimensionally reduction. An image texture is a set of metrics computed in image processing intended to enumerate the apparent texture of a leaf image. Leaf Image Texture gives information regarding the spatial arrangement of color or intensities in a leaf image or selected region of a leaf image. A diversity of techniques has been used for computing texture such as co-occurance matrix, Fractals, Gabor filters, variations of wavelet transform. Additional techniques have also been proposed for relating the local patterns using texture spectrum, for characterization of texture using a composition of edge information and co-occurance matrix properties. The recognition of explicit textures in an image is achieved principally by modeling texture as a two-dimensional gray level variation. The resultant two dimensional arrays are called as Gray Level Co-occurance Matrix (GLCM). First we must convert the RGB Image to Grayscale image. The 13 features extracted using GLCM Algorithm are, Standard Deviation, Mean, Variance, Contrast, Correlation, Energy, Entropy, Homogeneity, RMS ,IDM, Smoothness, Skewness and Sharpness. These 13 features includes all texture, color and shape features.

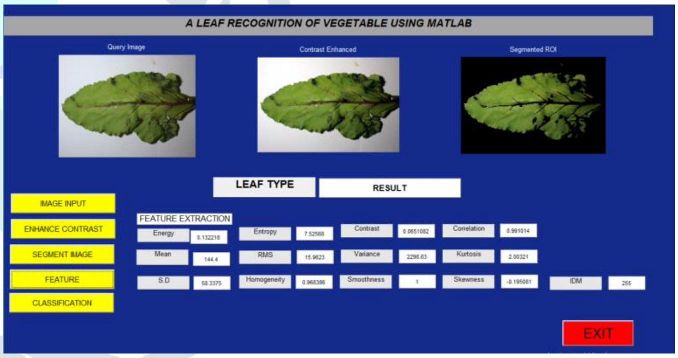


Fig.4: Extracted features along with values

3.5 Feature Classification

Feature classification is the grouping of features based on some criteria. Multi SVM is a binary linear classifier supervised learning models with associated learning algorithms that analyses data and recognizes patterns used for classification and analysis. Classification task usually involves separating data into training and testing sets. Each instance in the training set contains one \target value" (i.e. the class labels) and several attributes" (i.e. the features or observed variables). The goal of multi SVM is to produce a model (based on the training data) which predicts the target values of the test data given only the test data attributes. A multi Support Vector Machine (SVM) is a discriminative classifier formally defined by a separating hyper plane. In particular, by comparing with the 1-NN and k-NN classifiers, it can be found that the multi SVM classifier can not only save the storage space but also reduce the classification time under the case of no sacrificing the classification accuracy.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.0329	0.9939	0.1836	0.9849	147.3954	52.3838	7.4482	15.9686	1.2707e+03	1.0000	1.8041	-0.2717	255
2	0.0375	0.9950	0.1518	0.9813	157.1256	61.4065	7.5938	15.9679	866.0214	1.0000	1.8051	-0.0394	255
3	0.0316	0.9952	0.1636	0.9847	133.2233	57.6100	7.4763	15.9619	775.2858	1.0000	1.6443	-0.0816	255
4	0.0437	0.9932	0.1487	0.9787	146.0328	55.7324	7.5149	15.9666	1.5768e+03	1.0000	2.0701	0.0953	255
5	0.0421	0.9921	0.1648	0.9789	153.4851	52.2342	7.5069	15.9575	1.5420e+03	1.0000	2.6755	-0.0161	255
6	0.0363	0.9929	0.1656	0.9824	161.6958	50.0624	7.4218	15.9687	1.6423e+03	1.0000	2.3117	-0.2210	255
7	0.0424	0.9925	0.1565	0.9806	146.2162	52.3722	7.5661	15.9663	2.3440e+03	1.0000	2.2513	0.0626	255
8	0.0451	0.9931	0.1482	0.9792	129.3684	56.9091	7.7148	15.9247	2.3429e+03	1.0000	2.2925	0.1073	255
9	0.0242	0.9965	0.1431	0.9880	145.4770	58.4775	7.7014	15.9648	2.0584e+03	1.0000	1.9252	0.0060	255
10	0.0270	0.9943	0.1783	0.9867	138.6496	49.5134	7.4214	15.9687	700.7689	1.0000	1.8261	0.0076	255
11	0.0556	0.9912	0.1701	0.9729	145.1990	55.1263	7.2831	15.9684	1.5345e+03	1.0000	2.0636	0.1101	255
12	0.0612	0.9887	0.1911	0.9713	128.8439	52.2962	7.3260	15.9554	993.6451	1.0000	2.0438	0.2475	255
13	0.0450	0.9925	0.1605	0.9790	165.9863	54.8046	7.5365	15.9687	2.6200e+03	1.0000	2.5468	-0.4236	255
14	0.0358	0.9901	0.2204	0.9828	161.9175	43.1669	7.2530	15.9687	1.5452e+03	1.0000	3.1409	-0.7466	255
15	0.0465	0.9927	0.1475	0.9774	147.5457	56.3155	7.6891	15.9671	3.0491e+03	1.0000	2.1943	0.0548	255
16	0.0556	0.9895	0.1657	0.9741	150.3343	50.3892	7.5057	15.9679	2.4632e+03	1.0000	2.8276	-0.3471	255
17	0.0757	0.9871	0.2198	0.9668	154.4570	56.2255	7.0637	15.9687	2.5911e+03	1.0000	1.4655	-0.2501	255
18	0.0796	0.9837	0.1591	0.9640	157.3573	48.8967	7.3326	15.9687	1.8925e+03	1.0000	1.9511	-0.0818	255
19	0.0920	0.9850	0.1493	0.9569	126.4997	55.4485	7.5493	15.9574	1.7715e+03	1.0000	2.5370	0.2842	255
20	0.0588	0.9893	0.1657	0.9713	143.5383	52.3783	7.5483	15.9566	2.6421e+03	1.0000	2.7549	0.0794	255
21	0.0799	0.9873	0.1430	0.9634	127.3513	55.8711	7.6416	15.9540	2.0360e+03	1.0000	2.1663	0.3044	255
22	0.0892	0.9876	0.1554	0.9604	148.9813	58.6997	7.4030	15.9683	2.3174e+03	1.0000	1.8749	0.0789	255
23	0.0835	0.9875	0.1482	0.9642	140.3089	57.4717	7.5457	15.9650	2.8900e+03	1.0000	1.7242	-0.0120	255
24	0.0617	0.9896	0.1386	0.9694	142.4127	53.9591	7.6893	15.9596	1.3371e+03	1.0000	2.3116	-0.0135	255
25	0.0748	0.9879	0.1388	0.9646	149.0738	56.6420	7.7001	15.9634	2.3818e+03	1.0000	2.2699	0.0370	255

Fig.5: Trained data set values stored using SVM

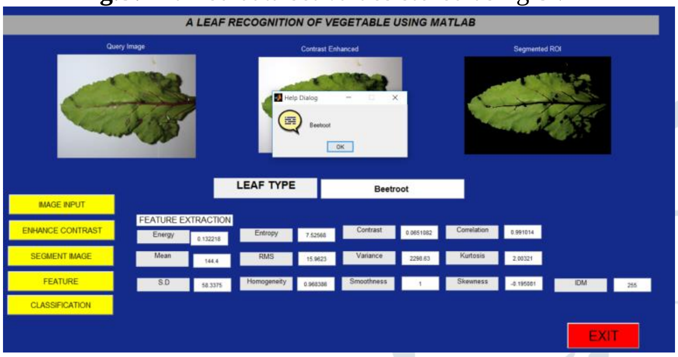
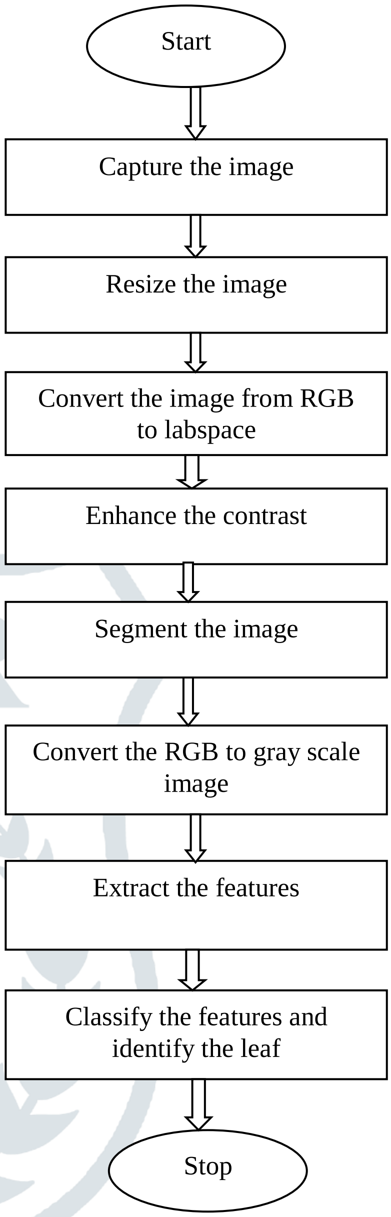


Fig.6: Display of result

4. SYSTEM FLOW



5. RESULT

The objectives of our project were to modernize the process of recognition of vegetable plants using its leaf images, by automating its classification and extracting its features from it's image.
The formula used to calculate the percentage is,
P=(X/N)*100;
Where, P= Percentage
X=No. of matched
N=Total no. of leaves
The accuracy of our project is approximately 94% where we have consider 5 classes with 20 samples each.

Trials	Tested leaves	Matched	Failed to match	Percentage
Cabbage	20	20	0	100
Cauliflower	20	20	0	100
Spinach	20	16	4	80
Beetroot	20	20	0	100
Radish	20	18	2	90

6. CONCLUSION AND FUTURE SCOPE

The objectives have been achieved.
Further work can be done by adding more classes and samples in each class and improve the database.
After the processing of data gathered in this study, we have come up with the following conclusion:
1 We designed the proposed system, A Leaf Recognition of Vegetable which can determine the identification of the leaf tested with the use of image processing and MATLAB as the primary programming language.
2 We tested the program by experimental evaluation. By having the computed information, the system gives an accuracy of ninety four percent (94%).
3 We can able to implement a useful system for the use of the student and future researchers.

7. REFERENCES

[1] "**A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network"** by Stephen Gang Wu1, Forrest Sheng Bao2, Eric You Xu3, Yu-Xuan Wang4, Yi-Fan Chang5 and Qiao-Liang Xiang4 1 Institute of Applied Chemistry, Chinese Academy of Science, P. R. China.

[2] "**A Leaf Recognition Of Vegetables Using Matlab"** by Nadine Jaan D. Caldito, Eusebelle B. Dagdagan, Mark G. Estanislao, Kim Leonard B. Jutic, Mary Regina B. Apsay, Marissa G. Chua, Jeffrey F. Calim, Florocito S. Camata.

[3] **"Plant Leaf Recognition using Shape based Features and Neural Network classifiers"** by Jyotismita Chaki, Ranjan Parekh School of Education Technology Jadavpur University Kolkata, India

[4] "**A Leaf Recognition Technique for Plant Classification Using RBPNN and Zernike Moments"** by A.H. Kulkarni, Dr. H.M.Rai, Dr. K.A.Jahagirdar, P.S.Upparamani.

[5] "**A Leaf Recognition Technique for Plant Classification Using RBPNN and Zernike Moments"** by A.H. Kulkarni, Dr. H.M.Rai, Dr. K.A.Jahagirdar, P.S.Upparamani.