Leaf Recognition System using Multi-Class Kernel Support Vector Machine

Abeer Toheed Quadri  
Department of Software Engineering  
Fatima Jinnah Women University  
The Mall, Rawalpindi, Pakistan  
abeertoheed@ymail.com

Mehreen Sirshar  
Department of Software Engineering  
Fatima Jinnah Women University  
The Mall, Rawalpindi, Pakistan  
msirshar@gmail.com

ABSTRACT—This paper discusses Leaf Recognition of different species of plants using multiclass kernel support vector machine, an efficient machine learning approach for the classification purpose. This proposed approach consists of three phases such as preprocessing, feature extraction and classification. The preprocessing phase involves a typical image processing steps such as transforming to gray scale and Binary conversion. The main contribution of this approach is the Multi-Class Kernel Support Vector Machine (SVM) for leaf recognition. Classifier tested with flavia dataset [1] and built on Visual Studio C# 2010.

Index Terms—Leaf Recognition; Support Vector Machine; Kernel Support Vector Machine; Multi Class Support Vector Machine; Pattern Recognition

I. INTRODUCTION

Plants play a significant role in our lives but unfortunately many of the leaves are at their edge of their extinction. So creation of a plant database is need of today to protect biosphere of earth. Apart from that different plants play medicine and can be used as bio-fuel. Plants can be classified on the basis of flower, leaf, root and fruits etc.

This is a digital era and different techniques have been designed have been applied to image processing with artificial intelligence for the classification plant. Different studies have been conduction using different variables but manipulation is easy with digital morphological features (DMFs).

II. RELATED WORK

Different attempts have been made by different researchers for plant recognition using leaf classification. Some attempt to classify leaf on the basis of its color histogram some have followed texture and edge information. Plants have also been classified as herbs, shrubs and trees. Research that is being followed in this paper classifies leaves by training Multi-class kernel support vector machine. Considerable work has been done on plant classification on the basis of its shape. Wu et al. [1], extracted 12 commonly used DMFs which were orthogonalized into 5 principal variables using PCA. They used 1800 leaves to classify 32 kinds of plants using a PNN system. Wang et al. [2], employed centroid contour distance (CCD) curve, eccentricity and angle code histogram (ACH). Fu et al. [3] also used centroid-contour distance curve to represent leaf shapes in which an integrated approach for an ontology-based leaf classification system is proposed.

Nature has made different leaves with a unique different shape and size. Using this feature Multiclass KSVM is trained and tested. First ten indexed leaves have been used from flavia data set [1]. Only well-defined and refined leaf images were selected training and testing of the system. Original dataset image size was 1024x1024 that has been reduced to 32x32 for less computation. Plant leave images that have been selected for training and testing of system are: Phyllostachys edulis (Carr.) Houz., Aesculus chinensis, Berberis anhweiensis Ahrendt, Cercis chinensis, Indigofera tinctoria L.,Acer Palmatum,Phoebe nanmu (Oliv.) Gamble,Kalopanax septemlobus (Thunb. ex A.Murr.) Koidz.,Cinnamomum japonicum Sieb.,Koelreuteria paniculata Laxm.

III. IMAGE PREPROCESSING

A. RGB to Gray Scale

All images are of 32x32 JPG and are converted to Gray Scale by using Equation 1:

\[ \text{Gray} = 0.2989 \times R + 0.5870 \times G + 0.1140 \times B \] (1)

R, G and B are color of pixels.

B. Gray Scale to Binary

All images are then converted to Binary by applying threshold mechanism mentioned in Equation 2:
IV. APPLIED CLASSIFICATION SCHEME

SVM is a classifier belonging to supervised learning techniques that is used for regression and classification. SVM classification training algorithm builds a model that is able to predict whether the new sample will lay in which category. SVM optimal hyper plane algorithm belongs to linear classifier, Equation 3:

\[ F(x) = \sum_{i=1}^{N} w_i z_i x + b \]  

Where \( z \) is given support vectors, \( w \) is set of weights.

Kernel Support Vector Machines (KSVM) is a non-linear classifier by applying kernel trick on maximum margin hyperplanes. It works on the fact that if the input data is first mapped into a higher dimensional space, the learning algorithm that is operating in this space will start to behave non-linear in the original input space. Kernel function can be defined as Equation 4:

\[ K(x, y) = (\phi(x), \phi(y)) \]  

Where \( \phi \) is a non-linear mapping function.

SVMs are binary classifiers and they are only able to take decision between two classes. To convert original SVMs to multi-class consider 3 classes A, B and C. Now divide multi-class into binary class problem i.e. A x B, A x C, C x B etc. Now it’s useless to compute A x B and B x A. So a decision between \( n \) classes can be can be decomposed as \( n(n-1)/2 \) binary problems.

This paper has taken 10 distinct leaves so the total classes are 10 and by decomposing those 45 binary SVMs are made.

The training algorithm that has been used is Sequential Minimal Optimization (SMO) algorithm. Applying kernel approach is itself very efficient as it can be applied to a problem that demand significant data pre-processing like dimensionality reduction in case of leaf recognition.
After system training is completed it is tested by classification.

Created SVMs can be viewed in machine tab. Dark vectors have more weight in decision process.

V. EXPERIMENTAL RESULTS

First 10 indexed leaves samples from flavia dataset [1] is been used to train and test the machine. Only well-defined and refined images were selected to test the proposed approach that gives 100% accuracy. Results would have been different if not well-defined dataset was used with leaves features not visible. But still the experimental results are better than previously used approaches for leaf recognition. Experimental results are mention in Table I while comparative results are mention in Table II.

<table>
<thead>
<tr>
<th>Table 1: Accuracy Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheme</td>
</tr>
<tr>
<td>Proposed in [6]</td>
</tr>
<tr>
<td>1-NN in [7]</td>
</tr>
<tr>
<td>MMC in [8]</td>
</tr>
<tr>
<td>BPNN in [8]</td>
</tr>
<tr>
<td>RBFNN in [9]</td>
</tr>
<tr>
<td>MLNN in [9]</td>
</tr>
<tr>
<td>k-NN (k = 4) in [10]</td>
</tr>
<tr>
<td>RBPNN in [7]</td>
</tr>
<tr>
<td>k- (k = 5) in [7]</td>
</tr>
<tr>
<td>PNN[11]</td>
</tr>
<tr>
<td>MMC in [10]</td>
</tr>
<tr>
<td>SVM [12]</td>
</tr>
<tr>
<td>Multiclass KSVM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2: Experimental Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flavia Index</td>
</tr>
<tr>
<td>--------------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>
VI. CONCLUSION
A new approach of plant classification based on leaves recognition is proposed in this paper based on Multi-class Kernel Support Vector Machine. The approach consisted of three phases that are the preprocessing phase, System Training and the classification phase. The system can automatically classify 10 kinds of plants via the leaf images loaded from digital cameras or scanners. The performance of the proposed approach is evaluated based on the accuracy and compared with other methods. For further research by incorporating system testing with different dataset can be checked.

REFERENCES

3 2 Berberis anhweiensis Ahrendt Anhui Barberry http://asaweb.huh.harvard.edu:8080/databases/specimens?id=277371 100%
4 3 Cercis chinensis Chinese redbud http://www.ag.auburn.edu/hort/landscape/dbpages/306.html 100%
5 4 Indigofera tinctoria L. true indigo http://www.itis.gov/servlet/SingleRpt/SimpleRpt?search_topic=TSN&search_value=565257 100%
6 5 Acer Palmatum Japanese maple http://en.wikipedia.org/wiki/Acer_palmatum 100%
7 6 Phoebe nanmu (Oliv.) Gamble Nanmu http://www.ars-grin.gov/cgi-bin/npgs/html/taxon.pl?28039 100%
8 7 Kalopanax septemlobus (Thunb. ex A.Murr.) Koidz. castor aralia http://www.itis.gov/servlet/SingleRpt/SimpleRpt?search_topic=TSN&search_value=503286 100%
9 8 Cinnamomum japonicum Sieb. Chinese cinnamon http://en.wikipedia.org/wiki/Cinnamomum_japonicum 100%
10 9 Koelreuteria paniculata Laxm. goldenrain tree http://www.itis.gov/servlet/SingleRpt/SimpleRpt?search_topic=TSN&search_value=56750 100%