Medicinal Plant Leaf Image Classification and Analysis using SVM Classifier Kernel Functions

[1] KAYATHIRI M, [2] KRISHNAVENI K, [3] PONMALAR K

[1] M.Phil Scholar, [2] Associate Professor, [3] Assistant Professor, Department of Computer Science.
[1] S.R.N.M College, Sattur, Virudhunagar Dist, Tamilnadu, India.

Abstract: Newsday's plant classification based on leaf identification is becoming a popular. Plants are fundamental for human beings, so it's play vital role categorize and preserve all the plants. Each leaf carries significantly information that can be used to identify and classify the origin of the plant. An image processing techniques based on leaves recognition can help to find the best features useful for plant representation and classification. In this research we propose a leaf recognition method which uses a new features set that incorporates shape, color and texture features. A total of 760 leaf features are extracted and used for training a SVM classification learner App. The method has been tested on input dataset showing excellent performance both in terms of accuracy that often reaches 96%, color features.

Keywords *— Leaf images, Preprocessing, Feature Extraction, shape, color and texture features, SVM Classification.*

1. INTRODUCTION

Image processing plays a major role in the identification of medicinal plants by extracting the features of herbal leaf and authenticating its medicinal traits. Due to the two dimensional representation of leaves, the medicinal plants are easily identified and recognized by analyzing the shape, texture, color, aspect ratio, vein structure of leaves rather than fruits, flowers etc. Since manual recognition requires expert botanist, an image processing technique using Support Vector Machine (SVM) classifier is proposed in this paper to recognize the medicinal plants.

The paper is organized into four sections. The literature survey is discussed in section II. The proposed methodology for the classification of medicinal plant leaf images is discussed in section III. The experimental results and analysis is presented in Section IV and Section V concludes this research work.

2. LITERATURE REVIEW

ArunPriya.C, et. al, proposed an efficient leaf recognition algorithm. Initially 12 shape and color features are extracted and reduced to 5 important features and classified using Support Vector Machine and K-Nearest Neighbour classifiers [1]. Jiandu Liu, et. al, proposed Plant leaf images category based on Wavelet convert and Support Vector Machines. Initially, the leaf images are distorted into the time-frequency field image by wavelet transforms not including any additional preprocessing such as image improvement and texture retreating [3]. Shitala Prasad, et. al, proposed Relative Sub-Image Based Features to predict 23 types of plant leaves by extracting 300 features [4]. S.Basavaraj, et. al, proposed a combined color, texture and edge features based approach for recognition and categorization of Indian medicinal Plants[5]. Lbrahim Z, et. al, proposed multiclass Support Vector Machine (SVM) classifier to identify the medicinal plants using the texture features namely Histogram of Oriented Gradients (HOG), Local Binary Pattern (LBP) and Speeded-Up Robust Features (SURF) with. On Flavia data set, HOG and LBP produce better results than SURF [6]. Jin L, et. al, proposed a routine categorization method to classify 12 different medicinal plant leaf images and achieved a standard successful identification rate of 93.3%. The result indicates that it is sufficient to mechanically classify medicinal plants by using multi-feature extraction of leaf images in permutation with SVM [7]. Juby George, et. al, proposed multi-layer perception and support vector machine classifiers. They produced 71.15% accuracy rate for color features, 73.07% for shape features, 66.34% for vein features, 91.34% for combining the color and shape features, and finally 96.13% for the combination of all three features[8]. Ruberto, Lorenzo Putzu (2014) proposed a fast Leaf identification algorithm based on SVM Classifier and High Dimensional Feature Vector [9]. Evanjali Gamit, Kazi A.R., Lokesh Gagnani P proposed supervised classification algorithm to classify the dissimilar families of plant leaves [10]. Ali Caglayan, Oguzhan Guclu, and Ahmet Burak Can proposed k-Nearest Neighbor, Support Vector Machines, Naive Bayes, and Random Forest classification algorithms to identify plant types with the Shape and color features of leaves. Random Forest technique yields 96% accuracy compared to other classifiers [11].

3. METHODOLOGY

Medicinal Plant Leaf Image Classification and Analysis using SVM Classifier Kernel Functions is proposed in this paper. The schematic diagram of the proposed research work is shown in Figure 1.

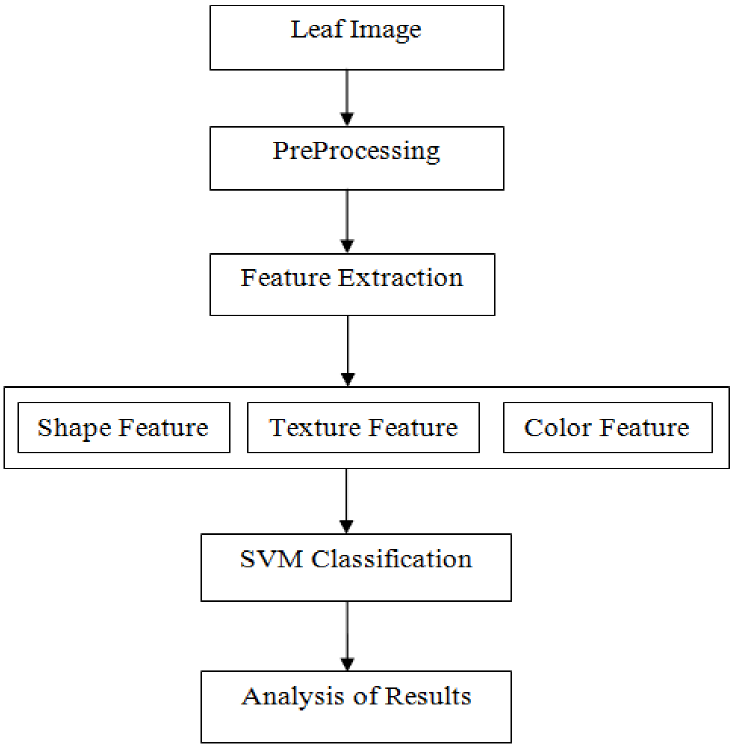


Figure 1: Schematic Diagram of the Proposed Method

A. *Medicinal Plant Leaf Image*

A medicinal plant leaf data base with 760 leaf images of 30 different plant classifications of size 400 x 600 stored in jpg format is created. The sample images of each category are shown in Figure 2.

Figure 2 Medicinal Plant Leaf Images

B. *Preprocessing*

The preprocessing step involves two main tasks, namely enhancement and segmentation. Real world input data always contains some amount of noise and therefore, enhancement techniques that reduce its effect are desirable. The leaf images may also have shadow effect during acquisition. So, initially the input leaf image is converted into HSV color space using (1) and the second channel S which is free from shadow effect can be taken for further processing.

$$H := \begin{cases} 0, & \text{if } MAX = MIN \leftrightarrow R = G = B \\ 60^0.\left(0 + \frac{G-B}{MAX-MIN}\right), & \text{if } MAX = R \\ 60^0.\left(2 + \frac{B-R}{MAX-MIN}\right), & \text{if } MAX = G \\ 60^0.\left(4 + \frac{R-G}{MAX-MIN}\right), & \text{if } MAX = B \end{cases}$$

$$\text{if } H < 0^0 \text{ then } H := H + 360^0$$

$$S_{HSV} := \begin{cases} 0, \text{if } MAX = 0 \leftrightarrow R = G = B = 0 \\ \frac{MAX-MIN}{MAX}, & \text{otherwise} \end{cases}$$

$$V := MAX \qquad \text{------- (1)}$$

The median Blur filter is a nonlinear digital filtering technique, often used to remove impulse and salt-and-pepper noise which can occur due to a random bit error in a communication channel from an image or signal. This operation is performed on the image using

Median Blur (src, dst, ksize) ---------- (2)

Where src -> input image, dst -> output image, ksize -> size of the kernel.

Finally fuzzy optimal Threshold value using Gaussian measure of fuzziness is calculated and applied to extract the leaf structure from the background.

C. *Feature Extraction*

Every plant leaf will be dissimilar from others due to its distinctive features. The following color, morphological shape and Texture features are extracted to train the classifier.

1. Shape and Texture Features

A set of morphological shape features namely Smooth factor, Form factor, Narrow factor, Perimeter ratio of diameter and Vein features are computed [2]. A texture feature which possess high variance between inter class images such as robust beside aging of images, clarification changes and other such factors has been proposed to extract the texture features of the leaf images. The Local Vector Pattern of size 256 x 256 is computed for different orientations (0o, 45o, 90o & 135o) and stored as a training feature dataset with reduced dimension [12].

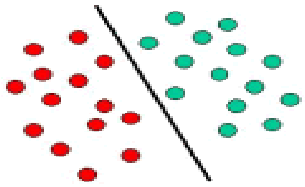
2. Color Feature Vector

The color features are generally represented by the color histogram. A novel color feature descriptor namely HSV Pattern Color Histogram for each color channel is computed, normalized and together as a color feature vector for classifying in the medicinal plant leaf images.

The feature set for all input images are computed and stored as a feature data set and given as input training feature vectors to the classifier.

D. Support Vector Machine Classifier (SVM)

SVM is a supervised machine learning algorithm which can be used for categorization or regression problems. SVM is built up as one of the pattern instrument for machine learning and information mining. Picture group, cash related time arrangement expectation, face recognition, biomedical sign survey, restorative diagnostics, and information mining utilizes SVM now a days. Normally SVM classifies the substance into two classes by figuring the most intense edge hyper-plane among the research substance of both specified classes.



The SVM algorithm provides a selection of six kernel functions:

- Linear SVM
- Quadratic SVM
- Cubic SVM
- Fine Gaussian SVM
- Medium Gaussian SVM
- Coarse Gaussian SVM

All these functions are applied on the input feature vector and the results are analyzed.

4. EXPERIMENTAL RESULTS

This research work was carried out in MATLAB R2017b. A medicinal plant leaf data base with 760 leaf images of 30 different plant classifications of size 400 x 600 stored in jpg format is created. The color, shape and texture features are calculated to form the feature vector to train the data set. The input data set is tested on various SVM Kernel functions and the accuracy based on various features is provided in the following Tables. The accuracy represents the number of medicinal plants correctly classified.

Shape Feature

The classification accuracy is very low when shape features alone are considered. The accuracy of Linear SVM accuracy is very low as 0%. Quadratic and Cubic SVM yields same accuracy in 8% and the Gaussian kernel functions are Fine, Medium and Coarse SVM is yield 20% of accuracy. The shape feature accuracy results shown in Table I. The comparative chart for shape feature accuracy is shown in Figure 3.

Table I Shape Feature Accuracy

SVM Kernel Functions	Shape Feature Accuracy
Linear SVM	0.0%
Quadratic SVM	8.0%
Cubic SVM	8.0%
Fine Gaussian SVM	20.0%
Medium Gaussian SVM	20.0%
Coarse Gaussian SVM	20.0%

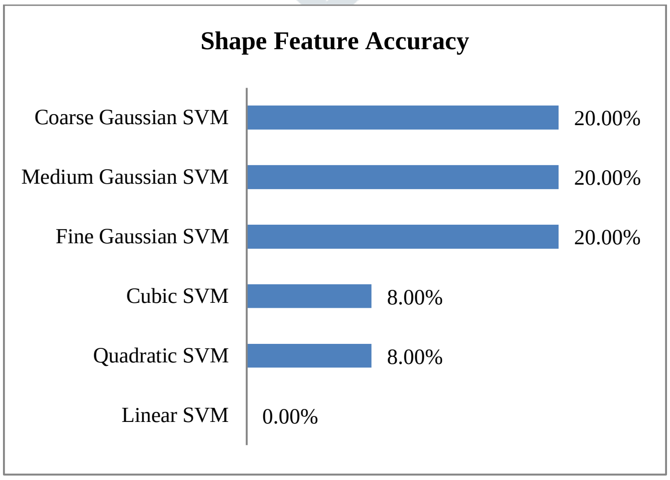


Figure 3 Accuracy of Shape Feature

Texture Feature

The classification accuracy of the texture features is shown in Table II.

Table II Texture Feature Accuracy

SVM Kernel Functions	Texture Feature Accuracy
Linear SVM	16.7%
Quadratic SVM	16.7%
Cubic SVM	16.7%
Fine Gaussian SVM	83.3%
Medium Gaussian SVM	83.3%
Coarse Gaussian SVM	83.3%

SVM kernel functions specified in the Texture Features. Linear, Quadratic and Cubic SVM yields same accuracy as 16.7% and the Gaussian kernel functions are Fine, Medium and Coarse SVM is yield 83.3% of accuracy. The comparative chart is shown in Figure 4.

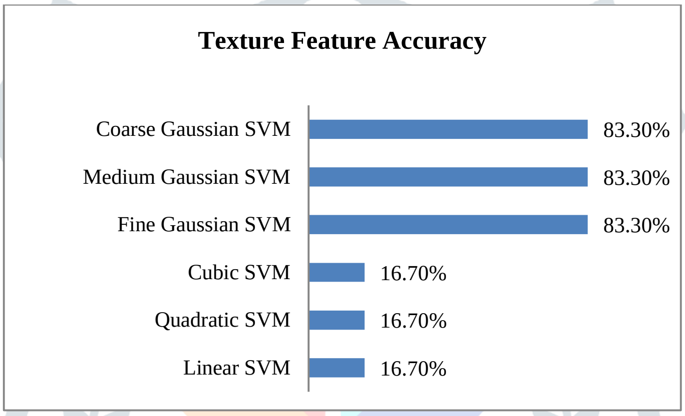


Figure 4 Accuracy of Texture Feature

Color Feature

The classification accuracy obtained by including all the three features Shape, Texture and Color is shown in Table III.

Table III Color Feature Accuracy

SVM Kernel Functions	Color Feature Accuracy
Linear SVM	88.0%
Quadratic SVM	84.0%
Cubic SVM	80.0%
Fine Gaussian SVM	84.0%
Medium Gaussian SVM	96.0%
Coarse Gaussian SVM	96.0%

When Color features are specified in addition, Linear SVM yields 88% accuracy, Quadratic and Fine Gaussian SVM yield same accuracy as 84%. Cubic SVM accuracy is 80%. Finally, the Medium and Coarse Gaussian SVM give same accuracy as 96%. The comparative chart is shown in Figure 5.

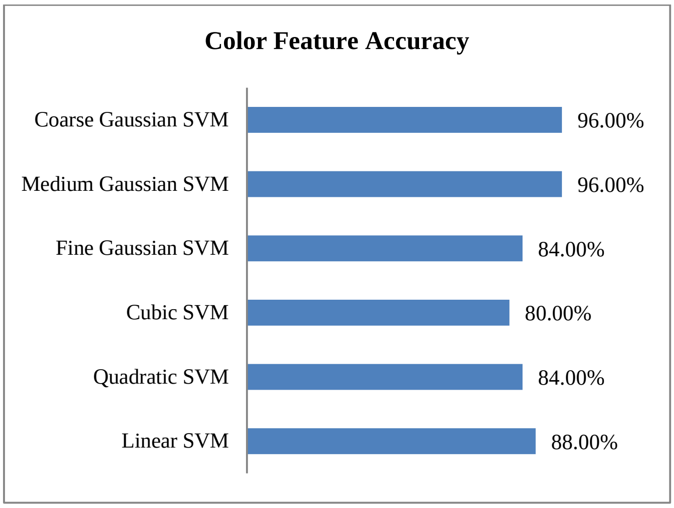


Figure 5 Accuracy of Shape, Texture & Color Features

Finally the input feature vector formed with the combination all the features color, shape and texture is used for classification and the accuracy results. The comparative analysis of the classification results based on shape, shape and texture and all the three features including color feature accuracy results and comparative charts are shown in Table IV and Figure 6.

Table IV Shape, Texture and Color Features Accuracy

SVM Classification	Features		
	Shape	Texture	Color
Linear	0.0%	16.7%	88.0%
Quadratic	8.0%	16.7%	84.0%
Cubic	8.0%	16.7%	80.0%
Fine Gaussian	20.0%	83.3%	84.0%
Medium Gaussian	20.0%	83.3%	96.0%
Coarse Gaussian	20.0%	83.3%	96.0%

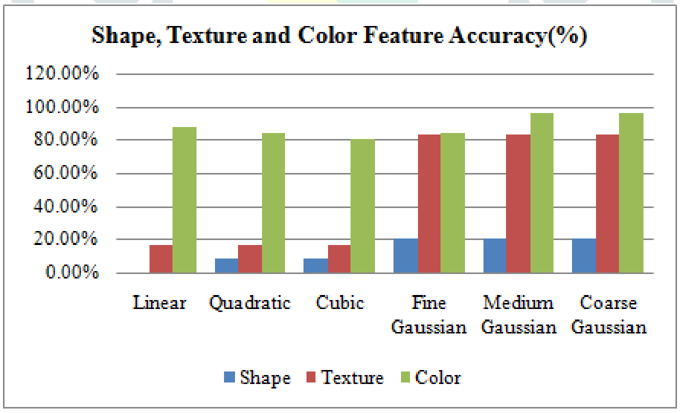


Figure 6 Accuracy Results for All Features

The SVM classifier with the kernel functions Medium Gaussian and Coarse Gaussian yield 96% for the combination of shape, texture and color features.

5. CONCLUSION

Medicinal plant leaf image classification and analysis using Support Vector Machine classifier with various kernel functions is proposed in this paper. A data set containing 760 medicinal plant leaf images of thirty different classes are pre-processed using image processing techniques, morphological shape, texture and color features of the leaf images are extracted and stored as a feature dataset to train the classifier. The feature set for input test image is created, mapped with the training feature vector for classification and the medicinal plant class and its scientific name are displayed. Finally the classification results obtained by various kernel functions are analyzed with different performance metrics. The experimental results show that the SVM classifier with the kernel functions Medium Gaussian and Coarse Gaussian outperform by yielding 96% for the combination of shape, texture and color features. Further analyses can be conducted to improve the current feature extraction process and to include additional features to make the algorithm robust for large leaf data base.

Acknowledgement

M.Kayathiri has received her M.Sc., Computer Science from Sri S.Ramasamy Naidu Memorial College, Sattur affiliated to Madurai Kamaraj University, Madurai, Tamil Nadu and pursuing M.Phil degree in Computer Science in the same Institute.

Dr.K.Krishnaveni, is an Associate Professor and Head of the Department of Computer Science, Sri S.Ramasamy Naidu Memorial College, Sattur, India. She has 25 years of teaching experience. She received her B.E. degree in Computer Science and Engineering from Bharathiar University, Coimbatore, India in 1990 and M.Tech. Degree in Computer and Information Technology from Center for Information Technology and Engineering of Manonmaniam Sundaranar University, Tirunelveli, India in 2004 and received Ph.D degree from the same University. Currently her research interests are Medical Image Processing, Data mining, Machine Learning and Deep Learning Neural Networks. She is a member of IEEE and Editor of various reputed journals.

K.Ponmalar is an Assistant Professor of the Department of Computer Science, Sri S.Ramasamy Naidu Memorial College, Sattur, India. she received her M.Sc degree in Computer Science from Sri S.Ramasamy Naidu Memorial College, Sattur affiliated to Madurai Kamaraj University, Madurai, Tamil Nadu. She completed her M.Phil degree in Computer Science at Vinayaka Mission University, Salem, Tamil Nadu, India. Now she is pursuing her Ph.D degree in Manonmaniam Sundaranar University, Tirunelveli, India. Her area of interests are Digital Image processing, Data mining and Machine Learning algorithms.

REFERENCES

[1] ArunPriya C., Balasaravanan T., Antony Selvadoss Thanamani: “An Efficient Leaf Recognition Algorithm for Plant Classification Using Support Vector Machine” Proceedings of the International Conference on Pattern Recognition, Informatics and Medical Engineering, March 21-23, 2012.

[2] K. Ponmalar, K. Krishnaveni, “Implementation and Evaluation of fuzzy Threshold machine Learning Classifiers to Recognize Medicinal Plant Leaves”, International Journal of Management, Technology And Engineering, Volume 8, Issue XI,2018.

[3] https://www.researchgate.net/publication/220776812_A_Method_of_Plant_Classification_Based_on_Wavelet_Transforms_and_Support_Vector_Machines

[4] Shitala Prasad, Krishna Mohan Kudiri, and R.C. Tripathi, “Relative Sub-Image Based Features for Leaf Recognition using Support Vector Machine”, Research Gate, January 2011.

[5] Basavaraj S. Anami, Suvarna S. Nandyal, and A. Govardhan, “A Combined Color, Texture and Edge Features Based Approach for Identification and Classification of Indian Medicinal Plants”, International Journal of Computer Applications, Vol 6–No.12, September 2010.

[6] https://www.researchgate.net/publication/322469380_Leaf_recognition_using_texture_features_for_herbal_plant_identification.

[7] H. X. Kan, L. Jin, and F. L. Zhou,” Classification of Medicinal Plant Leaf Image Based on Multi-Feature Extraction”, Pattern Recognition and Image Analysis, Vol. 27, No. 3, 2017, pp. 581–587, 1054-6618. © Pleiades Publishing, Ltd.

[8] Juby George, Gladston Raj S, “Comparison of Leaf Recognition using Multi-layer Perceptron and Support Vector Machine”, International Journal of Engineering Research in Computer Science and Engineering (IJERCSE), Vol 5, Issue 4, April 2018.

[9] Cecilia Di Ruberto, Lorenzo Putzu , “A Fast Leaf Recognition Algorithm Based on SVM Classifier and High Dimensional Feature Vector”, Research Gate, January 2014.

[10] Evanjali Gamit, A.R.Kazi, Lokesh P.Gagnani, "plant leaf classification using supervised classification algorithm", International Journal of Computer Science and Mobile Computing, Vol.5 Issue.5, May- 2016, pg. 307-311.

[11] Ali Caglayan, Oguzhan Guclu, and Ahmet Burak Can, "A Plant identification Approach Using Shape and Color Features in Leaf Images", pp.161–170, 2013.

[12] K.Ponmalar, K.Krishnaveni, "Random Forest Classification of Medicinal Plant Leaves Using Shape and Texture Features", Journal of Applied Science and Computations, Volume 5, Issue 8, 2018.

[13] K.Ponmalar, Dr.K.Krishnaveni, "A Novel HSVXOR Pattern Color Histogram Feature Descriptor to Recognize the Medicinal Plant Leaf Images", Universal Review, Volume 7, Issue XI, 2018.

[14] K.Krishnaveni, "A Study on Morphological Watershed Image Segmentation Algorithms and their Applications to Cervical Cytological Images", Ph.D Thesis, MS University, Tirunelveli, 2009.

[15] N.Krishnan, K.Krishnaveni, "A Fuzzy Thresholded Multiscale Morphological Segmentation of Digital Images", International Journal of Signal Processing, Special issue on "Soft Computing in Image Processing Techniques & Systems", Vol.3, no.1, 2005.

[16] N.Krishnan, K.Krishnaveni, "Fuzzy Optimal Thresholded Multiscale Morphological Segmentation of Digital Images", Proceedings of IEEE Third International Conference on Wireless and Optical Communications Network, 2006.