DESIGNING OF LEAVES CLASSIFIER FOR AYURVEDIC PLANTS USING HYBRID FEATURE EXTRACTION

By

Hinal J. Patel (170410754002)

Guided by

Dr. Nehal G. Chitaliya

Associate Professor, E & C Dept., SVIT, Vasad

A Thesis Submitted to
Gujarat Technological University
In Partial Fulfillment of the Requirements for
The Degree of Master of Engineering
In Electronics & Communication Engineering

May - 2019

Department of Electronics & Communication Engineering
Sardar Vallabhbhai Patel Institute of Technology
Vasad-388306, Gujarat, India.

CERTIFICATE

This is to certify that research work embodied in this thesis entitled **"Designing Of leaves Classifier for ayurvedic plants using Hybrid Feature Extraction"** was carried out by **Ms. Patel Hinal Jayeshbhai (170410754002)** at **Sardar Vallabhbhai Patel Institute of Technology, Vasad (041)** for Partial fulfillment of Master of Engineering degree in Embedded system Engineering to be awarded by Gujarat Technological University. This research work has been carried out under my guidance and supervision and it is up to my satisfaction.

Date: 27/04/2019

Place: SVIT – Vasad

Guide:

Nehal

Dr. Nehal G. Chitaliya
Associate Professor
E&C Department,
SVIT, Vasad.

Head:

29/04/2019

Dr. Y. B. Shukla
Head of Department
E&C Department,
SVIT, Vasad.

Sardar Vallabhbhai Patel Inst. Of Tech. VASAD 388 306

Seal of Institute

COMPLIANCE CERTIFICATE

This is to certify that research work embodied in this dissertation titled **"Designing of leaves Classifier for ayurvedic plants using Hybrid Feature Extraction"** was carried out by **Ms. Patel Hinal Jayeshbhai (170410754002)** at **Sardar Vallabhbhai Patel Institute of Technology, Vasad (041)** for partial fulfilment of Master of Engineering degree in Embedded System Engineering to be awarded by Gujarat Technological University. She has complied to the comments given by the Dissertation phase – I as well as Mid Semester Thesis Reviewer to my satisfaction.

Date: 27/04/2019

Place: SVIT - Vasad

Hinal J. Patel
(170410754002)

Dr. Nehal G. Chitaliya
Associate Professor
E&C Department,
SVIT, Vasad.

Sardar Vallabhbhai Patel Inst. Of Tech. VASAD 388 306

Seal of Institute

PAPER PUBLICATION CERTIFICATE

This is to certify that research work embodied in this dissertation titled **"Designing of leaves Classifier for ayurvedic plants using Hybrid Feature Extraction"** was carried out by **Ms. Patel Hinal Jayeshbhai (170410754002)** at **Sardar Vallabhbhai Patel Institute of Technology, Vasad (041)** for partial fulfilment of Master of Engineering degree in Embedded System Engineering to be awarded by Gujarat Technological University. Her article **Designing of leaves Classifier for ayurvedic plants using Hybrid Feature Extraction** has been published for publication by the **International Journal of Research in Electronics and Computer Engineering (IJRECE),** Print ISSN: 2393-9028, Online ISSN: 2348-2281 VOL. 7, ISSUE 2, APRIL - JUNE 2019.

Date: 27/04/2019

Place: SVIT-Vasad

Hinal J. Patel
(170410754002)

Dr. Nehal G. Chitaliya
Associate Professor
E&C Department,
SVIT, Vasad.

Signature/Stamp of Principal
(Dr. D. P. Soni)

Sardar Vallabhbhai Patel Inst. Of Tech. VASAD 388 306

Seal of Institute

THESIS APPROVAL CERTIFICATE

This is to certify that research work embodied in this dissertation titled **"Designing of leaves Classifier for ayurvedic plants using Hybrid Feature Extraction"** was carried out by **Ms. Patel Hinal Jayeshbhai (170410754002)** at **Sardar Vallabhbhai Patel Institute of Technology, Vasad (041)** is approved for the degree of Master of Engineering with specialization of Embedded System Engineering by Gujarat Technological University.

Date: 27/04/2019

Place: SVIT - Vasad

Examiner's Sign and Name:

....

( ) ( )

Sardar Vallabhbhai Patel Inst. Of Tech. VASAD 388 306

DESCLARATION OF ORIGINALITY

We hereby certify that we are the sole authors of this thesis and that neither any part of this thesis nor the whole of the thesis has been submitted for a degree to any other University or Institution.

We certify that, to the best of our knowledge, the current thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations or any other material from the work of other people included in our thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that we have included copyrighted material that surpasses the boundary of fair dealing within the meaning of the Indian Copyright (Amendment) Act 2012, we certify that we have obtained a written permission from the copyright owner(s) to include such material(s) in the current thesis and have included copies of such copyright clearances to our appendix.

We declare that this is a true copy of thesis, including any final revisions, as approved by thesis review committee.

We have checked write up of the present thesis using anti-plagiarism database and it is in allowable limit. Even though later on in case of any complaint pertaining of plagiarism, we are sole responsible for the same and we understand that as per UGC norms, University can even revoke Master of Engineering degree conferred to the student submitting this thesis.

Date:

Place: SVIT – Vasad

Signature of Student

Hinal J. Patel

(170410754002)

Sardar Vallabhbhai Patel Inst. Of Tech. VASAD 388 306

Signature of Guide

Dr. Nehal G. Chitaliya

Associate Professor

E&C Department,

SVIT (041), Vasad.

ACKNOWLEDGEMENT

First of all, I would like to thank my Guide **Dr. Nehal G. Chitaliya (Associate Professor, SVIT-Vasad)** for her guidance through several aspects of the system that only an experience person like her can give. Through her guidance I discovered several new ways of doing the same things, but in a more appropriate and systematic fashion. We thank her deeply.

Secondly, I would like to thank my Head of Department, **Dr. Y. B. Shukla** for being there for me whenever I needed his guidance and helping me not only through the dissertation but throughout our whole journey till present.

Thirdly, I would like to thank staff members of Electronics and Communication Engineering Department, S.V.I.T., Vasad, for their guidance and support. Last but not least I would like to thanks my friends, classmates and family members for their support and love throughout this dissertation work.

Hinal J. Patel
(170410754002)

TABLE OF CONTENTS

Index

List Of Figures

LIST OF TABLES

ABSTRACT

Today, An Ayurvedic leaf identification is most important task in ayurvedic medicinal industry. For the leaf identification, the Contourlet Transform (CT) and Canny edge detector technique using ANN (Artificial Neural Network), SVM (Support Vector Machine) and RF (Random Forest) Classifier is proposed. Firstly, apply pre-processing on the each image of dataset. Then leaf segmentation using CT and Canny edge detector on all images those are pre-processed by Gray-scalling and filtering. Contourlet transform is decomposed each image from the training dataset and Canny edge detector is extract contour of leaf region and extract vein features from training dataset and extract features like shape, color, GLCM based texture from the dataset of leaf. The features vector is matches with the classifier for classifying the final class and finally it identify the leaf species with accurate result.

Key Words: Leaf Classification, Contourlet Transform, Random Forest (RF), Artificial Neural Network (ANN), canny edge detector, Support Vector Machine (SVM)

CHAPTER 1

INTRODUCTION

1.1 Overview

Into Our proposed work we used Indian Ayurvedic plant's leaves for identifying and classifying plant species. Ayurvedic Medicinal plants are most widely used in the medicinal industry to make medicines. In the ancient past approximately 5000 years ago, The Indian peoples are used ayurvedic plants as medicine. The main element of ayurvedic plant is leaves, fruits, Flowers, roots and bark etc. But leaves are available in all season and we can easily identify its species. India has more than 8000 plants for medicinal value, the herbal medicines are used 1500 plants species, 500 plants are used in Commercial Ayurvedic preparation and an ayurvedic industry used 80% of plants species which are assembled from the forest and agricultural farm.

Today, the ayurvedic leaves are collected by a human from the forest area; those are not professionally trained to identify the correct ayurvedic plant species and those are very few in number. If manufacturing unit receives incorrect or surrogate medicinal plant and if there were Misunderstanding due to species name variation is also unchecked. If certain plants contain dry leaves, in this situation the manual identification task is too onerous. So, Medicines are abortive by uses of an incorrect medicinal plant that it makes difficulty and side effects of drugs on human's life [1].

In the previous work done, there were several methods in proposed work such as shape, color, texture and vein etc. But there were no used any combination of all features so we are trying to combine all features for increase accurate result and Machine learning classification approach also present to classifying the leaves. Many leaves features are the same to each other leaves so classifier makes groups for same characteristics that it can easily give an accurate result of the system.

Figure 1: Simple Plant Identification and Classification of system

Above figure is a simple identification and classification flow of system. The system is process by in this manner. There were many types of methods in image processing for increase accuracy result of the system.

1.2 Purpose of Research

Today, identification and classification of unknown plant species are performed manually by expert personnel who are very few in number. The important aspect is to develop a system which classifies the plants. This approach presents a new recognition approach based on Leaf Features Fusion and ML Classification algorithms for classifying the different types of plants.

1.3 Scope And Objective

Scope:

The Scope of research work is:

- Using different Segmentation Techniques like contour edge detection, Otsu's Thresholding to segment the Leaf and by different methods, we can extract the features like shape, color, Texture, Vein etc from leaves images.
- Using different classification Techniques Identifying and classifying the type of ayurvedic plant.

Objective:

The Objective of research work are:

- To confirm the importance of leaf length, width, area and perimeter.
- To improve learning of plant species using multi-classification.
- To improve the effective use of medicinal properties of the plant with its identification.
- To increase the accuracy of Plant Identification based on Hybrid Feature Fusion.
- To be suitably depicted with the assistance of different classification Approach Using SVM and RF Classifier.

1.4 Problem Definition

Plant Species knowledge is essential for protecting biodiversity. The identification of leaf by conventional keys is complex, time-consuming, and due to the use of the type of plant botanical terms frustrating for non-experts. This creates a hard to overcome hurdle for novices interested in acquiring Plant classification. In Existing systems, there are many feature extraction methods which is Contain feature combination of Shape-Color, Shape-texture or Color-Texture but there is not use any Venation feature or Invariant Feature with this Combination. So in our Proposed System trying to combine shape, color, texture, vein and invariant feature to overcome all problems and Machine Learning classification approach is present in the proposed system.

1.5 Application of System

1. Agriculture Industry
2. Biomedical Department
3. Ayurveda Department
4. Forensic Department

1.6 Thesis Layout

Chapter 1 is Introduction, which includes the major concepts of the thesis such as basics information about Ayurvedic leaves and plant identification system, Motivation, Problem Definition and Application of system.

Chapter 2 is a Literature Survey, in which discussion has been made on various kind of literature studied during the thesis.

Chapter 3 is a Proposed System, which includes proposed block diagram and methodology of the system.

Chapter 4 is an Implementation & Result Of system, which is experimented on MATLAB software.

Chapter 5 is a conclusion, which concludes the proposed system

References

Appendix

CHAPTER 2

LITERATURE REVIEW

The Literature Review is divided in to three phases, first phase discuss about the major area of research work that is leaf Segmentation Techniques. Second phase gives brief summary of feature extraction method with its merits and demerits and last section discuss about leaf classification of previous discussion on related work annotated in various research paper.

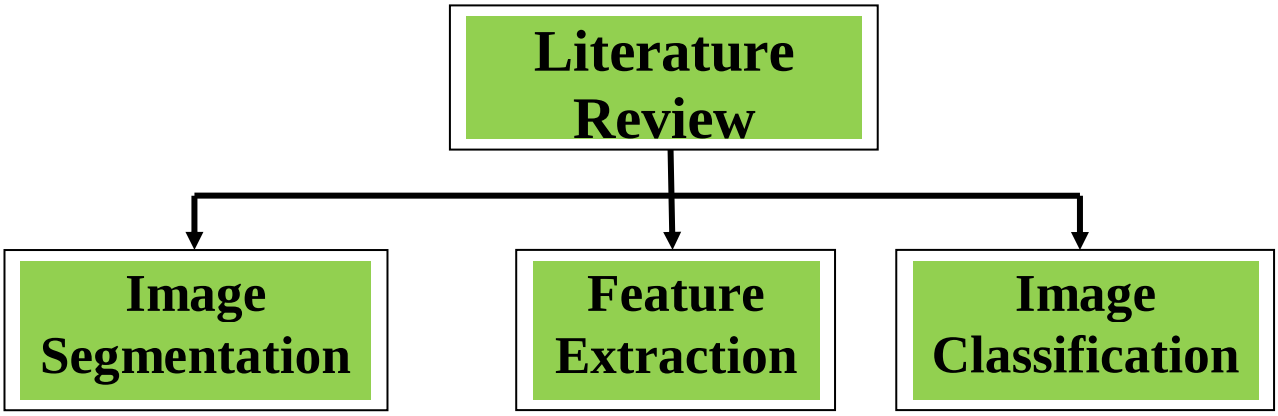


Figure 2: Types of Literature Review

2.1 Leaf Segmentation Methods

The segmentation methods are used for convert grayscale image into binary image. It is partitioning the digital image into multiple segments or sets of pixels so it can be easier to extract the features. There were many segmentation methods are used in Literature those are describe below:

2.1.1 Morphological Operation [1, 2, 5, 8]:

Morphological image processing is a collection of non-linear operations relating to the image's shape or morphology of features. Morphological operations rely solely on the relative ordering of pixel values and not on their numerical values and are therefore particularly suitable for binary image processing. Morphological operations can also be applied to gray images in such a way that their light transfer functions are unknown and their absolute pixel values are therefore of no or minor interest. Morphological techniques test an image that is called a structuring element with a small shape or template. The structuring element is positioned in the image at all possible locations and is compared to the corresponding pixel neighbourhood.

2.1.2 OTSU's Thresolding [3]:

Otsu's Thresholding method involves iterating across all possible threshold values and calculating a spread measure for each side of the threshold for the pixel levels and it is an image Thresholding algorithm to reduction of grayscale image into binary image. It is minimize the two intraclass variance of the threshold black (0) and white (1) pixels. It uses the image histogram to search for the threshold. It maximizes the segmented classes ' "between class variance." Otsu shows that minimizing "within class variance" is the same as maximizing the segmented classes "between class variance." And maximizing "between class variance" is less expensive computationally than minimizing "within class variance." It means the gray image pixels are divided into two classes either foreground or background. The threshold value is sum of the foreground and background spreads is at its minimum.

2.1.3 K-means Clustering [10]:

K-means clustering is a method of vector quantization that is popular for cluster analysis in data mining, originally from signal processing. K-means is one of the simplest uncontrolled learning algorithms to solve the well-known problem of clustering. The procedure follows a simple and easy way to classify a given set of data through a number of fixed apriority clusters (assuming k clusters). Defining k centers is the main idea, one for each cluster. These centers should be placed in a cunning manner because they cause different results from different locations. So, the better choice is to keep them as far from each other as possible. The next step is to take each point that belongs to a set of data and to associate it with the nearest center. The first step is completed when no point is pending and an early group age is completed. At this point, as a bary center of clusters resulting from the previous step, we need to re-calculate k new Centroid. A new binding must be made between the same data set points and the nearest new center after we have these k new Centroid. A loop was created and after that the k centers will change their location step by step until no changes are made or, in other words, centers will no longer move.

2.1.4 Canny edge detector [6, 16]:

The Canny Edge Detection is an operator that uses a multi-stage algorithm to detect a wide Range of edges in images. This System is applied for both color image and gray-scale image. The Canny edge detection is a technique to extract some useful structural information from different visual objects and dramatically reduce the amount of data to be processed.

2.2 Leaf Feature Extraction Methods:

2.2.1 Curvelet Transform Descriptors (CTD) [3]:

Curvelet transform (CT) is a multi-scale, multi-directional transformation capable of representing edges along curves effectively. It outperforms other transforms as the curvelet coefficients not only display the location and scale data, but also display the image feature orientation data. CT has a small number of curve handling coefficients, so the discontinuities can be treated well the decomposition of Curvelet Transform.

2.2.2 Local Binary Pattern (LBP) [3, 4]:

A local binary pattern (LBP) is a visual descriptor type that is used for computer vision classification. The LBP operator is a robust way to obtain a description of the texture. LBP operator extracted information is invariant against local grayscale variations as well as invariant of rotation. It has been discovered to be a strong feature for texture classification; it has also been determined that when LBP is combined with the descriptor Histogram of Oriented Gradients (HOG), the detection performance on some datasets is significantly improved.

2.2.3 Texture Feature (GLCM) [1, 2, 3, 11]:

The Gray Level Co-occurrence Matrix (GLCM) following features is calculated which is Energy, Contrast, Correlation and Entropy. Energy measures the orderliness how regularly and orderly pixels are in window. Contrast is the difference in color and light between parts of an image. Correlation is a single number that describes the degree of relationship between two Pixels. Entropy is the amount of information which must be coded for by a compression algorithm.

2.2.4 Zernike Moment [1]:

In computer science, the polynomials of Zernike are a sequence of orthogonal polynomials on the unit disk. With the help of regular shape descriptors, irregular objects such as plant leaves cannot be properly described. Moments of Zernike have a higher vector space feature and are of N order. We get higher probability of recognition when the order is high. The polynomials of Zernike defined in a circle x2 + y2 = 1 are a set of complex and orthogonal polynomials.

2.2.5 Shape Features [1]:

Table 1: Definition of Shape Features used in literature

Equivalent Diameter	$\sqrt{\frac{4 * Area}{\pi}}$
Area[5]	The Area Specifies the actual number of pixels in the region.
Solidity	$\frac{Convex\ area}{Original\ area}$
Convex Area[5]	That specifies the number of pixels in convex image. A convex image is a binary image that specifies the smallest convex polygon that can contain the region._ Each row of the matrix contains the *x* and *y* coordinates of one vertex of the polygon, With all of the pixels filled in within the polygon.
Eccentricity[5]	$\sqrt{1 - \frac{Minor\ axis\ length^2}{Major\ axis\ length^2}}$
Minor axis length & Major axis Length[5]	That specifies the length (in pixels) of the major axis and minor axis of the ellipse that has the same normalized second central moments as the region.
Extent	$\frac{no.of\ pixels\ in\ the\ region}{no.of\ pixels\ in\ the\ bounding\ box}$
Compactness[2]	$\pi * \frac{Area\ of\ leaf}{perimeter^2}$
Perimeter[2]	That specifies the distance between each adjoining pair of pixels around the boundary of the region.
Aspect Ratio[2]	$\frac{Maximum\ axial\ length}{Minimum\ axial\ length}$
Entirety	Convex Area – Area / Area
Perimeter Ratio	P/L+W; P is perimeter, L is length, W is width
Centroid	Calculate Centroid for connected components in the image.
Roundness[2]	It is based on the ratio between the inscribed and the circumscribed circles.

2.2.6 Color features:

1. HU Moment [1, 2]

Color moment are calculated by its means, Standard Deviation, Skewness and Kurtosis those equations are describe below and the first color moment can be interpreted as the image's average color and can be calculated by, The second moment of color is the standard deviation Obtained by taking the square root of the color distribution variance. The third moment in color is the skewness. It measures how asymmetrical the color distribution is and therefore provides information on the color distribution shape. Kurtosis is the fourth color moment, and it provides information about the color distribution shape, similar to skewness. More specifically, in comparison with normal distribution, kurtosis is a measure of how flat or high the distribution is there.

2.3 Leaf Classification Method:

2.3.1 Neural Network (NN) [2]:

A neural network is a neuron network or circuit, or an artificial neural network consisting of artificial neurons or nodes in a modern sense. Those receive input, modify its inner state (activation) by that input and produces output based on input and activation. An artificial neuron mimics the working of a biophysical neuron with inputs and outputs, but is not a biological neuron model. By connecting the output of certain neurons to the input of other neurons, the network forms a directed, weighted graph. The weights as well as the functions that calculate the activation can be modified by a learning process that is governed by a rule of learning, an artificial neural network is an interconnected node group inspired by a brain neuron simplification. Each circular node here represents an artificial neuron and an arrow represents a connection between one artificial neuron's output and another's input.

2.3.2 K-nearest neighbors (K-NN) [1, 3]:

K-nearest neighboring algorithm (k-NN) is a non-parametric classification and regression method. Non-parametric means that the underlying distribution of data is not assumed. In other words, from the dataset, the model structure is determined. It will be very useful in practice in which most mathematical theoretical assumptions are not followed by the adult world datasets. Lazy algorithm means that there is no need for model generation training data points. It helps to make this slower and more expensive to train and test phase. Phase of expensive testing means time and memory. In its worst situation, it takes more time for KNN to scan all datasets and it will take more memory to store training data to scan all data points.

2.3.3 Deep Neural Network (DNN) [5]:

A deep neural network (DNN) is an artificial neural network (ANN) with multiple layers between all the layers of input and output. The DNN finds the correct statistical exploitation to turn the input in to the output, be it a perfect correlation or a non-linear relationship. The system moves via layers to measure each output's probability. DNNs can model complex interactions that are not linear. DNN architectures create geometrical models in which the object is expressed as a layered geometry of primitives. The additional layers make it easier to
Compose features from lower layers, ultimately designing complex data with fewer units than a equally low-level network.

2.3.4 Support Vector machine (SVM) [1, 6]:

A support vector machine (SVM) will be an overseen machine taking in calculation that camwood is utilized to both order Also relapse purposes. SVMs are that's only the tip of the iceberg regularly utilized over order issues Furthermore as such, this is the thing that we will concentrate on in this post. SVMs are In light of those perfect from claiming finding a hyper plane that best partitions a dataset under two classes, similarly as demonstrated in the picture further down. Support vectors are the information indicates that are closest the hyper-plane, and the purposes of the informational collection that, on the off chance that it is expelled, it would modify the situation of the separating hyper-plane. Along these lines, they can be viewed as the basic components of an informational collection.

2.3.5 Multilayer Perceptron (MLP) [1]:

A multilayer Perceptron (MLP) is a class of artificial neural network feed forward. A MLP is made up of at least three node layers: an input layer, a hidden layer, and an output layer. With the exception of the nodes input, each node is a neuron using a nonlinear activation function. For training, MLP uses a supervised learning technique called back propagation. The multiple layers as well as non-linear activation differentiate between MLP and a linear Perceptron. It might distinguish data that could not be divided linearly.

2.3.6 Linear Discriminant Analysis (LDA):

In the linear Discriminant analysis, Discriminating function analysis is a generalization of Fisher's Linear Discriminant, a method used in statistics, pattern recognition and machine learning to find a linear combination of characteristics that characterizes or separates two or more classes of objects or events. LDA is closely linked to variance analysis and regression analysis,

This also attempts to express a dependent variable as a linear combination of other charactertics or measurements. LDA is also closely linked to the main component analysis and factor analysis, both seeking linear combinations of variables that best explain the data. LDA works when continuous quantities are measured on independent variables for each observation. The equivalent technique is discriminating correspondence analysis when dealing with categorical independent variables.

CHAPTER 3

PROPOSED WORK

3.1 Proposed Description

The main aim of this dissertation work is identifying the correct medicinal leaves and classifying leaves species also it gives the correct information about what are the uses of these leaves. Today, identification and classification of unknown plant species are performed manually by expert personnel who are very few in number. The important point is to develop a system which classifies the plant leaves. This approach presents a new recognition approach based on Leaf Features Fusion and ML Classification algorithms for classifying the different types of plants.

3.2 Block Diagram of Proposed System

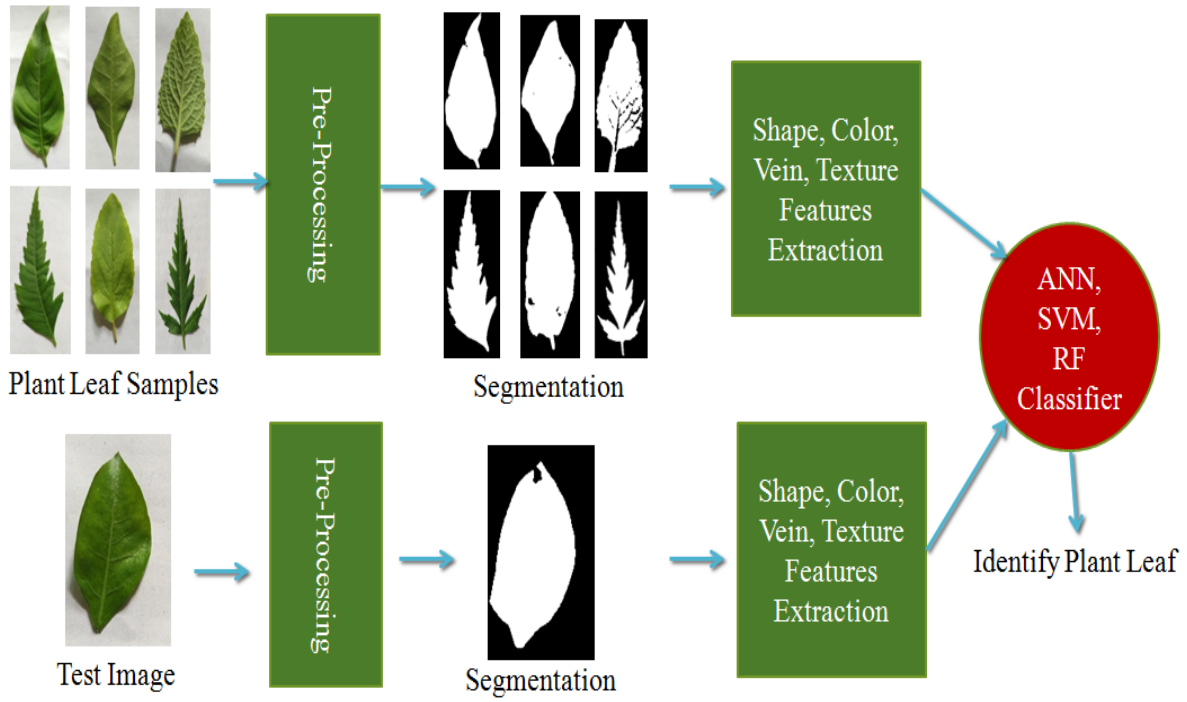


Figure 3: Block Diagram of Proposed System

In these above Proposed system, It is consists four step namely pre-processing, segmentation, feature extraction, classification. First, collect the leaves sample and performs pre-processing operation for removing noise and improvement of the image quality. After pre-processing apply segmentation on image. Among the various segmentation method Nonsubsampled Contourlet Transform (NSCT)and Contour edge detection. Contour edge detection has several advantages over other segmentation techniques is give the accurate result. Next step is the

Feature extraction and Feature selection, which is given the accurate result of classification. At last Leaf classification using ANN (Artificial Neuron Network) SVM (Support Vector Machine) and RF (Random Forest) algorithm has several advantages over other classification techniques for an accurate result.

STEPS:

Step 1: Prepare Training Dataset

a. Collect Plant Leaf Samples.
b. Acquisition of plant leaf images then Resizing the all images
c. Apply Pre-processing on each plan leaf image includes Gray conversion, median filtering and then binarization and segmentation using Contourlet transform and canny edge detector.
d. Extract Features of plant leaf such as shape, color, vein, texture etc.
e. Amalgamate the features based on combination.
f. Prepare features vector.

Step 2: Read the testing plant leaf image

Step 3: Apply Pre-processing on test image including same steps in step 1(c)

Step 4: Extract Features specified in Step 1(d) and Fuse them based on Combination.

Step 5: Train the training dataset and predicate testing image by using ANN, SVM & RF Classifier.

Step 6: Finally, identify the plan leaf.

Step 7: Stop.

3.3 Proposed Work Flow

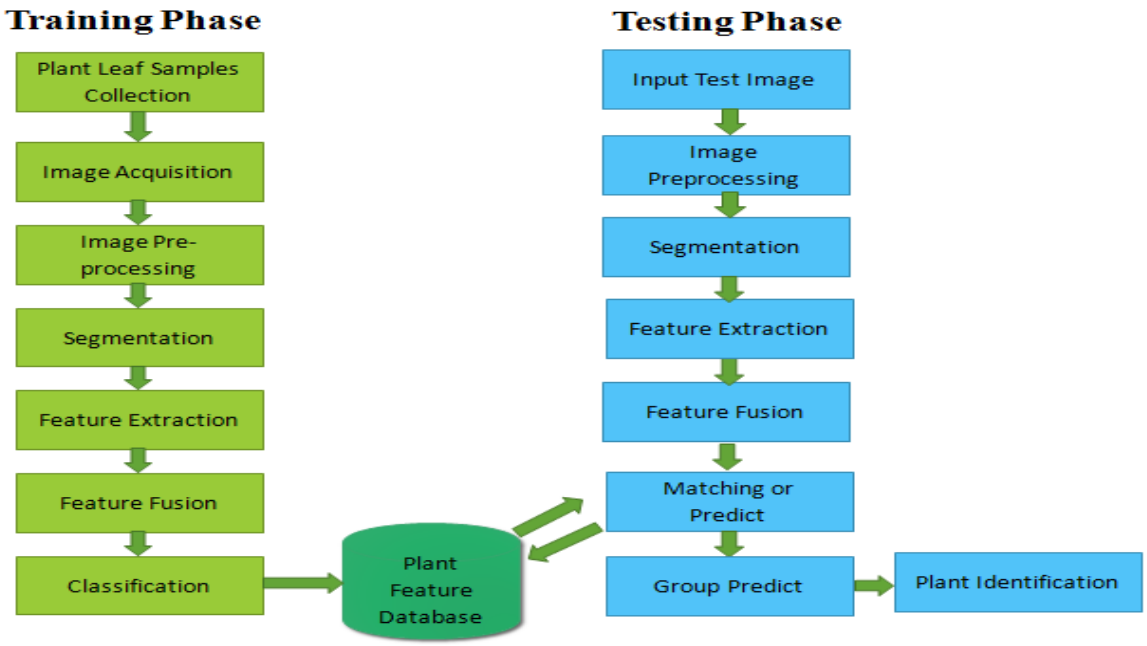


Figure 4: Proposed Work Flow

The Proposed Work Flow is as shown in the fig. 4; there were two stages operation for system: training phase and testing phase. In the Training phase, the images of leaves dataset are given as input to the system one by one. The Pre-Processing step verifies that all the images have same standard resolution. The next step is to convert color image to grayscale image and then into binary image. There were canny edge detector and Contourlet transform is used for segmentation. There were used canny edge detector; it is multistage algorithm to detect a wide range of edges. The Features Extraction stage extracts shape, texture, color and vein features from binary to color images respectively and store the values in another database. These are repeatable for all the leaves in the leaf dataset. The classifier is used for classifying the different type of leaf and trained with selected feature value from the feature dataset. There were ANN , SVM and RF methods are used for classification and the classes are separated by group of leaf. In one group all leaves are minor differentiate from each other. In the Testing Phase, first test the image then it passes through all stages and selected features are extracted from it. Then, these values are given to the input of ANN, SVM and RF Classifier. It decides whether the test sample image matches with any class of images. Then it gives perfect result of the leaf identification in the system.

3.4 Dataset of Leaf

In our proposed system 16 species with 191 leaves image are used for train dataset. There were we used Indian ayurvedic plant species for identification and classification. The dataset is downloaded from the public packages containing all the data of the Image CLEF 2013 plant retrieval task are now available (including the ground truth, an executable in order to compute new scores, the working notes and the oral and poster presentations at Image CLEF 2012 Workshop, and some additional information).In proposed system Sheet as Background (or uniform background) (42%) images are used for training and exclusively pictures of leaves in front of a white or colored uniform background produced with a scanner or a camera with a sheet. The list of Plant species are given below in table:

Table 2: Dataset of leaf

Sr. No.	Plant Species name	No. of images
1	Annick	45
2	Ardusi	6
3	Banana	6
4	Bay laurel	6
5	Buttonbush	6
6	Fresa	25

7	Heena	6
8	Maidenhair	6
9	Maple leaf	6
10	Mint	6
11	Neem	6
12	Pipal	5
13	Rabano	25
14	Tomato	25
15	Trillium	6
16	Tulsi	6

3.5 Methodology

A. Pre-Processing Technique

Firstly, select the image then resizing the images one by one in from of 128x128 picture frame. Then perform pre-processing if necessary. It is used for remove the noise by median filter and converts the RGB image into grayscale image in which the value of each pixel is a single sample representing only an amount of light. In fact, a dim shading is one in which the R, G, B planes have equal intensity, the intensity level represented as a number from decimal 0 to 255. Black is spoken to by R = G = B = 0 and white is spoken to by R = G = B = 255.We need to convert RGB to gray scale because there were we can't extract features like edges, shape, contrast, contours etc.

B. Segmentation Techniques

The segmentation is used for convert the gray image into binary image. After RGB to gray conversion of the leaves sample, the images are segmented by contour edge detection technique. There were many types of contour edge detection method like Sobel, Prewitt and Robert but in proposed method there were canny edge detector is used for image segmentation. The edge detector is used for separate two regions by contour. This method used in proposed because our aim is detect leaf shape and vain. By this technique we can easily extract vain.

i. Canny Edge Detection:

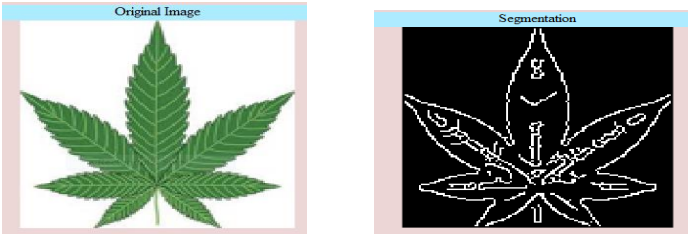
The Canny Edge Detection is an operator that uses a multi-stage algorithm to detect a wide Range of edges in images. This System is applied for both color image and gray-scale image. The Canny edge detection is a technique to extract some useful structural information from different visual objects and dramatically reduce the amount of data to be processed.

The general basis for edge detection includes:

1. The detection should accurately catch as many edges by detection of edge with low error rate.

2. The point of edges detect from the operator should accurately localize on the center of the edge.

3. The noise in image should not create false edges and given edge in the image should be marked once only



(a) (b)

Figure 5: (a) original image (b) Canny edge detection on image

Canny edge detection algorithm divided into five different steps:

Step 1: Apply Gaussian Filter to remove the noise for smooth the image.

Firstly, it is essential to filter out to avert false detection caused by noise because the noise to all edge detection result is easily affected by noise. This step will slightly smooth the image for reduce the noise on the edge detector by the Gaussian filter. The equation for Gaussian filter kernel of size (2k+1)*(2k+1) is given by,

$$H_{ij} = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{(i-(k+1))^2+(j-(k+1))^2}{2\sigma^2}\right) ; \ 1<= i, j<= (2k+1) \quad ...1$$

Step 2: Find the intensity gradients of the image.

The Canny algorithm uses four filters to detect horizontal, vertical and diagonal edges in the blurred image because the edges in an image May points are in a variety of direction. So, the edge gradient and direction can be calculated by,

$$G = \sqrt{G_X^2 + G_Y^2} \quad ...2$$

$$\Theta = a\tan 2(G_Y, G_X) \quad ...3$$

Where, G_X is First derivative in the X-axes direction and G_Y is First derivative in the Y-axes direction.

Step 3: Apply Non-maximum Suppression to get rid of specious response to edge detection.

This step is used for edge thinning technique and to find the largest edge. After applying step 2, the edge extracted from the gradient value is still quite blurred. So, the Non-maximum Suppression can used to suppress all the gradient values except the local maxima, which indicate locations with the sharpest change of intensity value.

Step 4: Apply double threshold to determine potential edges.

Step 5: Track edges by Hysteresis: Finalize the detection of edges by suppressing all the other edges that are week and not connected to strong edges.

ii. Contourlet Transform

The Discrete Contourlet transform (DCT) used a double filter bank structure to get the smooth contours of images. There were laplacian pyramid (LP) is first used for capture the point discontinuities, and a Double filter bank (DFB) is used for from those point discontinuities into linear structures.

The laplacian pyramid (LP) decomposition produce one band-pass image in a multidimensional signal processing, that can be avoid frequency scrambling and DFB is only suitable for high frequency since it will leak the low frequency of signals in its directional sub-bands. This method is amalgamated DFB with LP, that multiscale decomposition and remove the low frequency. Image signals therefore pass through LP sub-bands to receive Band-pass signal and pass through DFB to capture image directional information. This double filter bank structure combining LP and DFB is also referred to as the Pyramid Directional Filter Bank (PDFB),and this transformation approximates the original image by using basic contour, so it is also referred to as the discrete Contourlet transform.

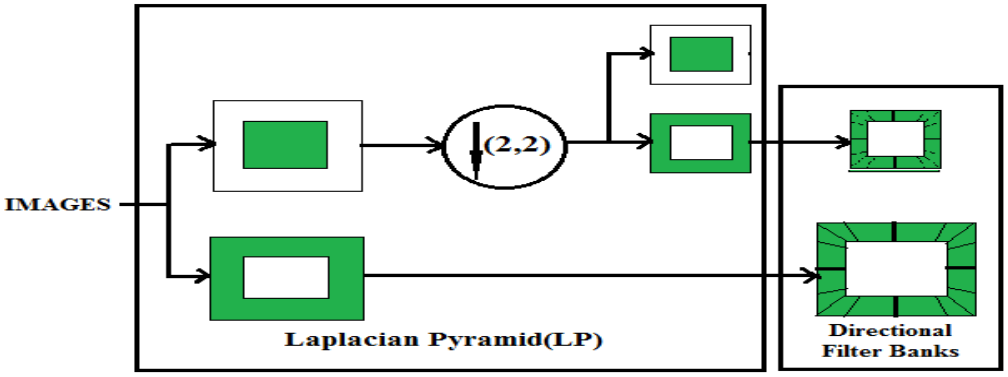


Figure 6: Discrete Contourlet Transform

Nonsubsampled Contourlet Transform:

The nonsubsampled contourlet transform (NSCT) was developed mainly because the contourlet transform is not invariant in shifts.The reason for this lies in the upsampling and downsamplig of the Laplacian Pyramid as well as the directional filter banks.The Laplacian Pyramid(LP)was replaced with a nonsubsampled pyramid structure to retain the multiscale property and a non-subsampled directional filter bank(DFB) for directionality to retain the directional and multi-scale properties of the transformation.

The first major difference is that both processes remove upsampling and downsampling.Instead the filters are upsampled in both the Laplacian Pyramid and the banks of the directional filter.

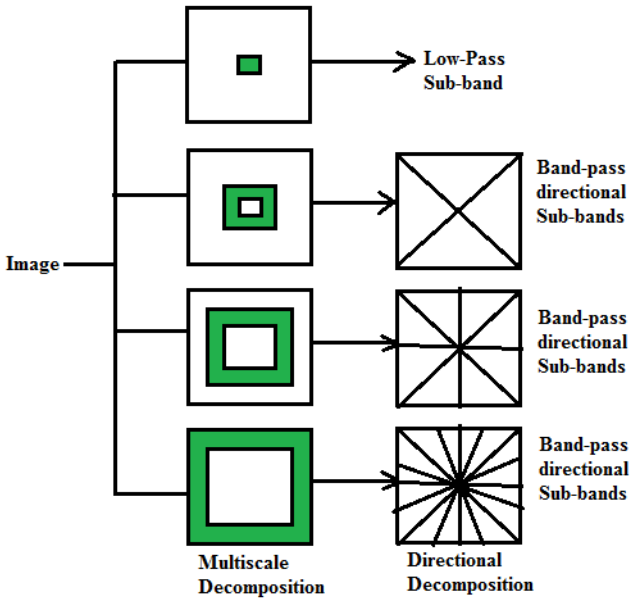
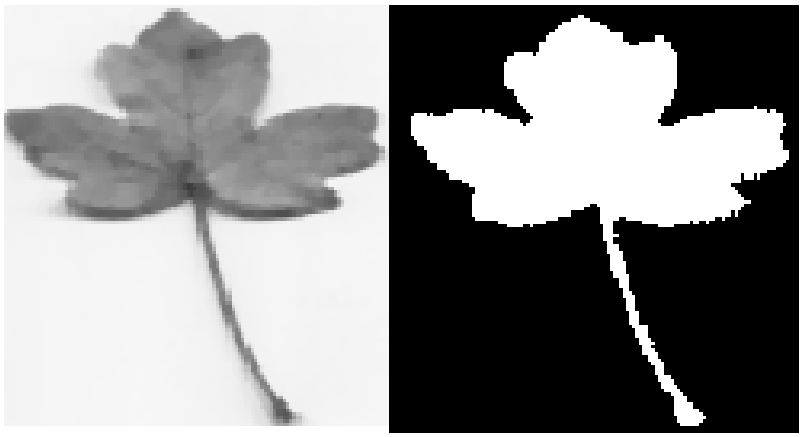


Figure 7: Nonsubsampled Contourlet Transform

While this mitigates the issue of shift invariance, a new issue with aliasing and the directional filter bank is now present. There is potential for aliasing and loss in resolution processing the coarser pyramid levels. However, this problem is avoided by sampling the direction filter bank (DFB) filters as was done with the pyramid filter bank filters.

The next issue with this transform is the filter design for both filter banks. With this transform, there were some properties they wanted such as: perfect reconstruction, a sharp frequency response, easy implementation, and linear phase filters. These features were implemented by removing the tight frame requirement first and then using a mapping to design the filters and then a ladder type structure was implemented. These changes lead to a transformation that is not only effective, but performs well when denouncing and enhancing images when compared to other similar and sometimes more advanced transformations.



(a) (b)

Figure 8: (a) 128x128 Size of Gray scale image (b) 128x128 Segmented image using Contourlet Transform

C. Feature Extraction Methods

i. Shape Features:

Table 3: List of Shape Features used in proposed method

Area	The Area Specifies the actual number of pixels in the region.
Minor axis length & Major axis Length	That specifies the length (in pixels) of the major axis and minor axis of the ellipse that has the same normalized second central moments as the region.
Perimeter	That specifies the distance between each adjoining pair of pixels around the boundary of the region.
Perimeter Ratio	P/L+W, where P is perimeter, L is length of image and w is width of image
Centroid	Calculate Centroid for connected components in the image.

ii.Texture Features: From the Gray Level Co-occurrence Matrix (GLCM) following features are calculated. Where,

X is the number of gray levels used.

Y is the probability distribution in the GLCM.

Energy measures the orderliness how regularly and orderly pixels are in window.

$$\text{Energy} = \sum_{i=0}^{X-1} \sum_{j=0}^{Y-1} X_{ij}^2 \quad \ldots 1$$

Contrast is the difference in color and light between parts of an image.

$$\text{Contrast} = \sum_{i=0}^{X-1} \sum_{j=0}^{X-1} (i-j)^2 Y_{ij} \quad \ldots 2$$

Correlation is a single number that describes the degree of relationship between two Pixels.

$$\text{Correlation} = \frac{\mathbf{1}}{\boldsymbol{\upsilon_x \upsilon_y}} \sum_{i=0}^{X-1} \sum_{j=0}^{X-1} [ijY_{ij} - \mu_x \mu_y] \quad \ldots 3$$

Entropy is the amount of information which must be coded for by a compression algorithm.

$$\text{Entropy} = \sum_{i=0}^{X-1} \sum_{j=0}^{X-1} Y_{ij} \log Y_{ij} \quad \ldots 4$$

iii.Color Features: Color moment are calculated by its means, Standard Deviation, Skewness and Kurtosis those equations are describe below and the first color moment can be interpreted as the image's average color and can be calculated by,

$$\text{Mean } X_I = \frac{\sum_{j=1}^{MN} X_{i,j}}{MN} \quad \ldots 1$$

The second moment of color is the standard deviation obtained by taking the square root of the color distribution variance.

$$\text{Standard Deviation } v_i = \sqrt{\frac{1}{MN}\sum_{j=1}^{MN}\frac{(X_{i,j}-X_i)^2}{(MN-1)}} \quad \ldots 2$$

The third moment in color is the skewness. It measures how asymmetrical the color distribution is and therefore provides information on the color distribution shape.

$$\text{Skewness } S_i = \sqrt[3]{\frac{1}{MN}\sum_{j=1}^{MN}(X_{i,j}-X_i)^3} \quad \ldots 3$$

Kurtosis is the fourth color moment, and it provides information about the color distribution shape, similar to skewness. More specifically, in comparison with normal distribution, kurtosis is a measure of how flat or high the distribution is,

$$\text{Kurtosis } K_i = \sqrt[4]{\frac{1}{MN}\sum_{j=1}^{MN}(X_{i,j}-X_i)^4} \quad \ldots 4$$

Where, $X_{i,j}$ is the value of image pixel j of color channel i. X_I is the mean for each channel i. v_i is the Standard Deviation, S_i is Skewness and K_i is Kurtosis for each Channel.

D. Classification Techniques

The Classification Technique is used for classify the leaves Features. It is used to separate the Features and make a group of leaves features which is similar to each other. Then, find the correct ayurvedic plant species.

There were many classification techniques for classifying but in proposed RF (Random Forest), SVM (Support Vector Machine) and ANN (Artificial Neural Network) classifier are used for classification.

i. Artificial Neural Network (ANN) Classifier:

A neural network is a neuron network or circuit, or an artificial neural network consisting of artificial neurons or nodes in a modern sense. Those receive input, change its internal state (activation) by that input and produces output depending on input and activation. An artificial neuron mimics the working of a biophysical neuron with inputs and outputs, but is not a biological neuron model. By connecting the output of certain neurons to the input of other neurons, the network forms a directed, weighted graph. The weights as well as the functions that calculate the activation can be modified by a learning process that is governed by a rule of learning.

An artificial neural network is an interconnected node group inspired by a brain neuron simplification. Each circular node here represents an artificial neuron and an arrow represents a connection between one artificial neuron's output and another's input.

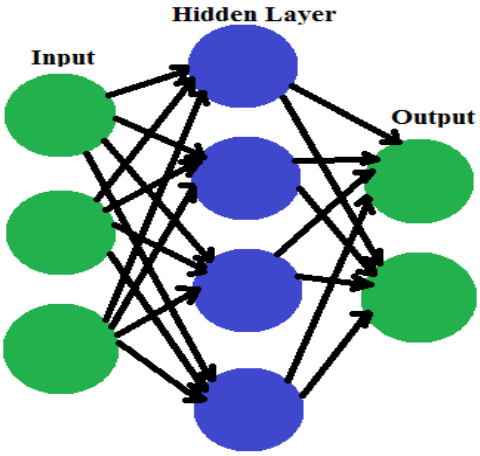


Figure 9: ANN classifier model

For our proposed system neural network training is shown in below figure. There were 16 hidden nodes are created for training and 29 Features are gives as input node.

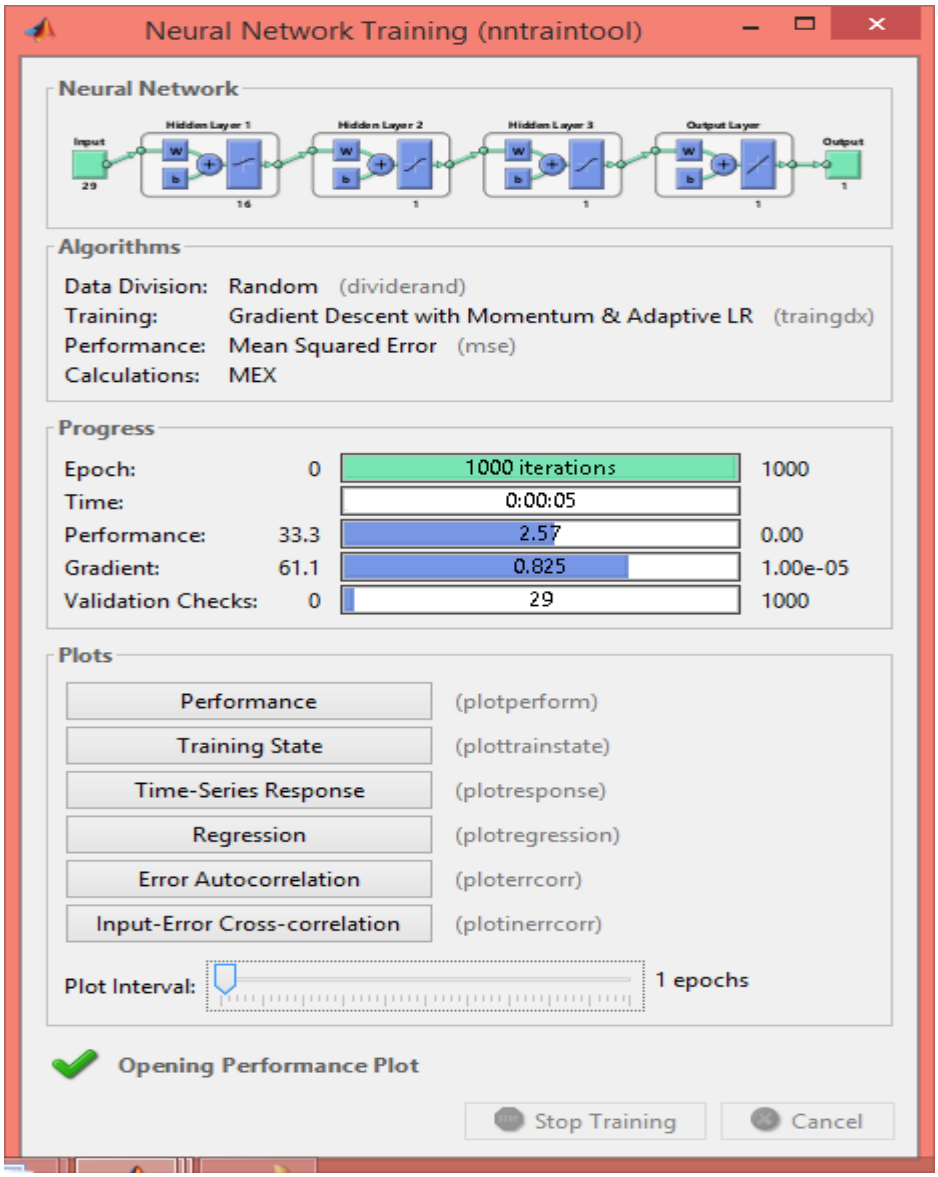


Figure 10: ANN training Tool

ii. Support Vector Machine (SVM) Classifier:

A support vector machine (SVM) will be an overseen machine taking in calculation that camwood is utilized to both order Also relapse purposes. SVMs are that's only the tip of the iceberg regularly utilized over order issues Furthermore as such, this is the thing that we will concentrate on in this post. SVMs are In light of those perfect from claiming finding a hyper plane that best partitions a dataset under two classes, similarly as demonstrated in the picture further down.

Support Vectors

Support vectors are the information indicates that are closest the hyper-plane, and the purposes of the informational collection that, on the off chance that it is expelled, it would modify the situation of the separating hyper-plane. Along these lines, they can be viewed as the basic components of an informational collection.

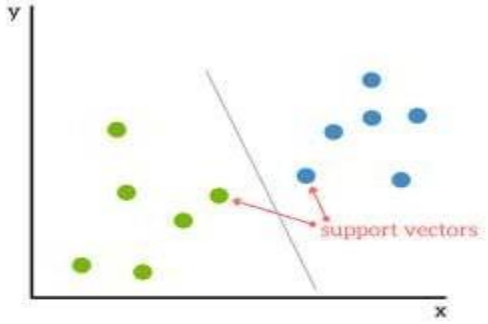


Figure 11: SVM classifier using hyper plane

What is a hyperplane?

As a basic example, to a arrangement undertaking for best two Characteristics (like those picture above), you might consider a hyperplane Concerning illustration a accordance that directly isolates and orders an arranged about information. Naturally, the further from the hyperplane our information focuses lie, that’s only the tip of the iceberg certain we need aid that they bring been effectively arranged. We In need our information focuses with make Concerning illustration much out starting with those hyperplane Similarly as possible, same time at present being on the right side from claiming it. Along these lines at new trying information is included, whatever side of the hyperplane it arrives will pick those class that we consign to it.

How would we locate the right hyperplane?

Those separations the middle of the hyperplane and the nearest information point from whichever set might be known as the edge. Those objective will be should decide a hyperplane with the best could reasonably be expected edge between the hyperplane Also At whatever side of the point inside the preparation set, providing for An more stupendous opportunity of new information being ordered effectively.

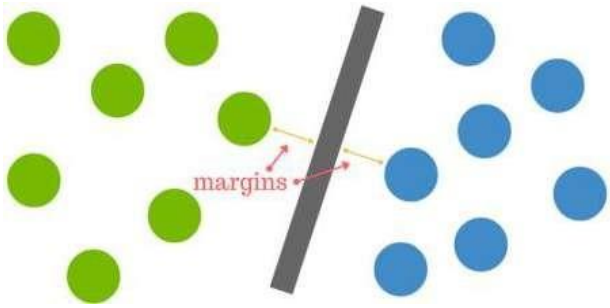


Figure 12: Margins of hyper plane

But it may happen that there is no clear hyperplane?

This may be the place it might get unreliable. Information will be infrequently ever concerning illustration clean Likewise our basic illustration over. A dataset will often gaze All the more in the cluttered balls the following which representable a linearly non distinct dataset. So as to arrange a dataset in those people over it's important to move out starting with a 2d perspective of the information with a 3d perspective. Demonstrating this may be easiest for an additional rearranged case. Envision that our two sets for hued balls over are sitting with respect to a sheet Furthermore this sheet may be lifted suddenly, propelling those balls under the air. Same time those balls would dependent upon noticeable all around, you use the sheet on independent them. This lifting of the balls addresses the mapping from claiming information under a higher size. This is known as kernelling.

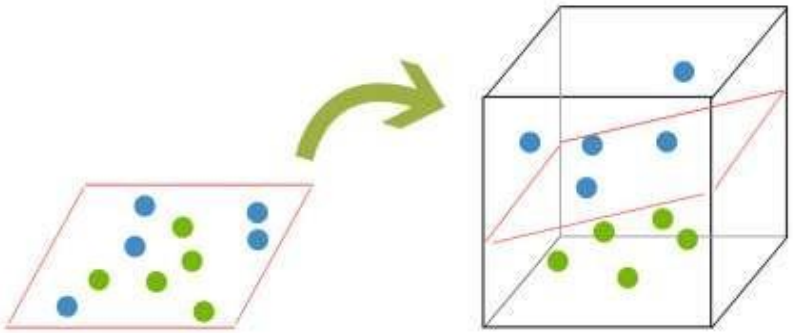


Figure 13: unreliable margin of hyper plane

A direct result we need aid presently clinched alongside three dimensions, our hyperplane might no more a chance to be a line. It must Right away make a plane likewise indicated in the illustration over. The ticket will be that the information will keep will make mapped under higher Also higher extents until a hyperplane camwood make structured to isolate it.

iii. Random Forest (RF) classifier:

Random Forest is a flexible, easy to use machine learning algorithm that generates a great result most of the time, even without hyper parameter tuning. It is also one of the most used algorithms, because its simplicity and the fact that it can be used for both classification and regression tasks. Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Random decision forests are correct for the habit of decision trees to overfit their training set.

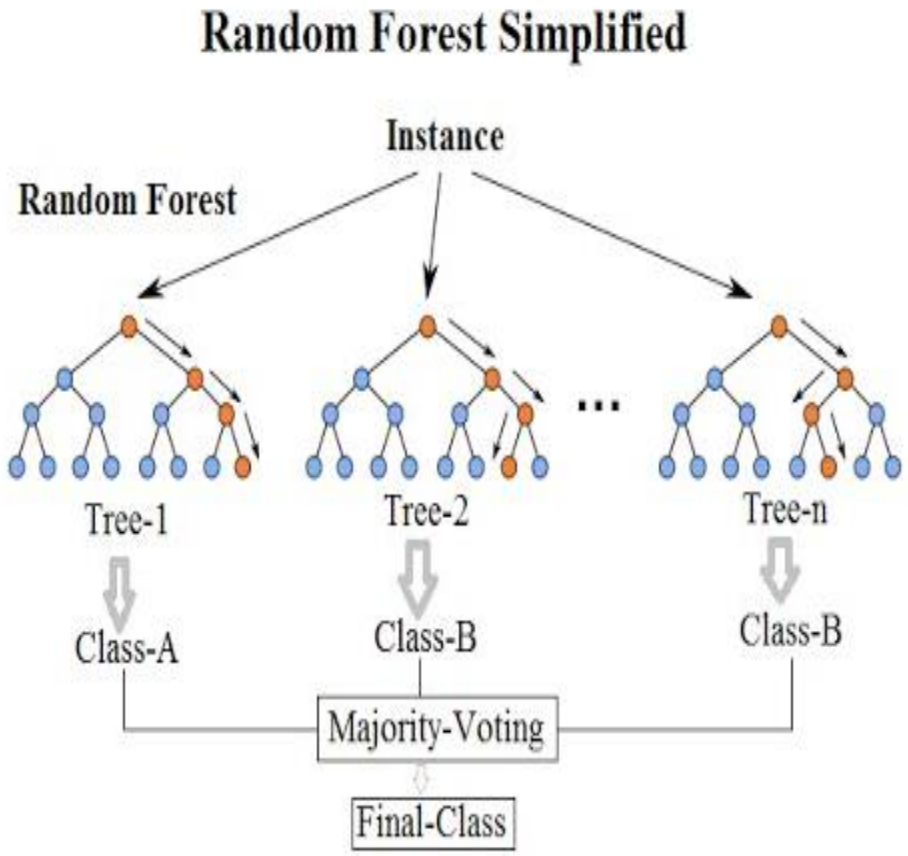


Figure 14: Random Forest Classifier [36]

In this above figure, We Extract the features from all different Classes and those are make a group or tree for classify the correct class. For i.e. if there were in class A and class B, the Class A has major voting then class B that the class A is classifying as final-class. RF is more accurate than other classifier techniques for proposed system.

3.6 Classification Measurement Analysis of System

The System Performance is calculated by its Accuracy, Precision and Recall. All Parameters are measured after the classification techniques.

The parameters are calculated by its True positive value, True negative value, false positive value and False Negative value.

Take, A = True positive value, no. of classify images which are similar to the query.

B = True negative value, no. of correct predictions when an instance is negative.

C = False positive value, no. of classify images those are dissimilar to the query.

D = False Negative value, no. of images in the database which are similar to query but not retrieved.

Accuracy is defined as; it is a representation of sampling error, a calculation of statistical bias, causing a differentiation between both the result as well as the "true" value.

$$\textbf{Accuracy} = \frac{A+B}{A+B+C+D} \quad ...1$$

Precision is defined as; this is the portion of relevant instances among the instances retrieved.

$$\textbf{Precision} = \frac{A}{A+C} \quad ...2$$

Recall is defined as; it is the fraction of the relevant instances retrieved over the total number of instances concerned.

$$\textbf{Recall} = \frac{A}{A+D} \quad ...3$$

CHAPTER 4

IMPLEMENTATION & RESULT OF SYSTEM

4.1 Requirement and Specification

4.1.1 Software Requirement

Operating System: Microsoft Windows 8 or high

System Type : 64 bit

4.1.2 Tools

- MATLAB R 2016a

4.2.1 Software Study

MATLAB

The chairman of the computer science department at the University of New Mexico was started developing MATLAB in 1970s. Matlab is trading software developed by Math works Inc. This is a logical and builds mathematical calculation of intelligent programming bundle [Inc90]. Matlab has several basic schedules that do arithmetic matrixing, plotting, etc.

Matlab Needs

Matlab has been used in many establishments as of now. It is used in the academic community and industry as part of the research. Method arrangements in Matlab are generally acquired faster than unravelling an issue from a language of programming.

Matlab is fast, considering that the Matlab centre's schedules are adjusted for specific PC models. Following the test, Matlab's speed and a program written in C were considered. Since the calculation of the back spread includes Matrix controls, the grid increase was selected for the test. While the following portion appears, Matlab was approx 2.5 times faster than a C programs both completing a network augmentation.

The artificial neural network back propagation algorithm is actualized in Matlab. The Matlab is chosen because it was about 2.5 times faster than a C program performing matrix multiplication and back propagation involves large matrix multiplication.

4.2 Experimental Results

STEPS:

1. Train the data set of leaves images which is divided into 16 classes.

16x1 struct with 5 fields

Fields	name	date	bytes	isdir	datenum
1	'Annick'	'18-Apr-201...	0	1	7.3753e+05
2	'Ardusi'	'18-Apr-201...	0	1	7.3753e+05
3	'Banana'	'18-Apr-201...	0	1	7.3753e+05
4	'Heena'	'18-Apr-201...	0	1	7.3753e+05
5	'Mint'	'18-Apr-201...	0	1	7.3753e+05
6	'Neem'	'18-Apr-201...	0	1	7.3753e+05
7	'Pipal'	'18-Apr-201...	0	1	7.3753e+05
8	'Tulsi'	'18-Apr-201...	0	1	7.3753e+05
9	'bay laurel'	'18-Apr-201...	0	1	7.3753e+05
10	'buttonbush'	'18-Apr-201...	0	1	7.3753e+05
11	'fresa'	'18-Apr-201...	0	1	7.3753e+05
12	'maidenhair'	'18-Apr-201...	0	1	7.3753e+05
13	'maple leaf'	'18-Apr-201...	0	1	7.3753e+05
14	'rabano'	'18-Apr-201...	0	1	7.3753e+05
15	'tomato'	'18-Apr-201...	0	1	7.3753e+05
16	'trillium'	'18-Apr-201...	0	1	7.3753e+05
17					

Figure 15: Prepare Dataset Of different type of ayurvedic leaves

Above figure shows about dataset of different ayurvedic plant species divided into 16 class.

2. Apply Pre-Processing like image resizing into 128x128 picture frame , gray scalling and median filter to remove the noise on original Leaves Images

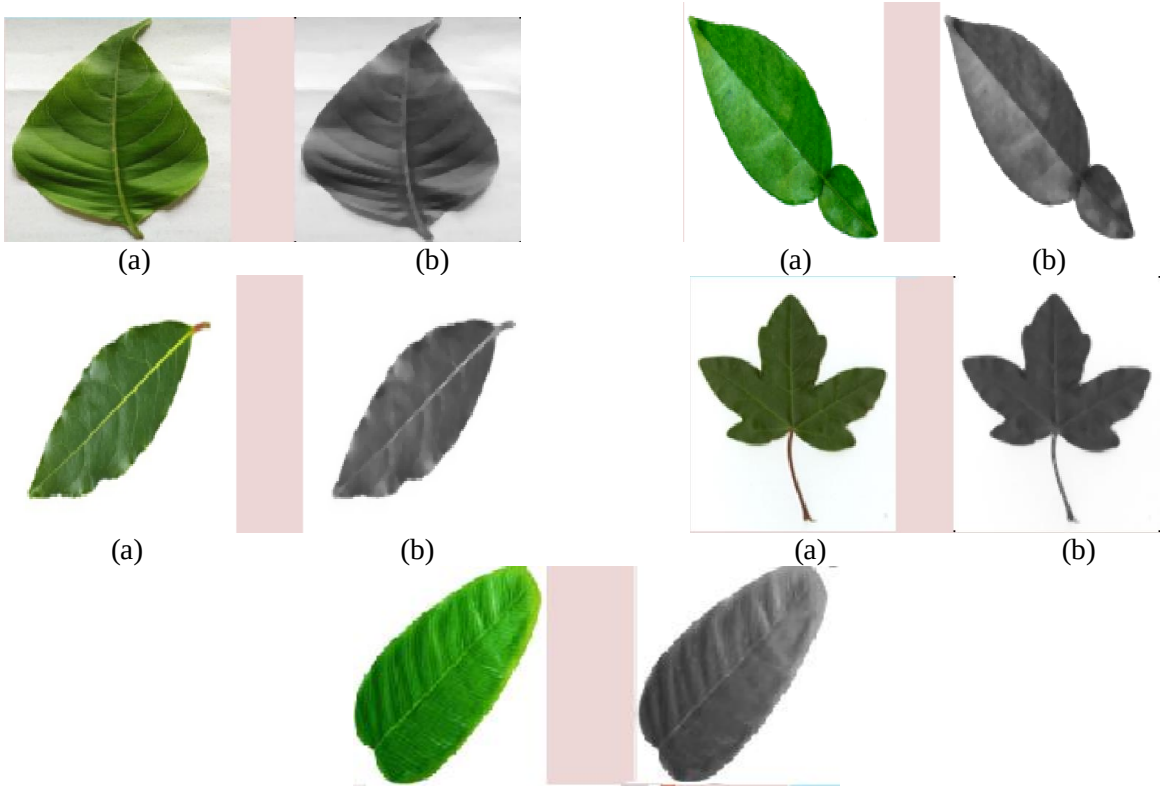


Figure 16: (a) original image (b) Pre-processing of image

Above figure is giving the result of after applying pre-processing method on original images. There were five leaves are used for pre-processing method. In this method Gray-scalling method and median filter is used for remove ‘salt and paper’ type noise in original image.

3. After pre-processing then applies segmentation method on the gray image by canny edge detector and Nonsubsampled Contourlet Transform on gray image for convert gray image into binary image.

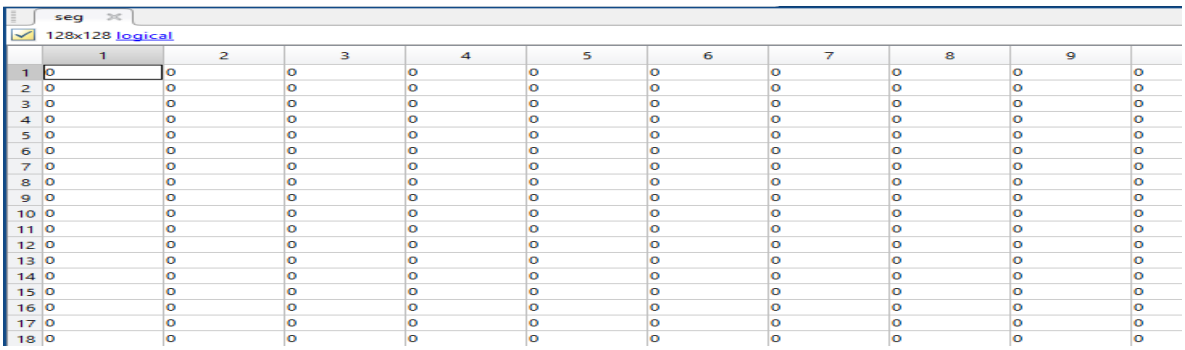


Figure 17: 128x 128 matrixes in Matlab by segmentation techniques

Above figure give the value of segmented image in form of matrix for 128 pixels value.

i. Apply Canny edge detector on gray image

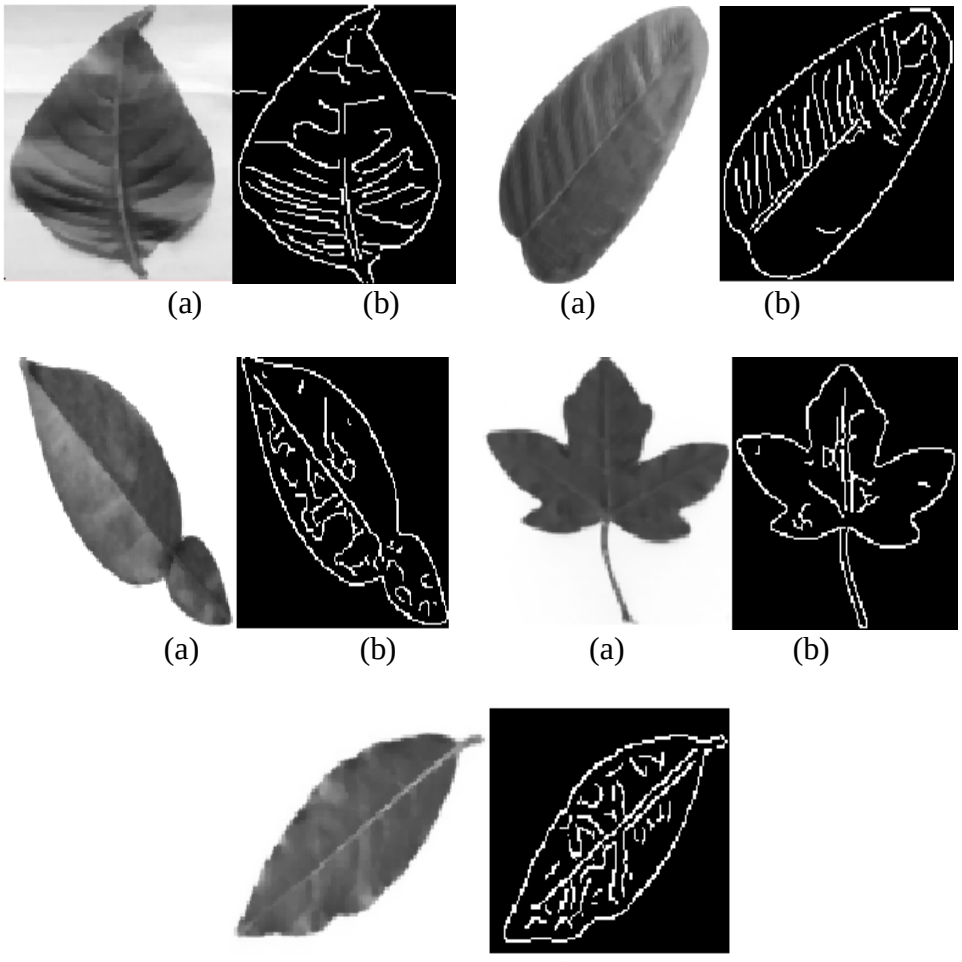


Figure 18:(a) Gray image (b) segmented image using canny edge detector

Above figure shows the output of canny edge detector from gray scale image. In this method we detect leaf shape boundary and veins on leaf gray image.

ii. Apply Contourlet transform on gray image by 2-level, 4-level and 6-level

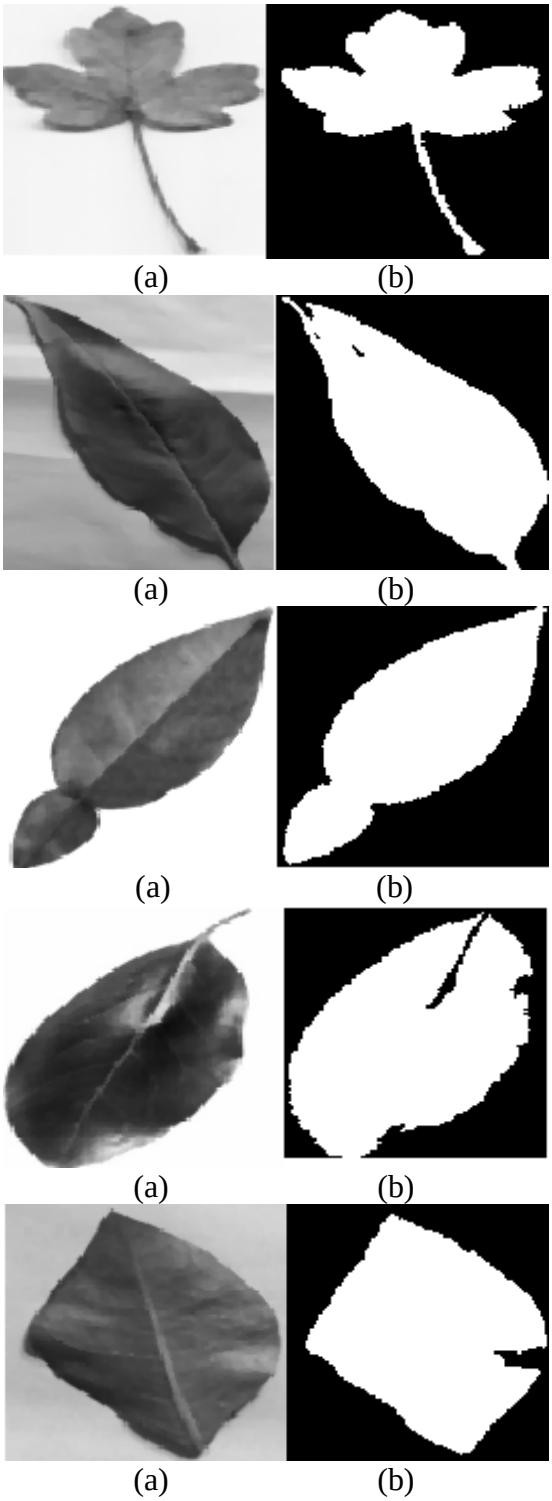


Figure 19: (a) Gray image (b) segmented image using Nonsubsampled Contourlet transform

Above figure shows the output of Nonsubsampled Contourlet transform from gray scale image. In this method we detect leaf shape. Region of leaf is converted into white pixel the intensity of white pixel is high.

4. After segmentation, Select the features like shape, texture on segmented image and also select color features on original image for accurate result

Table 4: No. Of features are used in proposed system

Features	No. Of Features
Shape	5
Texture	16
color	4
Vein	4

In proposed work we extract total 29 features from leaves image. There were 191 leaves are used for train dataset from different 16 plant species. So, the Feature vector makes 191*29 matrix for extract features at testing time.

trdataset_P

191x29 double

	1	2	3	4	5	6	7	8	9	10	11
1	64.8713	52.9678	570.3890	112.8783	75.4028	211.1340	53.9124	-1.0590	2.7873	0.0955	0.9733
2	65.3505	46.2821	512.8170	106.4868	74.4790	219.5979	42.2769	-1.6030	5.5872	0.0737	0.9602
3	65.4578	51.2465	445.6640	113.5029	73.0731	202.4489	64.5516	-0.8290	1.9308	0.1252	0.9798
4	61.6216	56.8833	404.9340	96.7589	75.6058	207.1956	59.2987	-1.0742	2.3519	0.1166	0.9782
5	64.3311	52.5351	379.8740	105.2337	61.0495	216.0348	50.3751	-1.3062	3.1296	0.0865	0.9738
6	65.0709	38.1766	400.2930	116.9990	54.8880	201.7691	67.9404	-1.0183	2.1475	0.1088	0.9848
7	72.9744	43.4248	394.1330	92.7419	60.6226	215.6720	57.6124	-1.3843	3.2373	0.1313	0.9725
8	65.2998	44.5181	476.5770	101.9283	71.7201	204.1746	63.6861	-1.0350	2.2345	0.1239	0.9794
9	63.3920	47.1532	442.1650	94.2166	71.1803	216.4850	51.9220	-1.5426	3.9643	0.1095	0.9720
10	64.9819	42.6009	441.0560	103.2695	69.1856	209.8764	57.3737	-1.2363	2.8220	0.1093	0.9778
11	63.0220	40.8068	475.6500	102.3500	72.0729	220.2025	49.1668	-1.4809	3.7794	0.0877	0.9698
12	65.9689	43.0109	479.2720	107.9391	71.7456	205.3211	64.8494	-1.0198	2.2063	0.1243	0.9795
13	66.5016	42.5852	438.2520	98.8757	69.2602	210.2973	62.9520	-1.1694	2.4874	0.1035	0.9820
14	63.5602	46.6474	360.3240	103.5163	62.9982	197.4168	88.5969	-0.9591	1.9782	0.1629	0.9862
15	64.5012	54.4904	458.2530	103.6256	74.7963	195.9371	88.6309	-0.8983	1.8618	0.2328	0.9802
16	65.2394	52.3030	396.0440	108.7136	70.8099	186.3888	94.7080	-0.7119	1.5456	0.2292	0.9848
17	63.5039	42.2578	348.8620	106.1250	56.4299	202.8935	84.9459	-1.0882	2.2459	0.1506	0.9858
18	64.4983	55.1009	471.3690	108.7803	74.4422	185.3494	95.3984	-0.7436	1.6060	0.2496	0.9832
19	64.4739	48.1574	462.8810	110.2328	68.2690	218.9560	46.7216	-1.4264	3.7480	0.0792	0.9641
20	63.7476	47.7450	472.3090	107.6709	67.5337	221.8253	40.9535	-1.5702	5.1289	0.0730	0.9513
21	63.1540	42.1845	463.0810	104.5975	71.9198	222.7898	40.0002	-1.7692	6.1311	0.0688	0.9572
22	69.8911	57.2497	377.9390	105.6379	61.0959	206.2050	64.9796	-1.0565	2.2584	0.1048	0.9819
23	62.3480	52.2536	490.3140	103.5184	69.9441	225.6473	35.8340	-1.6959	6.7316	0.0597	0.9405

Figure 20: Train the features set divided into 191x29

Above figure shows about 29 features extract from 191 leaves images which is initialized by one matrix. After then we can easily extract the features for testing one leaf which makes 1x29 matrix.

5. Apply RF, SVM and ANN classification algorithm for feature selection

i. Apply RF Classifier algorithm

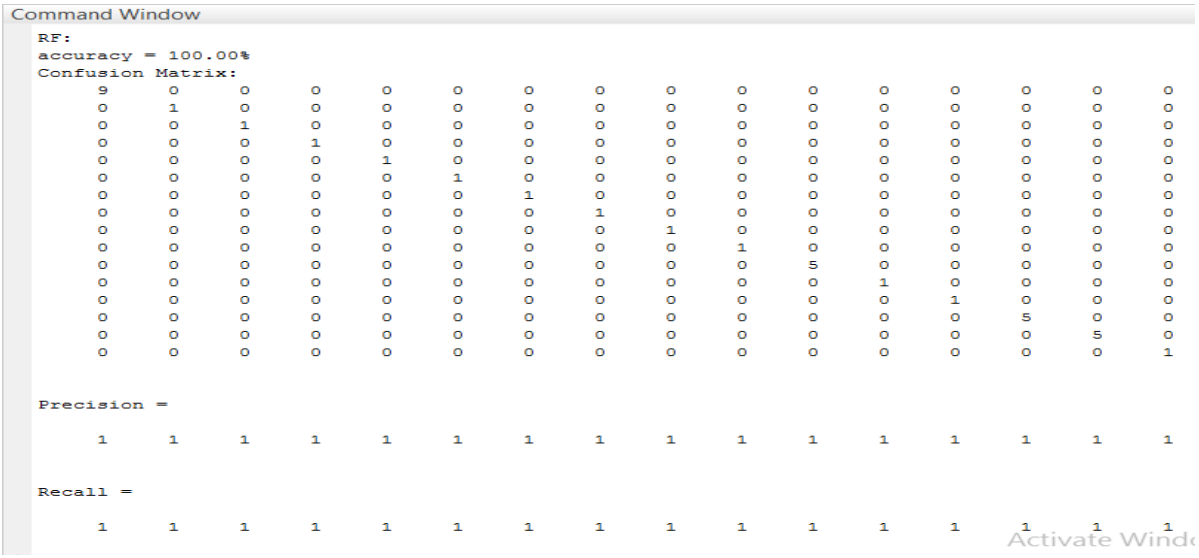


Figure 21: Calculation of RF Classifier

Above figure shows that, it is RF classifier and it generate confusion matrix and calculate its accuracy, precision and Recall by take 20% for testing and 80% for training. The accuracy by RF is 99.9999%.

ii. Apply SVM Classifier algorithm

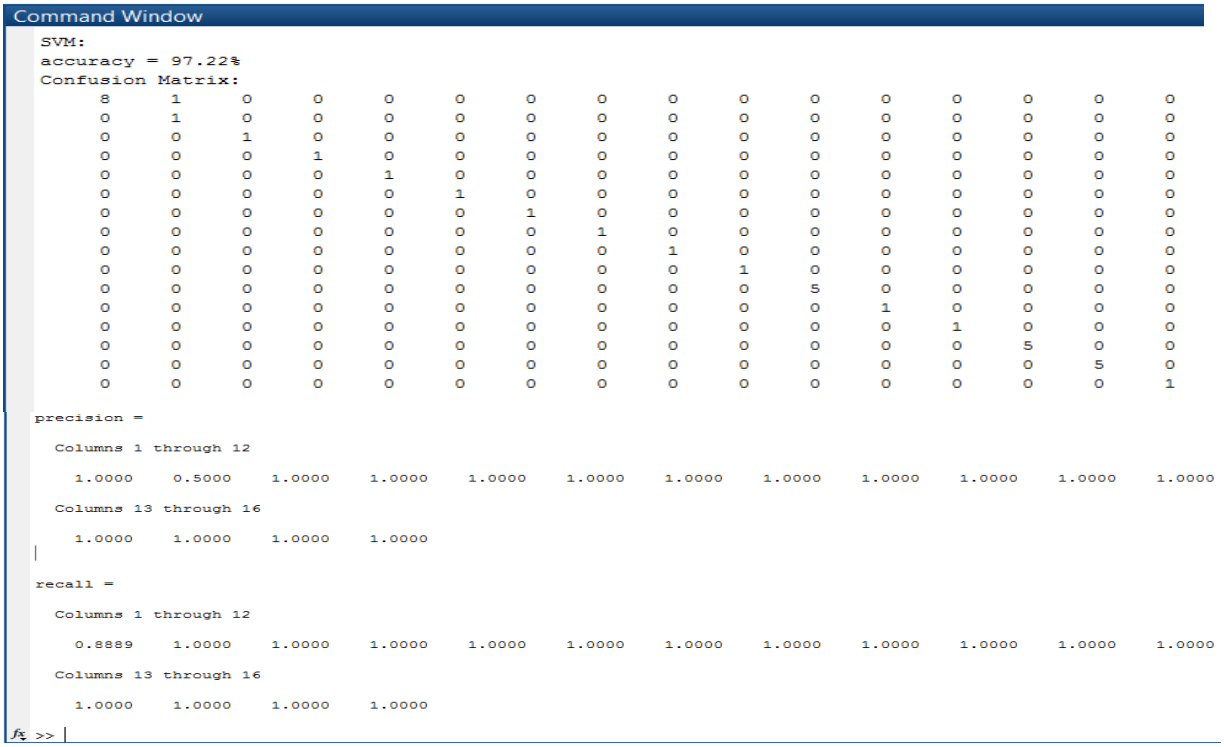


Figure 22: Calculation for SVM Classifier

Above figure shows that, It is SVM classifier and It generate confusion matrix and calculate its accuracy, precision and Recall. The accuracy by SVM is 97.2222%.

iii. Apply ANN Classifier algorithm

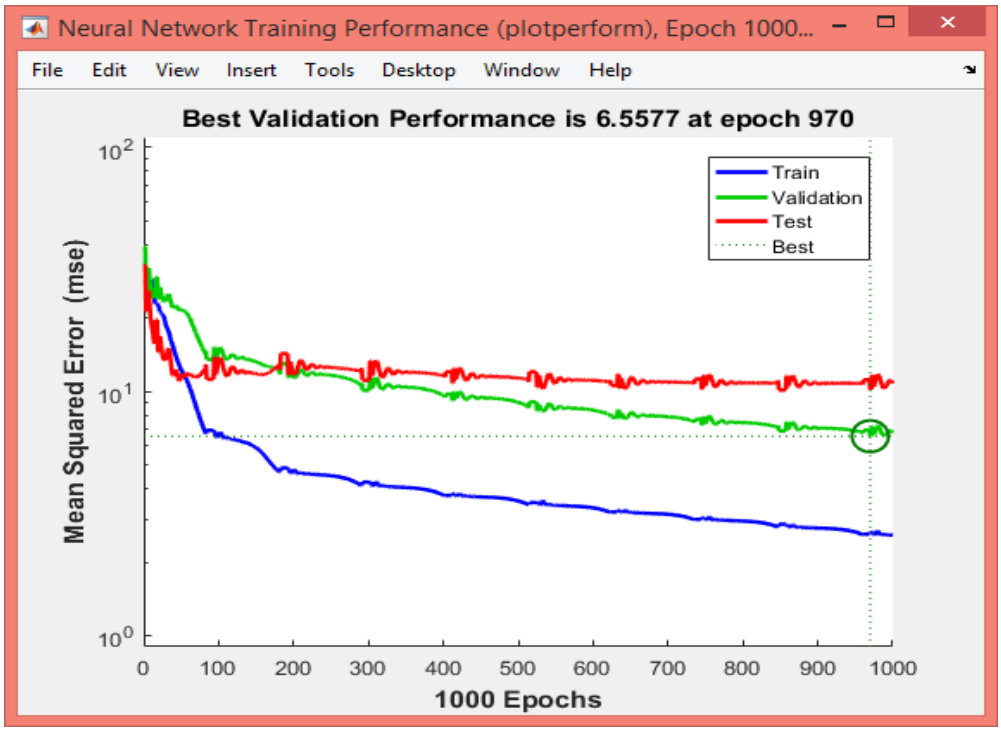


Figure 23: Performance of ANN on Matlab

In above figure, there were blue line for training, Red line for testing, Green line for validation and the dotted line is represent the best result of the performance of ANN classifier.

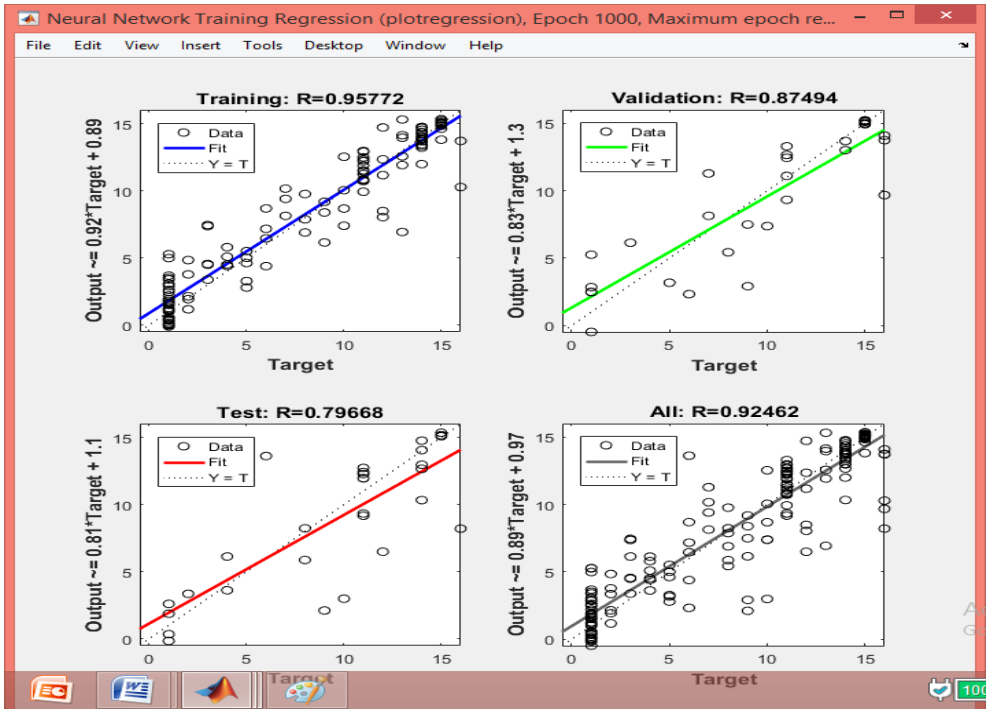


Figure 24: Regression of ANN training

Above figure shows that there were four plots for regression and compare with each other then conclude the best output of the ANN which is plot in fourth subplot with black line.

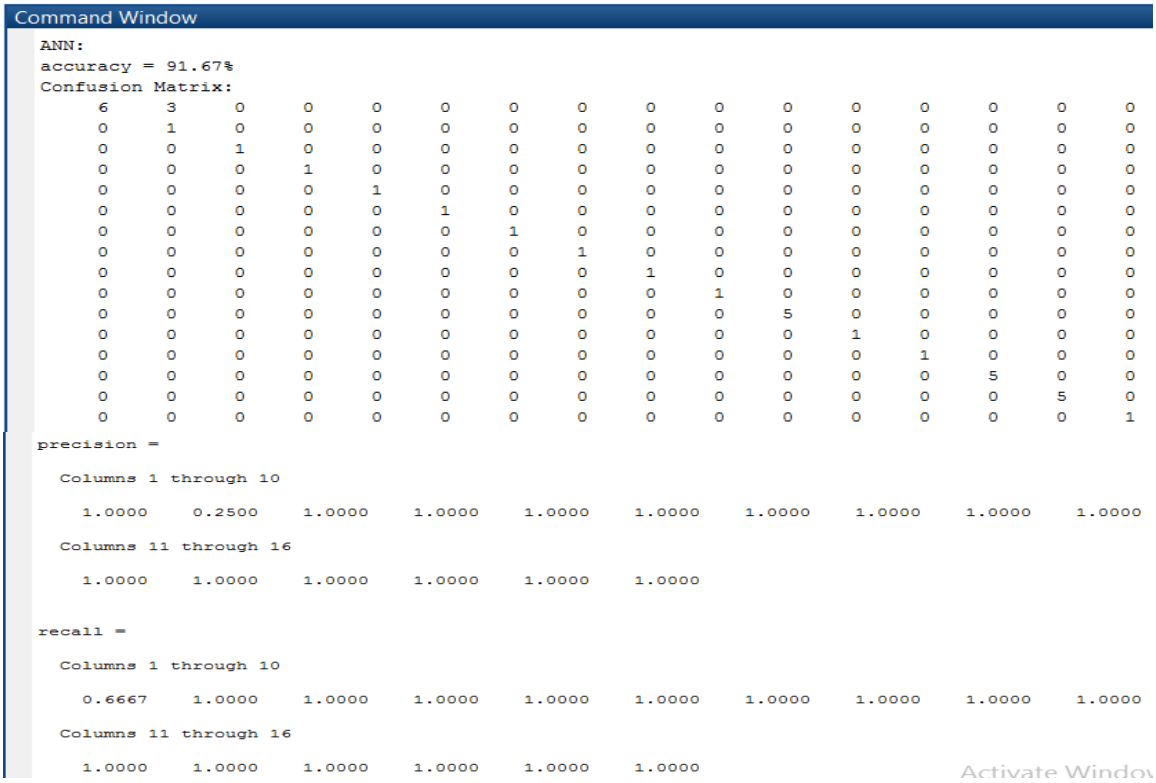


Figure 25: Calculation for ANN Classifier

Above figure shows that, It is ANN classifier and generate confusion matrix and calculate its accuracy, precision and Recall. The accuracy by ANN is 91.6667%.

4.3 GUI Design:

Here, in this below figure the GUI design for Designing of leaves classifier for ayurvedic plants using Hybrid Feature Extraction.

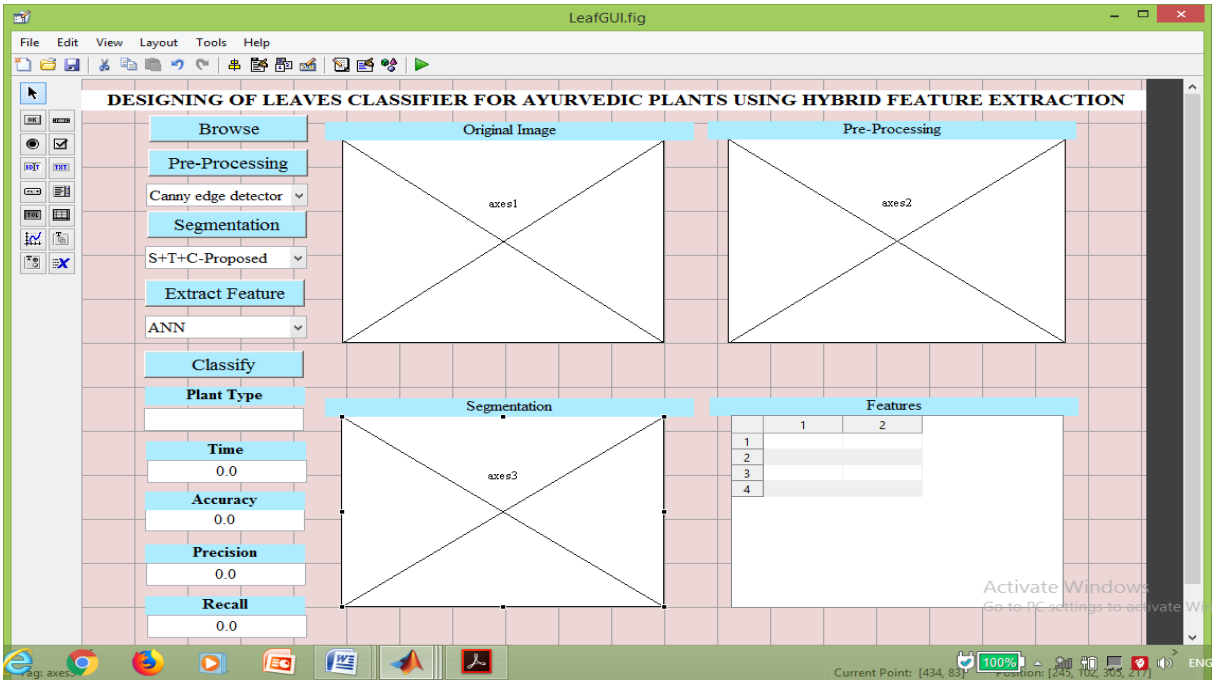


Figure 26: GUI Design for Proposed System

4.4 Performance of GUI for Testing

STEPS:

Step 1: Read the testing plant leaf image

Step 2: Apply Pre-processing on test image including gray scalling and filtering.

Step 3: Apply Segmentation method like canny edge detector to find boundary and vein on leaf image and Nonsubsampled Contourlet transform

Step 4: Apply feature extraction method for extract features from leaf image. There were shape, texture and vein features are extracted from segmented image or gray image and color features are extracted from color image

Step 5: Train the training dataset and predicate testing image by using ANN, SVM & RF Classifier

Step 6: Finally, identify the plan leaf with its accuracy, precision, Recall and timing of testing result

Step 7: Stop.

- **Canny edge detection with ANN Classifier**

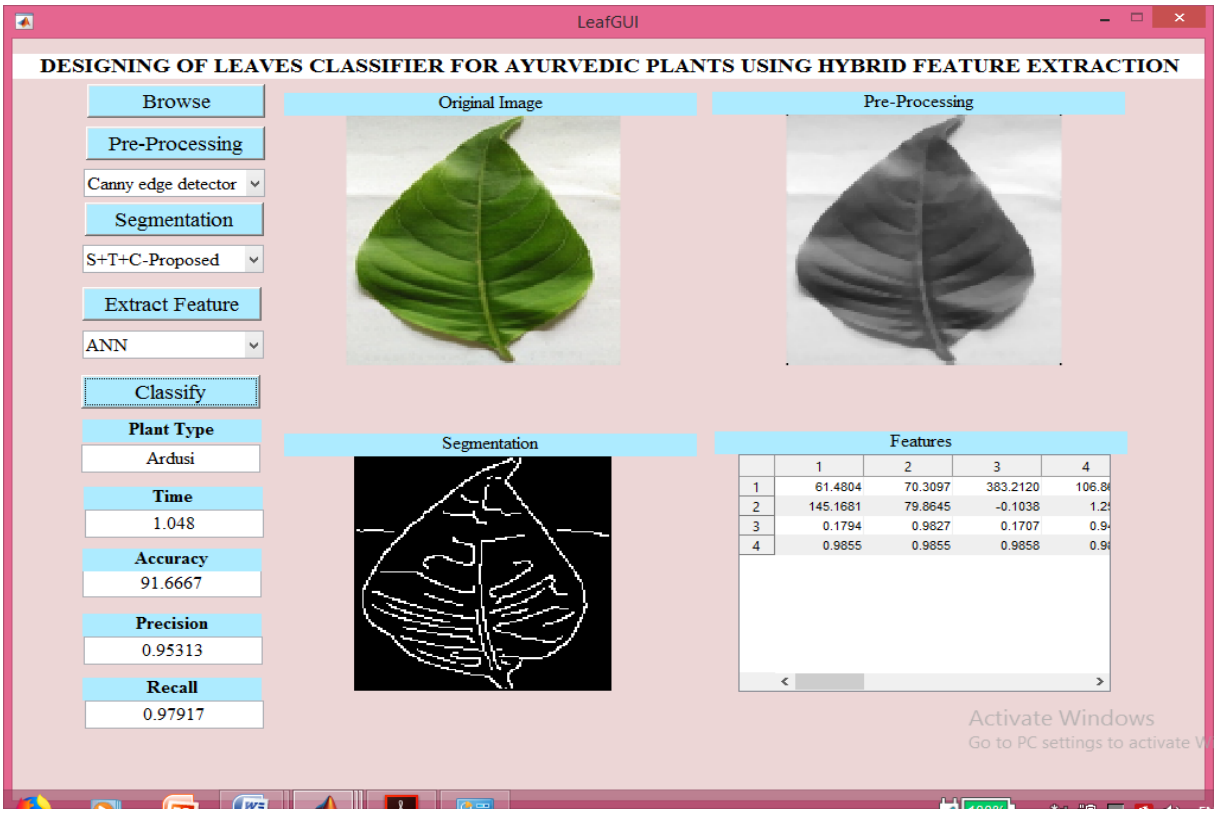


Figure 27: Canny edge detection with ANN Classifier

Above figure shows that, it is used canny edge detection with ANN classifier and identifying the type of plant species. It is give the value of its accuracy, Precision, Recall and time.

- **Canny edge detection with SVM Classifier**

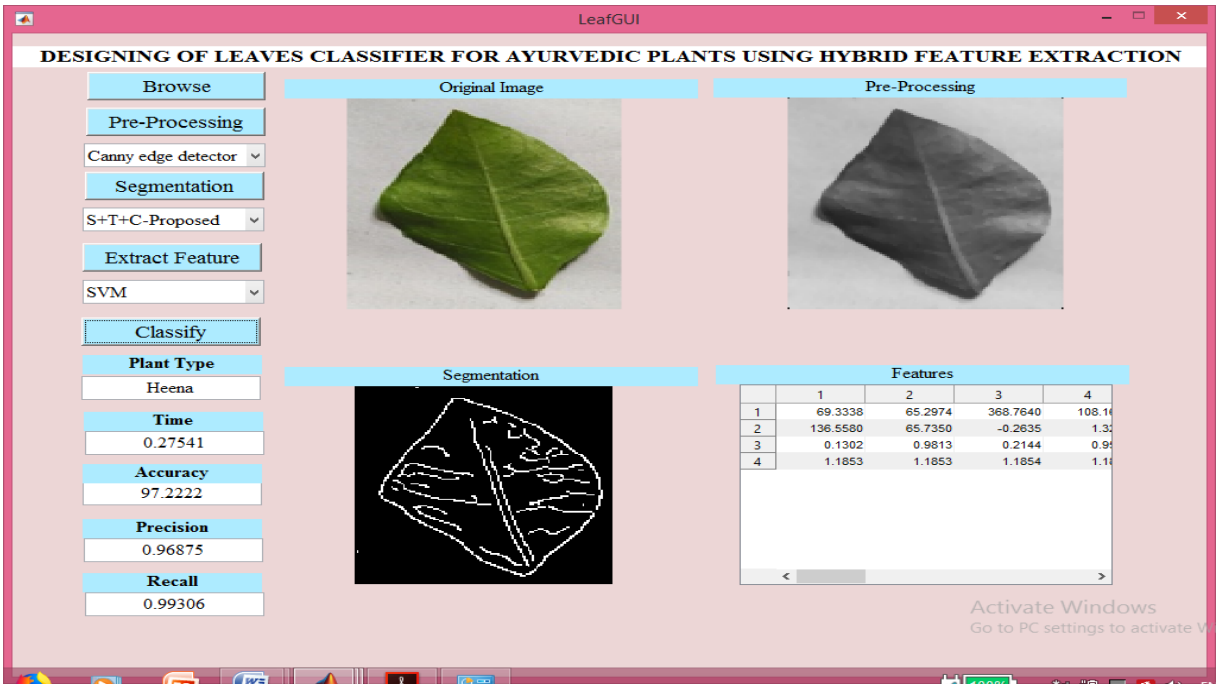


Figure 28: Canny edge detection with SVM Classifier

Above figure shows that, it is used canny edge detection with SVM classifier and identifying the type of plant species. It is give the value of its accuracy, Precision, Recall and time.

- **Canny edge detection with RF Classifier**

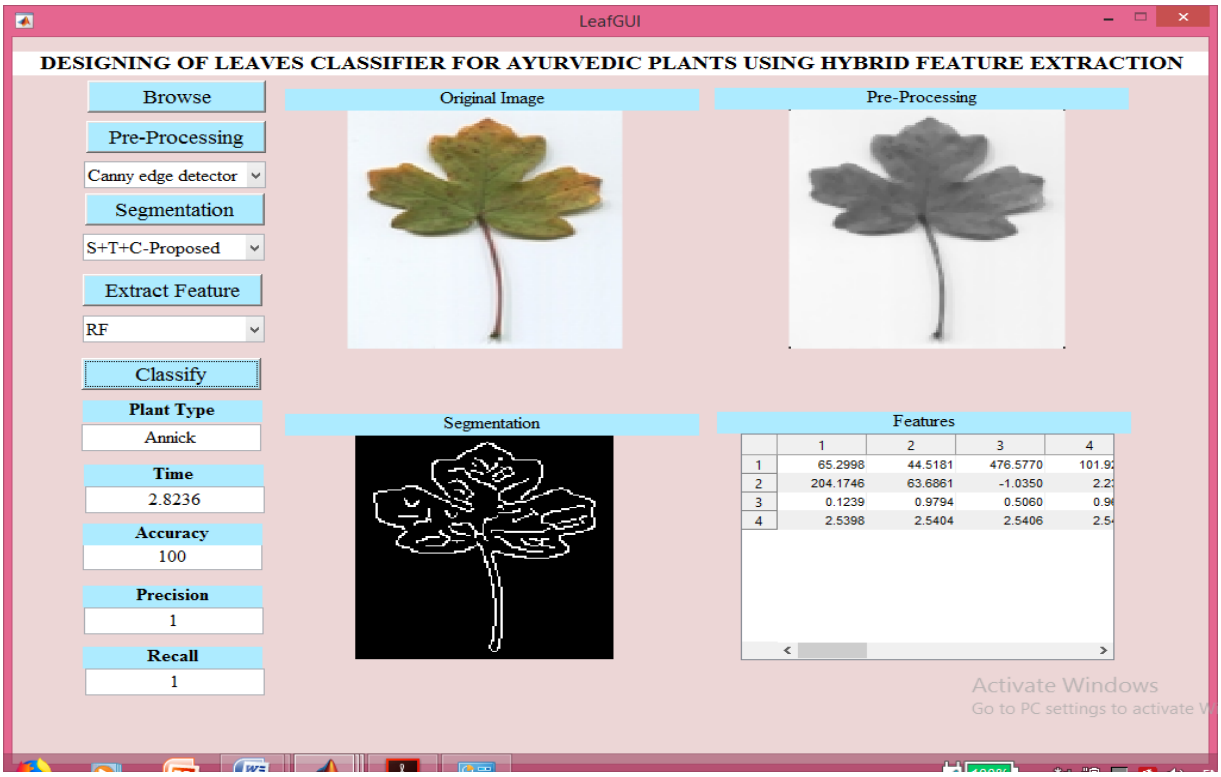


Figure 29: Canny edge detection with RF Classifier

Above figure shows that, it is used canny edge detection with RF classifier and identifying the type of plant species. It is give the value of its accuracy, Precision, Recall and time.

- **Nonsubsampled Contourlet Transform with ANN**

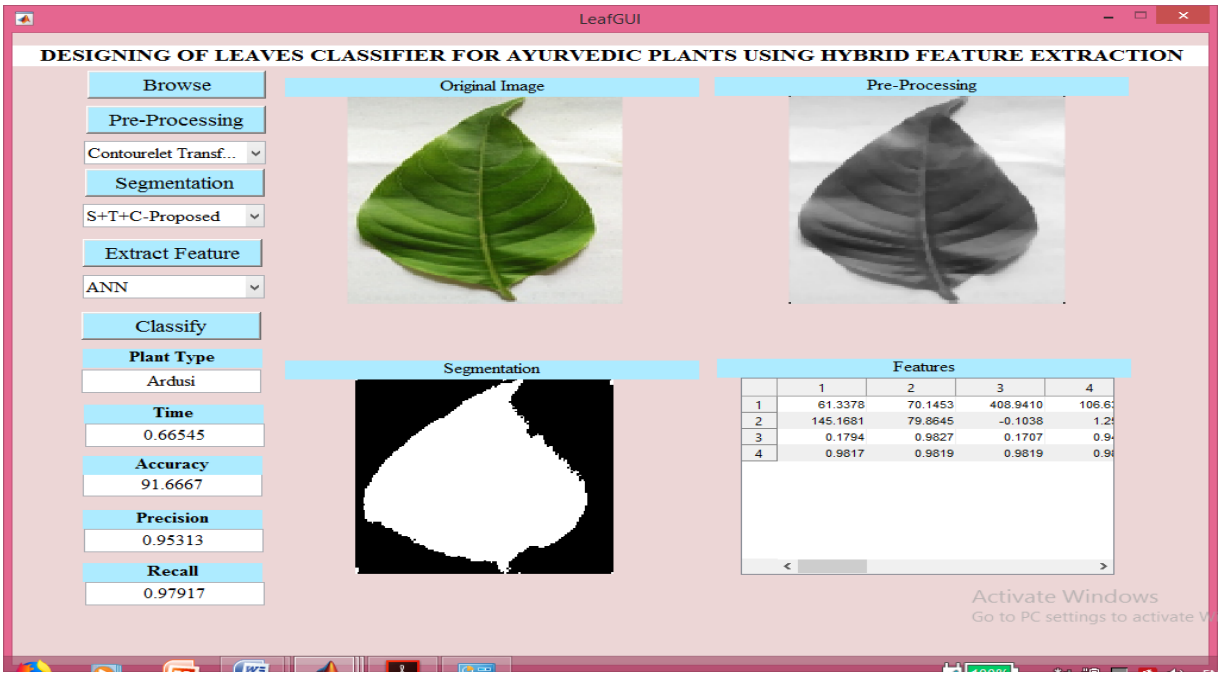


Figure 30: Nonsubsampled Contourlet Transform with ANN Classifier

Above figure shows that, it is used Nonsubsampled Contourlet Transform with ANN Classifier and identifying the type of plant species. It is give the value of its accuracy, Precision, Recall and time.

- **Nonsubsampled Contourlet Transform with SVM**

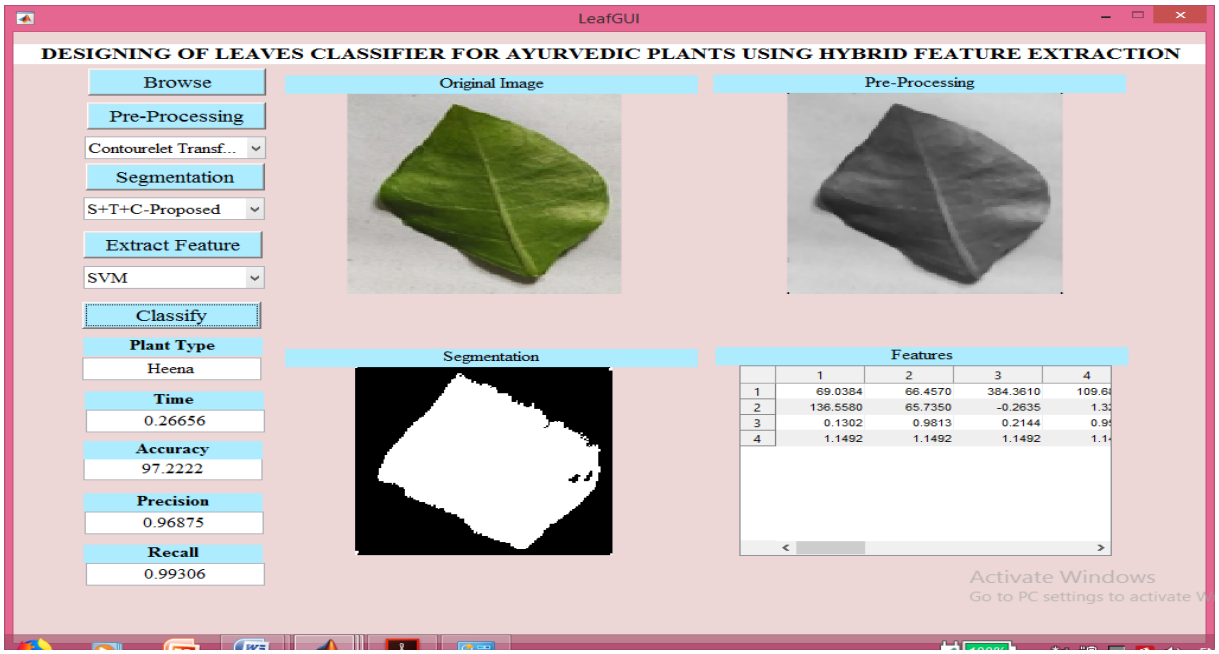


Figure 31: Nonsubsampled Contourlet Transform with SVM Classifier

Above figure shows that, it is used Nonsubsampled Contourlet Transform with SVM Classifier and identifying the type of plant species. It is give the value of its accuracy, Precision, Recall and time.

➢ **Nonsubsampled Contourlet Transform with RF**

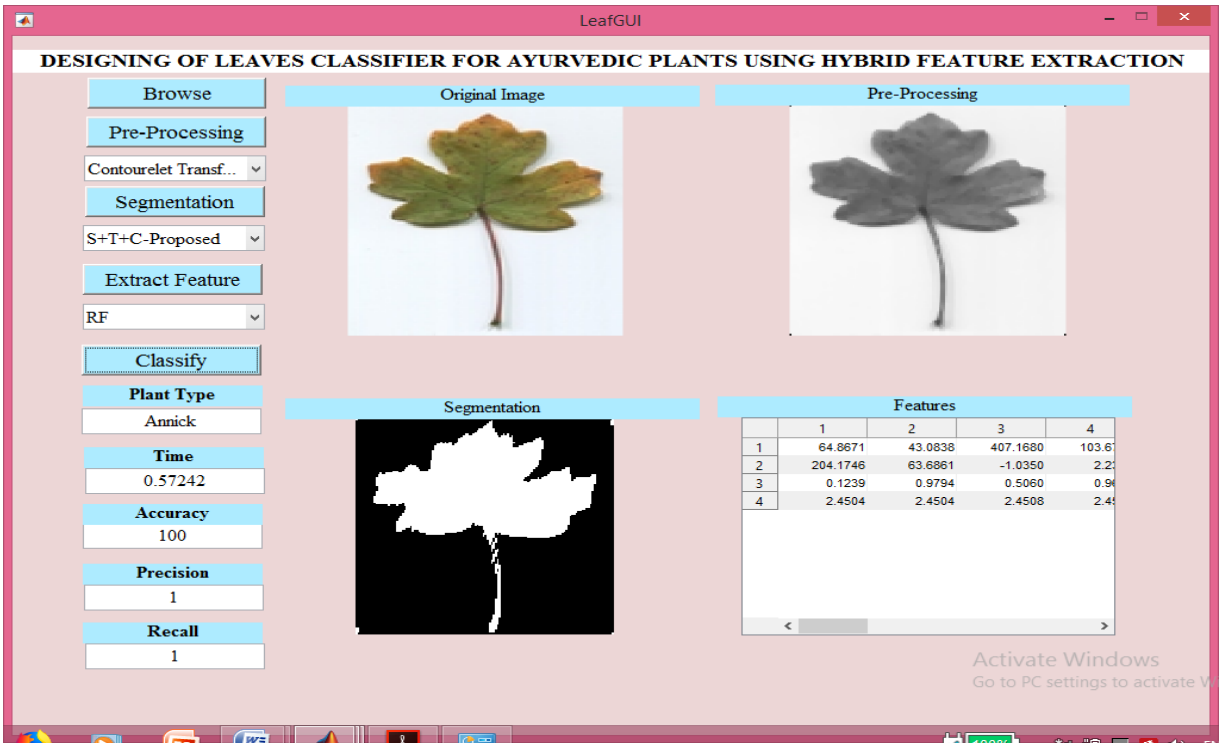


Figure 32: Nonsubsampled Contourlet Transform with RF Classifier

Above figure shows that, it is used Nonsubsampled Contourlet Transform with RF Classifier and identifying the type of plant species. It is give the value of its accuracy, Precision, Recall and time.

4.5 Comparison of Performance Analysis Of System

Table 5: Performance analysis of system with different segmentation technique and different classifier

Segmentation Technique	Classifier	Accuracy (%)	Precision	Recall	Time (sec.)
Canny Edge Detector (16 class) (191 images with 121x121 size)	ANN	91.6667	0.95313	0.97917	1.048
	SVM	97.2222	0.96875	0.99306	0.27541
	RF	99.9999	1	1	2.8236
Nonsubsampled Contourlet Transform (16 class) 191 images with 121x121 size)	ANN	91.6667	0.95313	0.97917	0.66545
	SVM	97.2222	0.96875	0.99306	0.57242
	RF	99.9999	1	1	0.26656

Above table shows that, It is parameter calculation of the proposed system which calculate accuracy, precision, Recall of the system with respect to different classifier and different segmentation techniques.

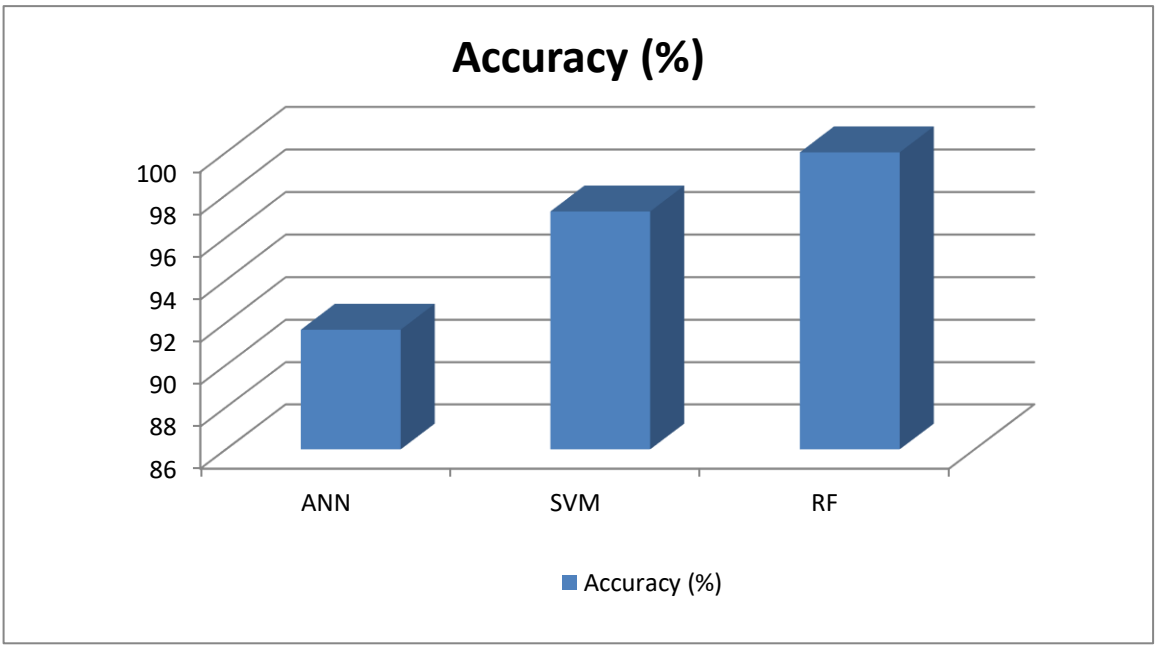


Figure 33: Graphical representation of accuracy of ANN, SVM and RF

Above graph is represented for Accuracy tends to Classification techniques which show different accuracy results of different classification techniques.

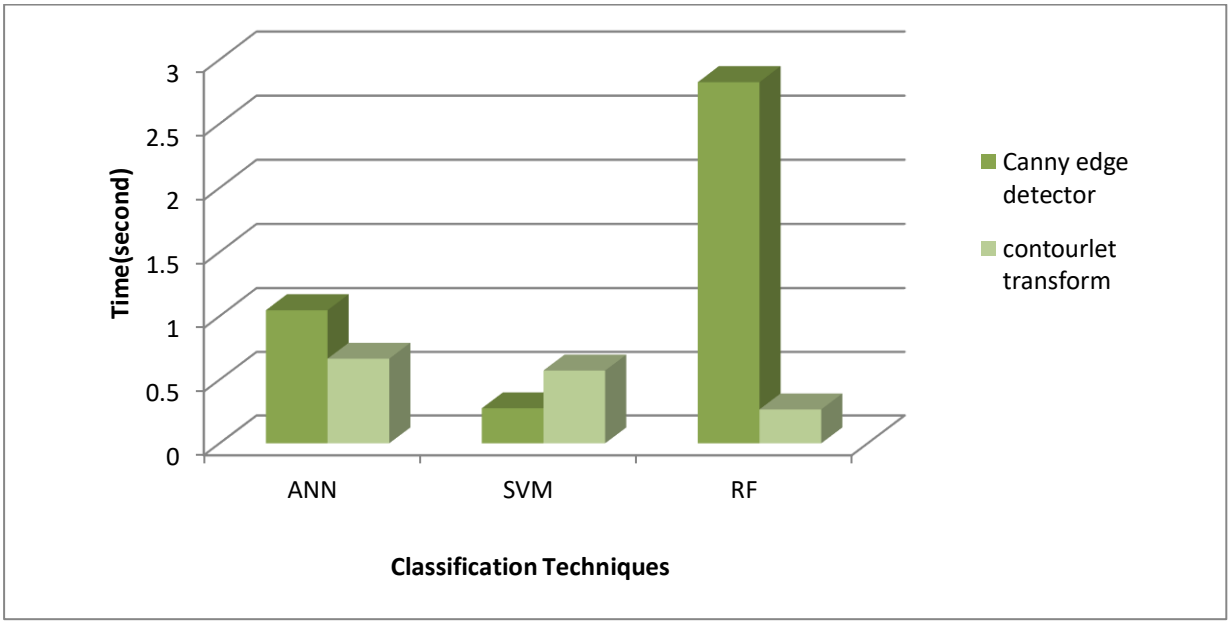


Figure 34: Processing time of canny edge detector and Nonsubsampled Contourlet Transform

Above graph is a show about timing difference of segmentation methods with respect to classification techniques.

CHAPTER 5

CONCLUSION

Here we conclude that for the segmentation the Contourlet transform as a robust method for both accuracy as well as processing time and also canny edge detector give accurate result, but the processing time is more than NSCT (Nonsubsampled Contourlet transform). By canny edge detector, there were extract vein of the leaves for accurate result. In proposed system, there were RF classifiers given 99.9999% accuracy. So, the RF is more accurate than other classifier for the proposed system and gives the accurate result for leaf identification.

REFERENCES

PAPERS

[1] Manojkumar P., Surya C. M., Varun P. Gopi, "Identification of Ayurvedic Medicinal Plants by Image Processing of Leaf Samples",2017 Third International conferences on research in computational intelligence and communication network(ICRCICN)

[2] Sigit Adinugroho, Yuita Arum Sari, "Leaves Classification Using Neural Network Based on Ensemble Features", 2018 5th International Conference on Electrical and Electronics Engineering, 978-1-5386-6392-9/18/$31.00 ©2018 IEEE

[3] Heba A. Elnemr, "Feature Selection for Texture-Based Plant Leaves Classification", IEEE#41458-ACCS'017&PEIT'017, Alexandria, Egypt

[4] Anilkumar.Muthevi, Dr.Ravibabu.Uppu, "Leaf Classification Using Completed Local Binary Pattern of Textures", 2017 IEEE 7th International Advance Computing Conference, DOI 10.1109/IACC.2017.169

[5] Anusha Rao, Dr. S.B. Kulkarni, "An Improved Technique of Plant Leaf Classification Using Hybrid Feature Modeling", International Conference on Innovative Mechanisms for Industry Applications(ICIMIA 2017),978-1-5090-5960-7/17/$31.00 ©2017 IEEE 5

[6] Adil Salman, Ashish Semwal & Upendra Bhatt, "Leaf Classification and Identification using Canny Edge Detector and SVM Classifier", International Conference on Inventive Systems and Control(ICISC-2017)

[7] Pankaj K & Dr.Thippeswamy G, "Survey on Leaf Recognization and Classification", International Conference on Innovative Mechanisms for Industry Applications(ICIMIA 2017)

[8] Arturo Oncevay-Marcos, Ronald Juarez-Chambi, Sofia Khlebnikov-Nunezand& Cesar Beltran-Castanon, "Leaf-Based Plant Identification Through Morphological Characterization in Digital Images", Springer International Publishing Switzerland 2015,G. Azzopardi and N. Petkov (Eds.): CAIP 2015, Part II, LNCS 9257, pp. 326–335, 2015,DOI: 10.1007/978-3-319- 23117-4 28

[9] Michael Rzanny, Marco Seeland, Jana Wäldchen & Patrick Mäder, "Acquiring and pre-processing leaf images for automated plant identification: understanding the trade-off between effort and information gain", Rzanny et al. Plant Methods (2017) 13:97,DOI 10.1186/s13007-017-0245-8

[10] D Venkataraman & Mangayarkarasi N, "Support Vector Machine Based Classification of Medicinal Plants Using Leaf Features", 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)

[11] Carlos Arrasco, Sofia Khlebnikov, Arturo Oncevay and Cesar Beltr´an Casta˜n´on, “Leaf Venation Enhancing for Texture Feature Extraction in a Plant Classification Task”, 2018 IEEE Latin American Conference on Computational Intelligence (LACCI) Guadalajara, Jalisco, Mexico, November 7-9, 2018

[12] Parul Mittal, Manie Kansal, Hardeep Kaur, “Combined Classifier for Plant Classification and Identification from Leaf Image based on Visual Attributes”, 2018 International Conference on Intelligent Circuits and Systems

[13] Feng Ni, Bin Wang,” INTEGRAL CONTOUR ANGLE: AN INVARIANT SHAPE DESCRIPTOR FOR CLASSIFICATION AND RETRIEVAL OF LEAF IMAGES”, 2018 25th IEEE International Conference on Image Processing (ICIP)

[14] A. López, R. Rojas, V. Trujillo, J. Cervantes, L. Rodríguez and F. García, “Detection of Compound Leaves for Plant Identification”, IEEE LATIN AMERICA TRANSACTIONS, VOL. 15, NO. 11, NOV. 2017

[15] Ionel-Bujorel Păvăloiu, Robert Ancuceanu, Claudia-Maria Enache, Andrei Vasilăţeanu, “Important Shape Features for Romanian Medicinal Herb Identification Based on Leaf Image”, The 6th IEEE International Conference on E-Health and Bioengineering - EHB 2017 , Grigore T. Popa University of Medicine and Pharmacy, Sinaia, Romania, June 22-24, 2017

[16] Adil Salman, Ashish Semwal, Upendra Bhatt, V. M Thakkar, “Leaf Classification and Identification using Canny Edge Detector and SVM Classifier”, International Conference on Inventive Systems and Control(ICISC-2017)

[17] Vijayashree.T, Dr. A. Gopal, “Leaf Identification for the Extraction of Medicinal Qualities Using Image Processing Algorithm”, 2017 International Conference on Intelligent Computing and Control (I2C2)

[18] R Rizal Isnanto, Ajub Ajulian Zahra, Patricia Julietta, “Pattern Recognition on Herbs Leaves Using Region-Based Invariants Feature Extraction”, Proc. of 2016 3rd Int. Conf. on Information Tech., Computer, and Electrical Engineering (ICITACEE), Oct 19-21st, 2016, Semarang, Indonesia

[19] Mathara Rojanamontien, Poomkawin Sihanatkathakul, Nicha Piemkaroonwong, Supanat Kamales and Ukrit Watchareeruetai, “Leaf Identification Using Apical and Basal Features”, 2016 8th International Conference on Knowledge and Smart Technology (KST)

[20] Xiaohan Yu, Shengwu Xiong, Yongsheng Gao, Yang Zhao, Xiaohui Yuana, “Multiscale Crossing Representation Using Combined Feature of Contour and Venation for Leaf Image Identification”, 2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA)

[21] Mathew. M. and Gopi V. P., “Transform based bleeding detection technique for endoscopic images”, 2nd International Conference on Electronics and Communication Systems (ICECS), pp. 1730-1734, Feb. 2015.

[22] M. Ramaraj, S. Raghavan, V. Raghunath and Wahid A. Khan, “Histogram variance controlled bleeding detectors for wireless capsule endoscopic images”, J. Med. Imaging Health Inf., vol. 4, no. 4, 2014.

[23] Tonmoy Ghosh, Syed Khairul Bashar, Md. Samiul Alam, Khan Wahid and Shaikh Anowarul Fattah, “A statistical feature based novel method to detect bleeding in wireless capsule endoscopic images”, 3rd International Conference on Informatics, Electronics & Vision (ICIEV), pp. 1-4, 2014.

[24] Tonmoy Ghosh, Shaikh Anowarul Fattah and Khan Arif Wahid, “Automatic bleeding detection in wireless capsule endoscopy based on RGB pixel intensity ratio”, International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT), pp. 1-4, 2014.

[25] Yanan Fu, Wei Zhang, Mrinal Mandal and Max Q.H.Meng, “Computeraided bleeding detection in WCE video”, IEEE journal of biomedical and health informatics, vol. 18, no. 2, Mar. 2014.

[26] A. López, R. Rojas, V. Trujillo, J. Cervantes, L. Rodríguez and F. García, “Detection of Compound Leaves for Plant Identification”, IEEE LATIN AMERICA TRANSACTIONS, VOL. 15, NO. 11, NOV. 2017

[27] Ionel-Bujorel Păvăloiu, Robert Ancuceanu, Claudia-Maria Enache, Andrei Vasilăţeanu, “Important Shape Features for Romanian Medicinal Herb Identification Based on Leaf Image”, The 6th IEEE International Conference on E-Health and Bioengineering - EHB 2017 , Grigore T. Popa University of Medicine and Pharmacy, Sinaia, Romania, June 22-24, 2017

[28] Adil Salman, Ashish Semwal, Upendra Bhatt, V. M Thakkar, “Leaf Classification and Identification using Canny Edge Detector and SVM Classifier”, International Conference on Inventive Systems and Control(ICISC-2017)

[29] Vijayashree.T, Dr. A. Gopal, “Leaf Identification for the Extraction of Medicinal Qualities Using Image Processing Algorithm”, 2017 International Conference on Intelligent Computing and Control (I2C2)

[30] Heba F. Eid, Aboul Ella Hassanien and Tai-Hoon Kim, “Leaf plant identification system based on hidden na¨ıve bays classifier”, 2015 4th International Conference on Advanced Information Technology and Sensor Application

WEBSITES

[31] https://en.wikipedia.org/wiki/Support-vector_machine

[32] https://en.wikipedia.org/wiki/Artificial_neural_network

[33] https://en.wikipedia.org/wiki/Canny_edge_detector

[34] https://en.wikipedia.org/wiki/Random_forest

[35] https://en.wikipedia.org/wiki/Contourlet

[36] https://images.app.goo.gl/PV4eAR55aYEPXMyv6

ABBERAVIATION

CTD	Curvelet Transform Descriptors
CT	Curvelet Transform
LBP	Local Binary Pattern
LP	Laplacian Pyramid
DFB	Directional Filter Bank
GLCM	Gray level co-occurrence matrix
RGB	Red-Green-Blue
RF	Random Forest
ANN	Artificial Neural Network
SVM	Support Vector Machine
DNN	Deep Neural Network
NN	Neural Network
K-NN	K-nearest Neighbour
MLP	Multilayer Perceptron
LDA	Linear Discriminant Analysis
GUI	Graphical User Interface
ML	Machine Learning
DCT	Discrete Contourlet Transform
NSCT	Non sub-sampled Contourlet Transform
PDFB	Pyramid Directional Filter Bank

APPENDIX

PAPER PUBLICATION

Following Research Paper is published in UGC Approved Journal:

- An article "Designing of leaves Classifier for ayurvedic plants using Hybrid Feature Extraction" has been published for publication by the International Journal of Research in Electronics and Computer Engineering (IJRECE), Print ISSN: 2393-9028, Online ISSN: 2348-2281 VOL. 7, ISSUE 2, APRIL - JUNE 2019. The Certificate of research paper is below:

ISSN (Online): 2348-2281 ISSN (Print): 2393-9028

IJRECE

I2OR

CERTIFICATE

This is to certify that **Hinal Patel**

of **Sardar Vallabhbhai Patel Institute of Technology, Vasad, Gujarat, India**

published a paper titled **Designing of Leaves Classifier for Ayurvedic Plants Using Hybrid Feature Extraction**

authored by **Hinal Patel, Dr. Nehal G. Chitaliya**

in the International Journal of Research in Electronics and Computer Engineering (IJRECE)

Vol. 7 Issue 2 **April to June 2019**.

Chief Editor

Managing Editor

ijrece.i2or@gmail.com

i2or-ijrece..com

❖ Comments For Internal Review (2730002) (Semester 3)

Exam Date : __/__/__

Sr. No.	Comments given by Internal review panel (Please write specific comments)	Modification done based on Comments
–	Detailed study is required.	Working Hard to fulfil the comments
–	Fundamentals for bale ground work is very weak.	
–	Good project no definition, but need extensive work on classifier & feature extraction.	Nehal
–	presentation skills are ~~bad~~ too much poor.	
		(Guide Sign.)

Particulars	Internal Review Panel	
	Expert 1	Expert 2
Name :	Dr. IC R Bhatt	Dr. M. V. Chitaliya
Institute :	SVIT-Vasad.	SVIT-Vasad
Institute Code :	041	041
Mobile No. :	9879191205	
Sign :		Nehal

Particulars	Internal Guide Details	
	Expert 1	Expert 2
Name :		
Institute :		
Institute Code :		
Mobile No. :		
Sign :		

Enrollment No. of Student : 1 7 0 4 1 0 7 5 4 0 0 2

Hall No. : 10

Exam Date : 17/12/2018

Sr. No.	Comments given by External Examiners (DP-I) : (Please write specific comments)	Modification done based on Comments
1.	Detailed study of different image masks are necessary.	Dataset is (downloaded) Prepared & work started on it.
2.	Good leaf database is desirable.	
		Nehl
		(Internal Guide Sign.)

- *Approved* ☑
- *Approved with suggested recommended changes* ☐
- *Not Approved* ☐

Please tick on any on. If approved/approved with suggession then put marks ≥ 50 %.

➢ Details of External Examiners :

Particulars	*Full Name*	*University / College Name & Code*	*Mobile No.*	*Sign.*
Expert 1	C. H. VITHALANI	GEC RAJKOT 020	9374836390	
Expert 2	A. H. Dafda	VGEC 017	9408481646	

Enrollment No. of Student : 170410754002

❖ **Comments of Mid Sem Review (2740001)** **(Semester 4)**

Exam Date : 26/02/2019 **Hall No :** 05

Sr. No.	**Comments given by External Examiners :** i) The appropriateness of the major highlights of work done; State here itself if work can be approved with some additional changes. ii) Main reasons for approving the work. iii) Main reasons if work is not approved.	**Modification done based on Comments**
1.	Segmentation concepts should be strengthened. OTSU segmentation is not known to candidate	Improving fundamentals
2.	Understanding of the concepts is weak. Must improve, as the basics of analysis and implementation is not known to the candidate.	
		Internal Guide Sign.

- *Approved* ☐
- *Approved with suggested recommended changes* ☑
- *Not Approved* ☐

Please tick on any on.
If approved/approved with suggession then put marks ≥ 50 %.

➢ ***Details of External Examiners :***

Particulars	*Full Name*	*University / College Name & Code*	*Mobile No.*	*Sign.*
Expert 1	C.R. PAREKH	GEC, Sec-28, (013)	9428277558	
Expert 2	Amit Upadhyay	CHARUSAT		

Turnitin Originality Report

Processed on: 24-Apr-2019 13:31 IST
ID: 958801134
Word Count: 10815
Submitted: 35

DESIGNING OF LEAVES CLASSIFIER FOR AYURVEDIC PLANTS USING HYBRID FEATURE EXTRACTION By Svit Vasad 041

Similarity Index	Similarity by Source
11%	Internet Sources: 8% Publications: 8% Student Papers: 6%

2% match (student papers from 06-May-2018)
Submitted to Charotar University of Science And Technology on 2018-05-06

2% match (Internet from 30-Jun-2018)
https://en.wikipedia.org/wiki/Contourlet

1% match (publications)
P. Manoj Kumar, C. M. Surya, Varun P. Gopi, "Identification of ayurvedic medicinal plants by image processing of leaf samples", 2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), 2017

1% match (Internet from 23-Apr-2019)
https://en.wikipedia.org/wiki/Artificial_neural_network

1% match (student papers from 27-Sep-2018)
Submitted to American University of the Middle East on 2018-09-27

1% match (Internet from 25-Nov-2017)
http://csis.pace.edu/~ctappert/srd2017/Proceedings/2017ProceedingsForPrint.pdf

1% match (student papers from 18-Apr-2019)
Submitted to Universiti Teknologi MARA on 2019-04-18

1% match (Internet from 05-Apr-2014)
http://www.imageclef.org/2013/plant

1% match (publications)
A. Jeyalakshmi, D. Ramya Chitra, "Comparative study of feature extraction using several wavelet transforms for source camera identification", 2018 2nd International Conference on Inventive Systems and Control (ICISC), 2018

1% match (publications)
D Venkataraman, N. Mangayarkarasi, "Support vector machine based classification of medicinal plants using leaf features", 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2017

1% match (student papers from 09-Feb-2019)
Submitted to Universiti Tenaga Nasional on 2019-02-09

1% match (Internet from 13-Mar-2019)
https://en.wikipedia.org/wiki/Deep_Learning

1% match (publications)
A. Hadi-Vencheh, "A new nonlinear model for multiple criteria supplier-selection problem", International Journal of Computer Integrated Manufacturing, 01/2011