ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

Identification of Leaf using DWT and SVM Classifier

Shreedhar , Anitha G

M. Tech Student, Dept. of Computer Science Engineering, UBDTC, Davanegere, India

Associate Professor, Dept. of Computer Science Engineering, UBDTC, Davanegere, India

ABSTRACT: In tropical regions, leaf identification is a challenging task, amongst the entire organ which are utilized in plant identification, leafs are probably the most used ones when compare to other organ. For instance, taxonomists can depend on leaves to search for patterns that can be used to identify a plant species. Color and Texture features were incorporated for the identification of the leaf; the proposed approach consists of four phases such as pre-processing, segmentation, feature extraction and classification. The pre-processing phase involves color transformation and resize. In order to extract leaf shape we are using DWT algorithm and shape features, leaf has to be segmented. Hence the DWT is used for segmentation. For feature extraction phase the color and texture features are derived. SVM classifier is used to classify and identify the type of leaf.

KEYWORDS : Discrete wavelet transform, SVM classifier and Shape features.

I. INTRODUCTION

Many methods have been proposed to identify plant leaves in an automated fashion. A large percentage of such works utilize shape recognition techniques to model and represent the contour shapes of leaves, however additionally, color and texture of leaves have also been taken into consideration to improve recognition accuracies. In tropical regions, amongst the entire organ which are utilized in plant identification, leafs are probably the most used ones. Mainly because they're commonly present when compared to flower and fruits. So far, many researchers have proposed various approaches in identifying leaves in order to develop a system that can help ordinary people to identify plants based on a leaf.

The features used in the identification systems can be classified into shape or morphological features, texture features, margin features, vein features and color features. An existing approach that combines Fourier descriptors with other shape features was investigated to identify 100 hundred kinds of leaves. They used Bayes's classifier to classify the image. The result shows that the combination of Fourier descriptors and several other shape features can be used to identify leaves. In this system they just identified the type of the leaf using shape and other features such as convexity, irregularity, roundness factor of the image. This approach may lags in terms of efficiency or accuracy. To overcome all these limitation a new approach is proposed in this paper, which consists of four phases such as pre-processing, segmentation, feature extraction and classification. The pre-processing phase involves color transformation and resize. In order to extract leaf shape and boundary features, leaf has to be segmented. Hence the adaptive clustering algorithm is used for segmentation. For feature extraction phase the color and texture features are derived from an trained SVM to knowledge base and SVM classifier will classifies the type of leaf according to the extracted feature properties.

II. LITERATURE SURVEY

Smita Patil et.al [1] has proposed a system on Identification of Growth Rate of Plant Based on Leaf Features using Digital Image Processing Techniques. In order to extract leaf shape and boundary features, leaf has to be segmented from a plant. Hence the watershed algorithm is used for segmentation. For feature extraction phase the color and texture features are derived. Total of 141 features are extracted in which 121 color features are taken using histogram and 20 texture features using Discrete Wavelet Transform and Fast Fourier Transform and are given as input vector to the Support Vector Machine classifier. The ultimate goal of this work is to develop a system where a user in the field can

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

take a picture of a plant leaf, feed it to the machine vision system carried on a portable computer, and have the system to classify and identify the growth rate.

The classification accuracy of 94.73% is observed. Jyotismita Chaki et.al [2] has proposed a system on Plant Leaf Recognition using Gabor Filter. This project addresses the problem of segmenting an image into different regions. They analyze two unsupervised learning algorithms namely the K-means and EM and compare it with a graph based algorithm, the Normalized Cut algorithm. The K-means and EM are clustering algorithms, which partition a data set into clusters according to some defined distance measure. The Normalized Cut criterion takes a measure of the similarity between data elements of a group and the dissimilarity between different groups for segmenting the images. Abdul Kadir et.al [3] has proposed a architecture on Performance Improvement of Leaf Identification System Using Principal Component Analysis. This system involved combination of features derived from shape, vein, color, and texture of leaf. PCA was incorporated to the identification system to convert the features into orthogonal features and then the results were inputted to the classifier that used Probabilistic Neural Network (PNN). This approach has been tested on two datasets, Foliage and Flavia, that contain various color leaves (foliage plants) and green leaves respectively. The results showed that PCA can increase the accuracy of the leaf identification system on both datasets.

Sofiene Mouine Et.al [4] has proposed a Advanced shape context for plant species identification using leaf image retrieval. This paper presents a novel method for leaf species identification combining local and shape-based features. Our approach extends the shape context model in two ways. First of all, two different sets of points are distinguished when computing the shape contexts: the voting set, i.e. the points used to describe the coarse arrangement of the shape and the computing set containing the points where the shape contexts are computed. This representation is enriched by introducing local features computed in the neighbourhood of the computing points. Experiments show the effectiveness of our approach.

III. METHODOLOGY

Figure 1 represents the proposed architecture. The algorithm consists of major two parts testing and training phase respectively. In training phase the system is trained to create a knowledge base. First stage of training phase is to read the sampled leaf images stored in the database. Then images are pre-processed to get resized and color conversion of images. When the images are pre-processed feature extraction is achieved by applying DWT feature extraction method and also ten shape features of leaves, SVM training algorithm will train the knowledge base according to the feature which we extracted in feature extraction step.

Testing consists of leaf image as a system input to the top layer of the architecture. An image pre-processing is used to change the color and size of the image. Discrete Wavelet Transform algorithm is used on segmented images in order to extract the features. These extracted features are subjected to SVM classifier where this classifier will classifies the type of leaf by compare it with the database present in the knowledge base which we stored during training phase.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

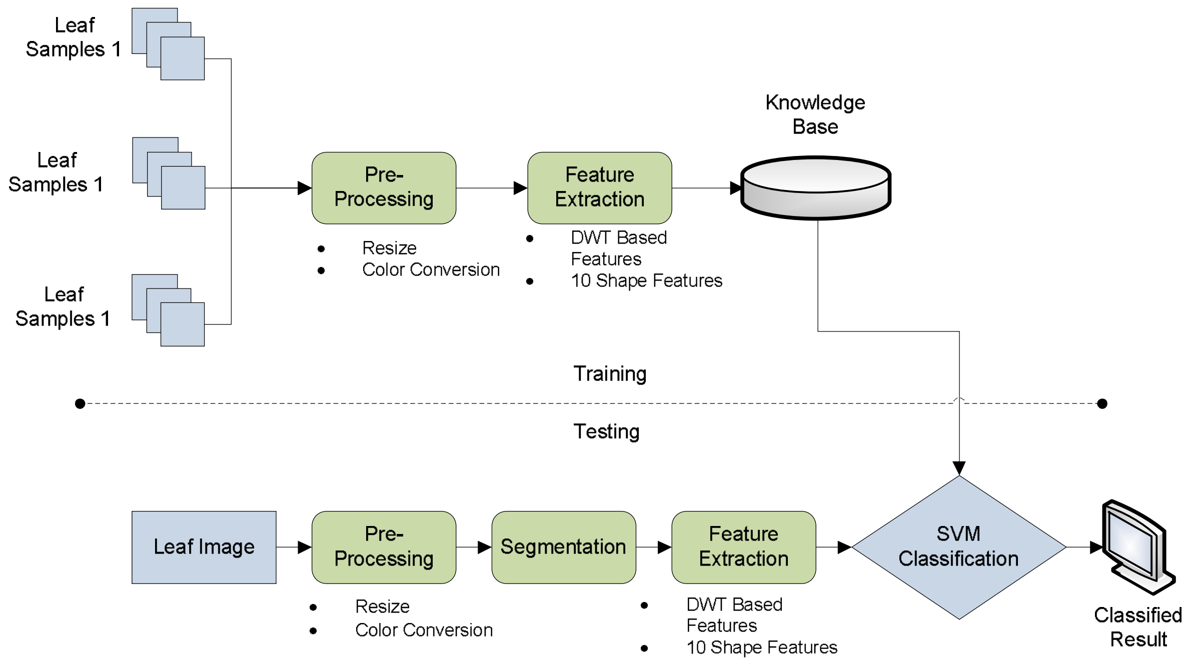


Figure 1: Block Diagram Of Proposed System.

A. PRE-PROCESSING

Pre-processing is mainly used to adjust the size of the image, removal of noise, color conversion and isolating objects of interest in the image. Pre-processing is any form of signal processing for which the output is an image or video, the output can be either an image or a set of characteristics or parameters related to image or videos to improve or change some quality of the input. This process will help to improve the video or image such that it increases the chance for success of other processes. In this paper we considered sampled leaf images as input and those images are subjected to pre-processing this will resulting in color conversion, resizing of input sample images.

B. FEATURE EXTRACTION

First stage of training phase is to read the sampled leaf images stored in the database. Then images are pre-processed. When the image is pre-processed feature extraction is achieved. For feature extraction we use DWT and 10 shape features.

1. DWT(DISCRETE WAVELET TRANSFORM)

A discrete wavelet transform is a type of transform in which the wavelets are discretely sampled. We are using DWT for text edge extraction process. In analyzing a real world signals DWT will provide a multi-resolution description of a signal. Essentially, a discrete multi-resolution description of a continuous-time signal is obtained by a DWT. DWT converts a series a0, a1, a2……. am into one low pass coefficient series known as approximation and one high pass coefficient series known as detail. Length of each series is m/2. In real life situations, such transformation is applied recursively on the low-pass series until the desired number of iterations is reached. The discrete wavelet transform has applications ranging from data compression to signal coding. The pre-processed segmented images are then subjected to SVM classifier which is explained in next section.

2. SHAPE FEATURES

Feature extraction is used to extract relevant features for recognition of plant leaves. Figure 2 represents the steps how we are processing the feature extraction.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

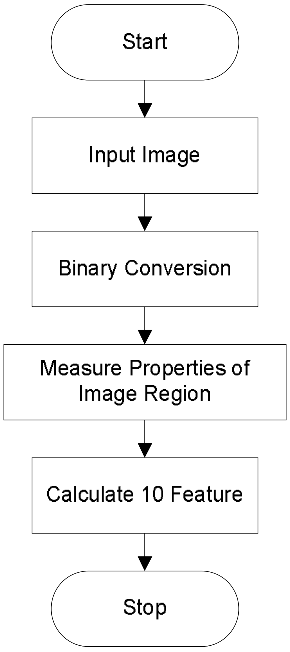


Figure 2: Steps for extracting features.

The redundancy is removed from the image and the leaf images are represented by a set of numerical features. Figure 2 represents the steps how we are processing the feature extraction. These features are used to classify the data. The following features are used for doing leaves classification .

Extraction of Eccentricity: Region-based method is used to estimate the best fitting ellipse for the extraction of eccentricity, scalar that specifies the eccentricity of the ellipse that has the same second-moments as the region. The eccentricity is the ratio of the distance between the foci of the ellipse and its major axis length. The value is between 0 and 1. (0 and 1 are degenerate cases; an ellipse whose eccentricity is 0 is actually a circle, while an ellipse whose eccentricity is 1 is a line segment.) This property is supported only for 2-D input label matrices.

Extraction of Aspect Ratio: For a rectangular shape, aspect ratio is defined as the ratio of its breadth by length. For leaves having irregular shapes, the aspect ratio is defined as the ratio of the minor axis to the major axis of the best fitting ellipse.

$$\text{Aspect Ratio} = \frac{\text{Width of leaf}}{\text{Length of leaf}} \quad (1)$$

Leaf Area: The value of leaf area is evaluated by counting the number of pixels having binary value 1 on smoothed leaf image. It is denoted as A.

Leaf Perimeter: Leaf perimeter is defined as distance around the boundary of the region. It is evaluated by counting the number of pixels containing leaf margin. It is denoted by P.

Area ratio of perimeter: Ratio of Area to perimeter is defined as the ratio of leaf Area A and leaf perimeter P, is calculated by A/P.

Length of the major axes: Scalar specifying the length (in pixels) of the major axis of the ellipse that has the same normalized second central moments as the region. This property is supported only for 2-D input label matrices.

Length of minor axes: Area of the ellipse that just encloses the region. The length (in pixels) of the minor axis of the ellipse that has the same normalized second central moments as the region. This property is supported only for 2-D input label matrices.

Solidity: Scalar specifying the proportion of the pixels in the convex hull that are also in the region computed as Area/ Convex Area. This property is supported only for 2-D input label matrices.

$$\text{Solidity} = \frac{\text{Area of leaf}}{\text{Area of convex hull}} \quad (2)$$

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

Along with these features will calculate the black pixels and white pixels on the pre-processed image, next according to the features which we extracted will identify and SVM classifier will classify the spices of plant depending on the leaf shape.

C. SVM CLASSIFIER

SVM stands for Support Vector Machines. In order to achieve the optimal separating hyper-plane in the higher dimensional feature space, SVM first maps the input vector into a higher dimensional feature space.

The kernel function on two samples X and X', represented as feature vectors in some input space, is defined as

$$K(X, X') = \exp\left(\frac{\|X - X'\|^2}{2\sigma^2}\right) \quad (3)$$

$\|X - X'\|^2$ is the squared Euclidean distance between the two feature vectors. σ is a free parameter.

Furthermore, a decision boundary, i.e. the separating hyper-plane, is determined by support vectors rather than the whole training samples and thus is extremely robust to outliers. Exactly an SVM classifier is designed for binary classification. That is, to separate a set of training vectors which belong to two different classes. Note that a decision boundary similar to the support vector i.e. training samples. To provide the required mapping to ice-water labels a soft-margin SVM classifier is used. An SVM works by computing a linear decision boundary in a high dimensional space using the subset of labelled training samples near the decision boundary (called the support vectors). The SVM decision boundary equation is

$$f(x) = \sum_{\forall i} y_i \, \alpha_i K(x_i, x) \quad (2)$$

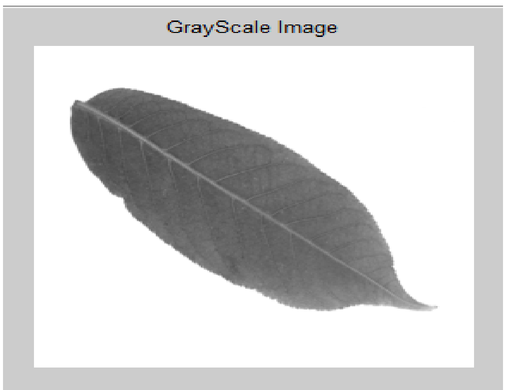
Where $K(x_i, x)$ Kernel function is defined by

$$K(x_i, x) = \exp(-\gamma |x_i - x|^2) \quad (3)$$

Then according to the result which we stored in the knowledge base at training phase SVM classifier will classify the segmented image by comparing it with sampled input image.

IV.EXPERIMENTAL RESULT

Figure 3 represent the experimental result of the proposed system. Firstly will consider sampled leaf image as input image, which is shown in Figure 3(a) after considering sampled leaf images are pre-processed, then images are converted into gray scale images as shown in Figure 3(b), these pre-processed images are taken for feature extraction and SVM training will train to knowledge base. Lastly the SVM classifier classifies the images according to database and result is shown in Figure 3(c).The effective experimental results shows how effectively our proposed system will work.

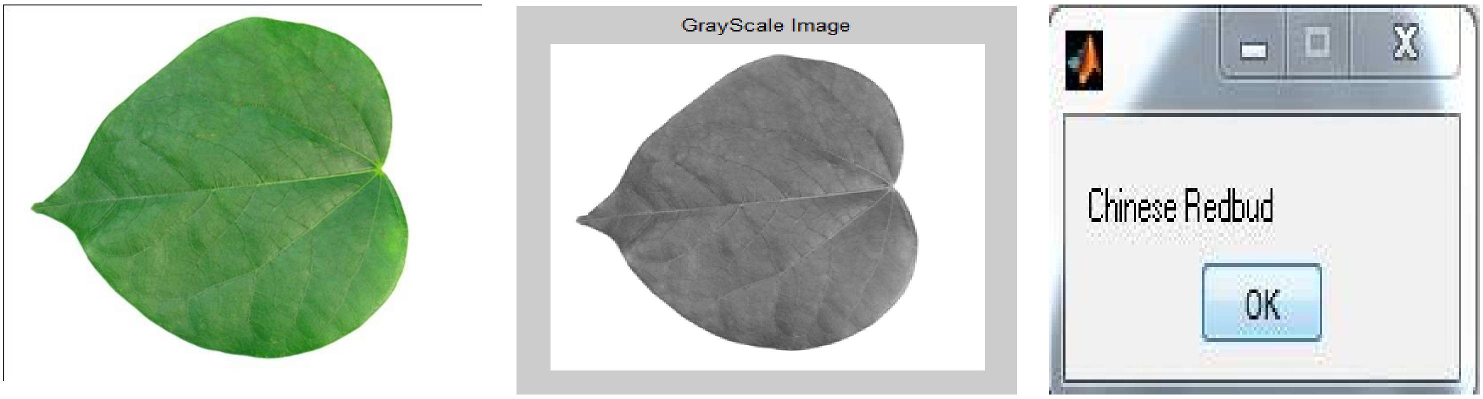


ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016



(a) (b) (c)

Figure 2: (a) Input leaf Images; (b) Gray Scale Image; (c) Result

V. CONCLUSION

In this paper propose an end-to-end leaf identification by shape features extraction. The leaf identification component uses SVM classifier based on rich shape feature extraction such as Area, perimeter, Ratio of aspect, Minor axes length, Major axes length and solidity, this leading to SVM speed improvement. Detection is done using DWT (Discrete wavelet transform) and morphological operations. For identification of type of leaf process SVM is used and for feature extraction DWT is used. Extensive evaluation on a large dataset illustrates the efficiency of our approach.

REFERENCES

[1] Smita Patil, Shridevi Soma and Suvarna Nandyal, "Identification of Growth Rate of Plant Based on Leaf Features using Digital Image Processing Techniques", Volume 3, Issue 8, August 2013.

[2] Suman Tatiraju and Avi Mehta, "Plant Leaf Recognition using Gabor Filter", Jyotismita Chaki and Ranjan Parekh, Volume 56, Issue 10, October 2012.

[3] Abdul Kadir, Lukito Edi Nugroho, Adhi Susanto and Paulus Insap Santosa, "Performance Improvement of Leaf Identification System Using Principal Component Analysis", Volume 44 2012.

[4] Sofiene Mouine, Itheri Yahiaoui and Anne Verroust-Blondet, “Advanced shape context for plant species identification using leaf image retrieval", 2012.

[5] Jyotismita Chaki and Ranjan Parekh, “Plant Leaf Recognition using Shape based Features and Neural Network classifiers”, Volume 2, Issue 10, 2011.

[6] C.L. Lee and S.Y. Chen. Classification for leaf images. In 16th IPPR Conference on Computer Vision, Graphics and Image Processing, volume 8, pages 17–19, 2003.

[7] B. S. Bama, S. M. Valli, S. Raju, and V. A. Kumar, “Conten based leaf image retrieval using shape, color and texture features”, Indian Journal of Computer Science and Engineering, vol. 2, no. 2, pp. 202-211, 2011.

[8] A.F. Ali and D.M. Shawky. A novel approach for protein classification using Fourier transfer. World Academy of Science, Engineering and Technology, 4:193–197, 2010.