IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 20, 2019, accepted October 30, 2019, date of publication November 7, 2019, date of current version November 20, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2952176

Plant Species Recognition Using Morphological Features and Adaptive Boosting Methodology

MUNISH KUMAR[1], SURBHI GUPTA[2], XIAO-ZHI GAO[3], AND AMITOJ SINGH[1]
[1]Department of Computational Sciences, Maharaja Ranjit Singh Punjab Technical University, Bathinda 151001, India
[2]Department of Computer Science and Engineering, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad 500090, India
[3]School of Computing, University of Eastern Finland, 70211 Kuopio, Finland

Corresponding author: Munish Kumar (munishcse@gmail.com)

ABSTRACT Plant species detection aims at the automatic identification of plants. Although a lot of aspects like leaf, flowers, fruits, seeds could contribute to the decision, but leaf features are the most significant. As a plant leaf is always more accessible as compared to other parts of the plants, it is obvious to study it for plant identification. The present paper introduced a novel plant species classifier based on the extraction of morphological features using a Multilayer Perceptron with Adaboosting. The proposed framework comprises pre-processing, feature extraction, feature selection, and classification. Initially, some pre-processing techniques are used to set up a leaf image for the feature extraction process. Various morphological features, i.e., centroid, major axis length, minor axis length, solidity, perimeter, and orientation are extracted from the digital images of various categories of leaves. Different classifiers, i.e., k-NN, Decision Tree and Multilayer perceptron are employed to test the accuracy of the algorithm. AdaBoost methodology is explored for improving the precision rate of the proposed system. Experimental results are obtained on a public dataset (FLAVIA) downloaded from http://flavia.sourceforge.net/. A precision rate of 95.42% has been achieved using the proposed machine learning classifier, which outperformed the state-of-the-art algorithms.

INDEX TERMS Leaf recognition, feature extraction, k-NN, decision tree, multilayer perceptron, plant leaf classification, plant species identification, AdaBoost.

I. INTRODUCTION

As the future is moving to an artificially intelligent world, machines are replacing human experts in every domain. One such significant domain is agriculture, where the human experts are looking for intelligent machines, which may make their task easier and perform even better than human experts. Such intelligent systems are very crucial, as they are likely to eliminate any chances of ambiguity. Leaf recognition for plant species detection is a significant research zone in the field of image processing and computer vision. Although a lot of methods have been developed so far, the existing computational models for leaf recognition must address a couple of challenging issues. One of these challenges is the extraction of features of plant leaf and their representation so that accurate classification of plant species could be made. Out of many features, leaf shape is a conspicuous element that most algorithms rely on to perceive and describe a plant [1]. In addition, leaf shading, surface, and vein can also be considered for more accurate classification [2]. Each of these components is significant for the recognition and classification of a leaf image. Because of the availability of effortless cameras and remarkable computer vision frameworks, plant/leaf recognition has become an active area of research. The popular frameworks mainly comprise pre-processing, feature extraction & selection and classification. This paper aims to propose a novel classifier for plant species recognition using morphological features enhanced with the adaptive boosting methodology. The major contributions of our paper are:

- Fast and accurate leaf classification for plant species identification
- Utilization of morphological leaf features with low dimensionality
- Evaluation of different classifiers
- Optimize the classification results using AdaBoost

This paper is subdivided into seven sections. An introduction to plant leaf recognition has been presented in Section 1.

The associate editor coordinating the review of this manuscript and approving it for publication was Md. Asikuzzaman.

Section 2 presents a review of existing techniques. The diagram of the proposed system is depicted in Section 3. The exploratory outcomes are portrayed in Section 4. Furthermore, a performance comparison with state-of-the-artwork is made in Section 5. Finally, Section 6 provides the conclusions and some notes on the future research work topics.

II. RELATED WORK

Numerous approaches have been proposed to recognize plant leaves in an automatic manner. A large portion of such attempts used the feature extraction from the leaf, trailed via training a model based on these features. The shape, color and textural features are widely used for feature extraction and classification. The major contributions are as follows:

A. SHAPE FEATURES BASED CLASSIFICATION

Im *et al.* [3] presented a system of representing the leaf shapes by their polygonal approximations and used an expanding number of nearby points of interest for consequent steps. Pietikäinen *et al.* [4] used shape and texture features for leaf classification using neural networks and achieved 83.00% accuracy. Kulkarni *et al.* [5] proposed technique based on color, texture, vein and shape features combined with Zernike moments. Radial Basis Probabilistic Neural Network (RBPNN) classifier was used for classification. Prasvita and Herdiyeni [6] used shape features and neural network classifier and acquired a classification accuracy of 90.00%. Ekshinge and Andore [7] achieved 85% accuracy by elliptic Fourier analysis using shape features.

B. MORPHOLOGICAL FEATURES BASED CLASSIFICATION

There are some popular methods that have been used to extract the information of leaf which include digital morphological features. Neto *et al.* [8] employed PNN together with image and data processing techniques. The author considered five categories of geometrical features, namely, perimeter, physical width, length, area and diameter of leaves for recognition. Wu *et al.* [9] used morphological features with a PNN classifier to characterize 32 types of green leaves. Later, external characteristics of leaf such as leaf shape, venation, leaf margin, and texture were used for plant morphological research [10], [11]. [12] has employed Multilayer Perceptron (MLP) classifier for leaf recognition. The proposed classifier obtained recognition accuracy of 94.0% for leaf recognition using morphological features. Kadir *et al.* [2] produced good results with a classification accuracy of 93.75%. ArunPriya *et al.* [13] used morphological features, geometric features, vein structure features with SVM classifier and achieved 94.20% accuracy.

C. COLOR-FEATURE BASED CLASSIFICATION

Timmermans and Hulzebosch [14] presented a neural network system for the successful classification of the cactus plant. Perez *et al.* [15] used color and geometrical features to recognize weeds in crop fields. The K-NN classifier was used for classification. Fuzzy logic decision making was also employed to recognize weeds in an agricultural field (Yang *et al.* [16]). Zulkifli [17] proposed a general regression neural network to classify 10 different species of plants with leaves of different green shades. A couple of leaf classification frameworks considered surface components like entropy and homogeneity (Man *et al.* [18]) to enhance the accuracy of detection. A similar approach to using color information was proposed for plant recovery by Kebabci *et al.* [19]. Principal Component Analysis (PCA) was used to obtain preeminent features and a higher precision was obtained for characterizing 60 sorts of plants. Anami *et al.* [20] proposed a leaf classification system based on edge histogram, color histogram, and area of leaves. A recognition accuracy of 93.6% was obtained for the proposed system. Bama *et al.* [21] proposed an efficient content-based leaf image retrieval method using texture and color features.

D. TEXTURE FEATURES BASED CLASSIFICATION

Man *et al.* [18] proposed shading as a major component for leaf recognition. The author claimed that the proposed framework could classify 24 species of plants with a precision rate of 92.2%. Chaki and Parekh [22] used texture features using GLCM and obtained a classification accuracy of 78.00%. Chaki *et al.* [23] used preprocessing to make the leaf image invariant to translation, rotation, and scaling. A recent plant species detection algorithm by Zhang *et al.* [24], [25] used a two-stage local similarity-based classification learning method. This method performed classification based on cluster analysis.

E. OTHER APPROACHES

Abbasi *et al.* [26] used the Curvature Scale Space (CSS) technique and k-NN classifier to classify chrysanthemum leaves. Recently, Lee *et al.* [27] investigated the use of deep learning model for leaf classification. The author obtained interesting and surprising results. The different orders of venation came out to be the best representative instead of shape. Moreover, the author used multi-level representation in leaf data corresponding to species classes.

The existing machine learning-based approaches are mainly dependent on the shape and texture features. Our algorithm achieved comparable accuracy used morphological features even without using textural features. It is observed that most of the algorithms has used either PNN or SVM. The K-NN, decision tree and multilayer perceptron with the Adaboost technique are explored which achieved better accuracy with a lower dimensionality as compared to the current algorithms.

III. DESIGN OF THE PROPOSED SYSTEM

The structure of the proposed plant leaf recognition system is delineated in Fig. 1. The proposed framework consists of different stages, specifically, data acquisition, digitization, pre-processing, feature extraction and classification based on the extracted features.

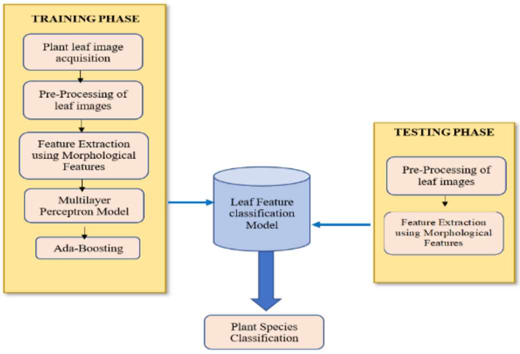
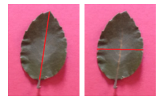


FIGURE 1. Block diagram of the proposed model.

Explanation:

- In data acquisition and digitization phase the samples of plant leaves are collected and digital images of these samples are produced. The current work employed a public dataset of leaf images, which consists of the acquired digital images of different plant leaves.
- In the pre-processing phase, digital images of leaves are converted into a grayscale format.
- In the feature extraction phase, the morphological features, namely, major axis, minor axis, centroid, solidity, perimeter, and orientation are extracted.
- For classification, k-NN, decision tree, and multilayer perceptron classifier are used.
- The AdaBoost methodology is used to improve the precision rate of the proposed system

The proposed pseudo-code and algorithm for plant species detection via leaf recognition is as follows:

Step 1. Read the leaf image	
Step 2: Convert RGB image to grayscale	
Step 3. Extraction of features Major axis and Minor axis $M = \frac{(x_1 - x_c)^2}{rx^2} + \frac{(y_1 - y_c)^2}{ry^2}$	
Centroid $C_x = \frac{\sum C_{ix} A_i}{\sum A_i}, C_y = \frac{\sum C_{iy} A_i}{\sum A_i}$	
Boundary extraction for calculation of Area, Perimeter, Solidity $S = \frac{\text{Area}}{\text{Convex area}}$ $P = 2L + 2W$	
Step 4: Training and testing	
Step 5: Adaboosting	

IV. FEATURE EXTRACTION TECHNIQUES

Leaf classification is being carried out via computational models of leaf recognition methods. Outer qualities of leaf, for example, leaf shape, venation, edge, vein, shape, skeleton, and surface are being utilized for plant morphological research [2], [10], [11]. Numerous specialists have applied rice seed morphological highlights to distinguish and investigate rice seed quality [28]–[30]. In the present paper, morphological features such as major and minor axis length, centroid, solidity, perimeter, and orientation are extracted for leaf recognition. They are defined in the following section:

The major axis is the line connecting at one end called a base point to the tip of the leaf. For drawing the major axis two points are selected. Then the line will be drawn on to the selected points that represent the major perpendicular axis of the image. This major axis length measures the length of the image in width-wise as follows.

$$\text{Major axis length} = \frac{(x_1 - x_c)^2}{rx^2} + \frac{(y_1 - y_c)^2}{ry^2} \quad (1)$$

where x_1, y_1 is the point along the major axis and x_c and y_c are the center point. rx, ry is the radius along x-axis and y-axis, respectively. **Minor Axis** is the line drawn perpendicular to the major axis.

Centroid is defined as the center of mass of the region. An image can be sub-divided into various small regions. Each small region can have its individual centroid point. The centroid of a polygon is calculated using (2)

$$C_x = \frac{\sum C_{ix} A_i}{\sum A_i}, \quad C_y = \frac{\sum C_{iy} A_i}{\sum A_i} \quad (2)$$

where C_x and C_y are the centroids for where A_i is the contour area given by (3)

$$A = \frac{1}{2} \sum_{i=0}^{N-1} \left(x_i y_{i+1} - x_{i+1} y_i \right) \quad (3)$$

Solidity is the extent of the pixels in the convex hull that is additionally in the area. It is computed as equation (4)

$$S = \frac{\text{Area}}{\text{Convex area}} \quad (4)$$

Generally, there are two types of leaf images hollow and solidly filled image. The solidly filled image always consists of a single color. All the pixels around the center of mass are filled with high-intensity colors. In a hollow image, the pixels around the center of mass are only partially filled. That means small pixels are left empty.

$$\frac{c}{s} = \frac{10}{(D_h/D_t)\,(N_s/1000)^{1.5}} \quad (5)$$

C is the centroid, *S* is the solidity, D_h is the diameter, and N_s is the pixel count in a specific area. When all the centroid points are connected, it will generate a clearly drawn line. This line will be random and the whole image center of mass stands around this centroid point.

Perimeter (or circumference) of a region R is characterized as the length of its external shape, where R must be associated [28]. The perimeter is determined by estimating the whole of the separations between progressive limit pixels as in equation (6) [30]. The simplest measure of the perimeter is obtained by tallying the number of limit pixels that belong to an object [10]. The separation around the limit of the region is called a perimeter. It is the total circumference around the image of the leaf. The total number of pixels around the boundary points is calculated which will give information about the total amount of pixels that have been used to fill the boundary pixels.

$$P = 2L + 2W \tag{6}$$

where P is the perimeter, L is the length of the major axis and W is the length of the minor axis.

Orientation is the angle between the x-axis and the major axis of the ellipse. It denotes the alignment of the image along with the major axis and minor axis. Orientation along the coordinate axis will automatically shorten the length of major and minor axis as defined in equation (7).

$$O = \cos\left(m_j\right) + \sin\left(m_i\right) + \mathrm{sqrt}\left(\tan\left(m_j \times m_i\right)\right) \tag{7}$$

where O is the orientation, m_i is the length of the minor axis, and m_j is the length of the major axis.

V. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this section, experimental results dependent on the proposed framework are introduced. In this paper, a public dataset taken from http://flavia.sourceforge.net/ is used for experimentation. This dataset contains 32 different types of plant leaves. This dataset has leaf images of 32 common plants in China, such as Phyllestachys Pubescens, Aesculus Chinensis, Berberis Ferdinandi-coburgii Schneid etc., clicked using Apple iPad 2 device. It comprises 1907 images 720 × 960 pixels for all 32 categories. The size of the dataset is about 1 GB. In this work, the authors have considered 10 images of each category for the experimental work. Three different approaches were used for evaluation. First 80-20 approach is used in which 80% images are considered randomly as training dataset and remaining 20% are considered as testing dataset. Another approach is 3-fold and 5-fold cross-validation. In 3 fold cross-validation, the whole dataset is randomly partitioned into three groups. Training is done on two groups and testing is done on the third group. A similar approach is used for 5 fold cross-validation. Python platform is used for the experimentation on intel i3 with 8GB RAM machine. Various classifiers i.e. k-NN, decision tree, and Multilayer perceptron classifiers are used for evaluation. K-NN classifier using one nearest neighbor is used for experimentation. In multilayer perceptron number of layers is calculated as average of sum of total number of features and classes i.e.

$$\text{No. of layers} = \frac{(\text{number of features} + \text{numberofclasses})}{2}$$

TABLE 1. Precision rate of proposed leaf recognition system.

Classification Technique	80-20*	3-Fold CV**	5-Fold CV
k-NN	91.62%	91.40%	92.19%
Decision Tree	76.22%	81.37%	86.28%
ML perceptron	94.89%	94.11%	95.38%
AdaBoost_MLP	95.08%	94.85%	95.42%

TABLE 2. RMSE of proposed leaf recognition system.

Classification Technique	80-20*	3-Fold CV**	5-Fold CV**
k-NN	12.09%	10.50%	9.83%
Decision Tree	17.74%	14.87%	13.7%
ML perceptron	11.28%	10.10%	9.35%
AdaBoost_MLP	11.16%	9.96%	9.24%

TABLE 3. Far of proposed leaf recognition system.

Classification Technique	80-20*	3-Fold CV**	5-Fold CV**
k-NN	0.5%	0.6%	0.4%
Decision Tree	1.9%	1.7%	1.2%
ML perceptron	0.4%	0.6%	0.4%
AdaBoost_MLP	0.4%	0.5%	0.3%

*80% Training Set and 20% Testing Set
** Cross Validation

TABLE 4. Accuracy of ML perceptron.

Classification Technique	80-20*	3-Fold CV**	5-Fold CV**
ML perceptron	94.89%	94.11%	95.38%

*80% Training Set and 20% Testing Set
** Cross Validation

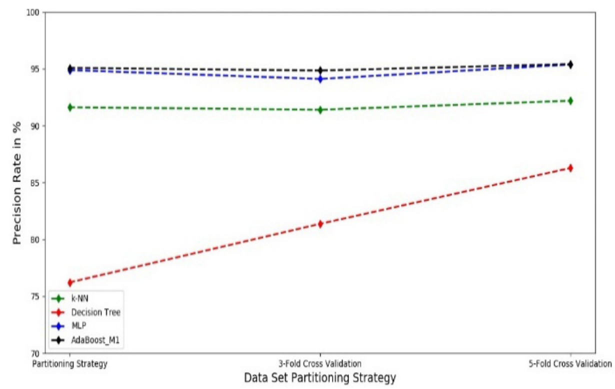


FIGURE 2. Precision rate of the proposed system.

Each layer represents one neuron for every feature. The Adaptive Booting approach is likewise explored with the same dataset. Boosting is an approach to manage machine learning in light of making an exact desire rule by joining numerous tolerably feeble and wrong runs. The AdaBoost calculation of Freund and Schapire [31] is the most generally used boosting calculation with applications in different fields. We have considered this algorithm in the present work to enhance the classifier performance. The base classifier is MLP. Number of estimators used are 50.

TABLE 5. Comparison of proposed machine learning based approach with state-of-the-art.

Features	Authors	Machine Learning Classifier	Accuracy
Shape, Texture	Pietikäinen *et al.* [4]	Neural Network	83.00%
Texture	Chaki and Parekh [23]	GLCM	78.00%
Shape	Prasvita and Herdiyeni [6]	Neural Network	90.00%
Shape	Kulkarni *et al.* [5]	RBPNN	93.80%
Shape	Ekshinge and Andore [7]	Elliptic Fourier analysis	85.00%
Edge, Color, Area	Anami *et al.* [20]	Neural Network	93.60%
Texture, Shape, Color	Ananthi *et al.* [32]	Probabilistic Neural Network	56.30%
Shadow features	Ehsanirad and Kumar [33]	Radial basis exact fit neural network	90.00%
Wavelet transforms GLCM, geometric features	Zhang *et al.* [24]	SVM	93.10%
Morphological features	Neto *et al.* [8]	Probabilistic Neural Network	90.00%
Morphological features, Texture	Kadir *et al.* [2]	Probabilistic Neural Network	93.75%
Morphological features, geometric features, vein structure	ArunPriya *et al.* [13]	SVM	94.20%
Morphological features	Kumar [12]	Neural Network	94.00%
Histograms of oriented gradients (HOG)	Olsen *et al.* [34]	Gaussian SVM	94.70%
Morphological Features	Proposed System	Multilayer Perceptron	95.38%
Morphological Features	Proposed System	Multilayer Perceptron & AdaBoost	95.40%

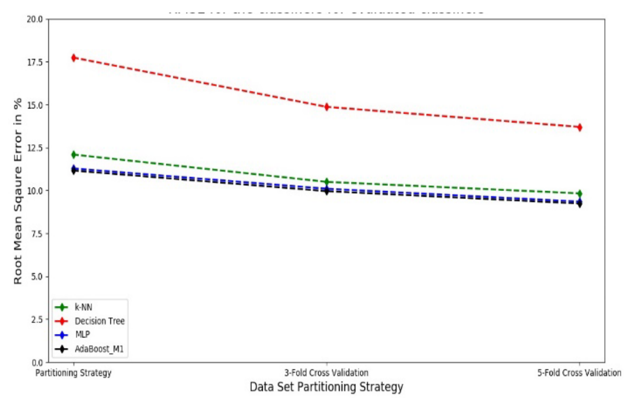


FIGURE 3. RMSE of the proposed system.

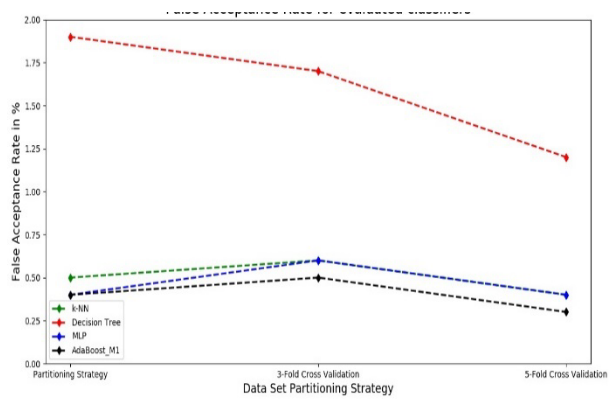


FIGURE 4. False acceptance rate of the proposed system.

The Performance of the proposed framework for plant leaf recognition is analyzed based on precision rate, RMSE (Root Mean Square Error) and FAR (False Acceptance Rate). Experimental results based on these parameters are illustrated in TABLE 1-4. In TABLE 1, the precision rate of the proposed leaf recognition system is presented. It is clear that the MLP and Adaboost_MLP are performing much better than k-NN and Decision tree classifiers.

Table 2 and 3 depict the performance of the proposed system based on RMSE and FAR respectively. Again, the MLP's and adaboost's RMSE and FAR are lower as compared to another classifier which clearly indicates that they are performing better. As MLP and Adaboost_MLP are comparative, the accuracy for Adaboost_MLP is marginally higher than MLP (Table 4). It has been observed that the maximum accuracy of 95.40% is achieved using the AdaBoost methodology with a 5-fold MLP classifier. The Precision, RMSE and FAR are plotted in Figs. 2-4 for analysis. Confusion matrix for accuracy of 95.40% using AdaBoost methodology with MLP classifier is presented in Fig. 5.

VI. COMPARATIVE STUDY OF PROPOSED SYSTEM AND STATE-OF-THE-ART METHODS

In this section, the authors have presented a comparative study of proposed work and state-of-the-artwork. A good number of techniques for plant species recognition are compared with the proposed technique. The analytical comparison based on technique, classifier used, and accuracy achieved has been presented in TABLE 5. Our proposed

```
  a  b  c  d  e  f  g  h  i  j  k  l  m  n  o  p  q  r  s  t  u  v  w  x  y  z aa ab ac ad ae af   <-- classified as
  2  0  0  0  0  0  0  0  0  0  0  0  0  1  4  0  0  0  0  0  0  0  0  0  0  0  0  1  2  0  0  0 |  a = pubescent bamboo
  0  8  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  2  0  0 |  b = Chinese horse chestnut
  0  0  8  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  1  0 |  c = Anhui Barberry
  0  0  0  5  0  0  0  0  0  0  0  0  0  0  0  0  0  2  1  0  0  0  0  0  0  0  0  0  0  0  0  2 |  d = Chinese redbud
  0  0  0  0  9  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0 |  e = true indigo
  0  0  0  0  0 10  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 |  f = Japanese maple
  0  0  0  0  0  0 10  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 |  g = Nanmu
  0  0  0  0  0  0  0 10  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 |  h = castor aralia
  0  0  0  0  0  1  0  0  9  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 |  i = Chinese cinnamon
  0  0  0  0  0  0  0  0  0  9  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 |  j = goldenrain tree
  0  0  0  0  0  0  0  0  0  1  9  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 |  k = Big-fruited Holly
  0  0  0  0  0  0  0  0  0  0  0  9  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0 |  l = Japanese cheesewood
  0  0  0  0  0  0  0  0  0  0  0  0  6  1  0  0  0  0  0  0  0  0  0  0  0  0  2  1  0  0  0  0 |  m = wintersweet
  1  0  0  0  0  0  0  0  0  0  0  0  0  9  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 |  n = camphortree
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  9  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0 |  o = Japan Arrowwood
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  9  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0 |  p = sweet osmanthus
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 10  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 |  q = deodar
  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  8  0  0  0  0  1  0  0  0  0  0  0  0  0  0 |  r = maidenhair tree
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 10  0  0  0  0  0  0  0  0  0  0  0  0  0 |  s = Crape myrtle
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  9  0  0  0  0  1  0  0  0  0  0  0  0 |  t = oleander
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  9  0  0  0  0  0  0  0  0  0  0  0 |  u = yew plum pine
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 10  0  0  0  0  0  0  0  0  0  0 |  v = Japanese Flowering Cherry
  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  8  0  0  0  0  0  0  0  0  1 |  w = Glossy Privet
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  9  0  0  0  0  0  0  0  0 |  x = Chinese Toon
  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  9  0  0  0  0  0  0  0 |  y = peach
  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  9  0  0  0  0  0  0 |  z = Ford Woodlotus
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 10  0  0  0  0  0 | aa = trident maple
  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  9  0  0  0  0 | ab = Beales barberry
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 10  0  0  0 | ac = southern magnolia
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 10  0  0 | ad = Canadian poplar
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 10  0 | ae = Chinese tulip tree
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 10 | af = tangerine
```

FIGURE 5. Confusion matrix false acceptance rate of the proposed system.

approach is statistically significantly significant than other existing approaches as depicted in TABLE 5.

As illustrated in TABLE 5, most of the proposed techniques used the shape, texture, and morphological features. Out of these, shape-based features [4]–[7] were used at earlier stages with good accuracy to start with. After that, the texture features [4], [23], [33] were added. However, these features do not contribute much to higher accuracy. Later, edge, color, and area-based classification [20] is proven to be much more accurate in comparison. Wavelets and texture [25] of a leaf with geometric features are also explored. Unfortunately, due to the redundant set of features, it affected the overall accuracy. The morphological features [8], [12], [13] improved the accuracy further in comparison to all the other existing features. Their combination with SVM is claimed to provide the best accuracy of 94.20%. The proposed system explored the possibility of accuracy improvement for species detection with morphological features in combination with different classifiers. Moreover, the system can be enhanced using the Adaboost technique to obtain an accuracy of 95.40%. Thus, it has achieved the highest accuracy as compared to all the machine learning-based state-of-the-art methodologies.

VII. CONCLUSION AND FUTURE SCOPE

In this paper, an efficient plant leaf recognition system using morphological features and adaptive boosting methodology has been presented. Experimental results are performed using three different classification techniques, namely, k-NN, decision tree, and multilayer perceptron. The AdaBoost methodology is considered to improve the precision rate of the proposed system. In our work, the maximum precision rate of 95.42% has been achieved for 32 kinds of plant leaves. The authors have observed that the proposed system performed better than the existing techniques for plant leaf recognition in agricultural research. In the future work, our model can be extended for use in the fields of herbal cosmetic industry and natural corrective industry.

COMPLIANCE WITH ETHICAL STANDARDS

The authors herewith declare that they have no conflict of interest. A public dataset downloaded from http://flavia.sourceforge.net/ is considered for experimental work.

REFERENCES

[1] J. Hossain and M. A. Amin, "Leaf shape identification based plant biometrics," in *Proc. 13th Int. Conf. Comput. Inf. Technol. (ICCIT)*, Dec. 2010, pp. 458–463.

[2] A. Kadir, L. Nugroho, A. Susanto, and I. P. Santosa, "Performance improvement of leaf identification system using principal component analysis," *Int. J. Adv. Sci. Technol.*, vol. 44, pp. 113–124, May 2014.

[3] C. Im, H. Nishida, and T. L. Kunii, "A hierarchical method of recognizing plant species by leaf shapes," in *Proc. IAPR Workshop Mach. Vis. Appl.*, Nov. 1998, pp. 158–161.

[4] M. Pietikäinen, T. Mäenpää, and J. Viertola, "Color texture classification with color histograms and local binary patterns," in *Proc. Workshop Texture Anal. Mach. Vis.*, 2002, pp. 109–112.

[5] A. H. Kulkarni, H. M. Rai, K. A. Jahagirdar, and P. S. Upparamani, "A leaf recognition technique for plant classification using RBPNN and Zernike moments," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 2, no. 1, pp. 984–988, 2013.

[6] D. S. Prasvita and Y. Herdiyeni, "MedLeaf: Mobile application for medicinal plant identification based on leaf image," *Int. J. Adv. Sci., Eng. Inf. Technol.*, vol. 3, no. 2, pp. 5–8, 2013.

[7] S. Ekshinge, I. Sambhaji, and M. D. Andore, "Leaf recognition algorithm using neural network based image processing," *Asian J. Eng. Technol. Innov.*, vol. 2, no. 2, pp. 10–16, 2014.
[8] J. C. Neto, G. E. Meyer, D. D. Jones, and A. K. Samal, "Plant species identification using Elliptic Fourier leaf shape analysis," *Comput. Electron. Agricult.*, vol. 50, no. 2, pp. 121–134, 2006.
[9] S. G. Wu, F. S. Bao, E. Y. Xu, Y.-X. Wang, Y.-F. Chang, and Q.-L. Xiang, "A leaf recognition algorithm for plant classification using probabilistic neural network," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol.*, Dec. 2007, pp. 11–16.
[10] J. S. Cope, D. Corney, J. Y. Clark, P. Remagnino, and P. Wilkin, "Plant species identification using digital morphometrics: A review," *Expert Syst. Appl.*, vol. 39, no. 8, pp. 7562–7573, 2012.
[11] M. F. Bong, G. B. Sulon, and M. S. M. Rahim, "Recognition of leaf based on its tip and base using centroid contour gradient," *Int. J. Comput. Sci. Issues*, vol. 10, no. 2, pp. 477–482, 2013.
[12] S. Kumar, "Leaf color, area and edge features based approach for identification of Indian medicinal plants," *Indian J. Comput. Sci. Eng.*, vol. 3, no. 3, pp. 436–442, 2012.
[13] C. A. Priya, T. Balasaravanan, and A. S. Thanamani, "An efficient leaf recognition algorithm for plant classification using support vector machine," in *Proc. Int. Conf. Pattern Recognit., Inform. Med. Eng. (PRIME)*, Mar. 2012, pp. 428–432.
[14] A. J. M. Timmermans and A. A. Hulzebosch, "Computer vision system for on-line sorting of pot plants using an artificial neural network classifier," *Comput. Electron. Agricult.*, vol. 15, no. 1, pp. 41–55, 1996.
[15] A. J. Pérez, F. López, J. V. Benlloch, and S. Christensen, "Color and shape analysis techniques for weed detection in cereal fields," *Comput. Electron. Agricult.*, vol. 25, no. 3, pp. 197–212, 2000.
[16] C. C. Yang, S. O. Prasher, J. A. Landry, J. Perret, and H. S. Ramaswamy, "Recognition of weeds with image processing and their use with fuzzy logic for precision farming," *Can. Agricult. Eng.*, vol. 42, no. 4, pp. 195–200, 2000.
[17] Z. Zulkifli, "Plant leaf identification using moment invariants & general regression neural network," M.S. thesis, Univ. Teknologi Malaysia, Johor Bahru, Malaysia, 2009.
[18] Q.-K. Man, C.-H. Zheng, X.-F. Wang, and F.-Y. Lin, "Recognition of plant leaves using support vector machine," in *Proc. Int. Conf. Intell. Comput.*, 2008, pp. 192–199.
[19] H. Kebapci, B. Yanikoglu, and G. Unal, "Plant image retrieval using color, shape and texture features," *Comput. J.*, vol. 54, no. 9, pp. 1475–1490, Sep. 2011.
[20] B. S. Anami, S. S. Nandyal, and A. Govardhan, "A combined color, texture and edge features based approach for identification and classification of Indian medicinal plants," *Int. J. Comput. Appl.*, vol. 6, no. 12, pp. 45–51, 2010.
[21] B. S. Bama, S. M. Valli, S. Raju, and V. A. Kumar, "Content based leaf image retrieval (CBLIR) using shape, color and texture features," *Indian J. Comput. Sci. Eng.*, vol. 2, no. 2, pp. 202–211, 2011.
[22] J. Chaki and R. Parekh, "Plant leaf recognition using shape based features and neural network classifiers," *Int. J. Adv. Comput. Sci. Appl.*, vol. 2, no. 10, pp. 41–47, 2011.
[23] J. Chaki, R. Parekh, and S. Bhattacharya, "Plant leaf recognition using texture and shape features with neural classifiers," *Pattern Recognit. Lett.*, vol. 58, pp. 61–68, Jun. 2015.
[24] H. Zhang, P. Yanne, and S. Liang, "Plant species classification using leaf shape and texture," *Amer. J. Eng. Technol. Res.*, vol. 11, no. 9, 2011.
[25] S. Zhang, H. Wang, and W. Huang, "Two-stage plant species recognition by local mean clustering and weighted sparse representation classification," *Cluster Comput.*, vol. 20, no. 2, pp. 1517–1525, 2017.
[26] S. Abbasi, F. Mokhtarian, and J. Kittler, "Reliable classification of chrysanthemum leaves through curvature scale space," in *Proc. Int. Conf. Scale-Space Theories Comput. Vis.*, vol. 1252, 1997, pp. 284–295.
[27] S. H. Lee, C. S. Chan, S. J. Mayo, and P. Remagnino, "How deep learning extracts and learns leaf features for plant classification," *Pattern Recognit.*, vol. 71, pp. 1–13, Nov. 2017.
[28] A. Dell'Aquila, "Application of a computer-aided image analysis system to evaluate seed germination under different environmental conditions," *Italian J. Agronomy*, vol. 8, pp. 51–62, Sep. 2004.
[29] B. K. Yadav and V. K. Jindal, "Monitoring milling quality of rice by image analysis," *Comput. Electron. Agricult.*, vol. 33, no. 1, pp. 19–33, Dec. 2011.
[30] C. V. Maheshwari, "Quality assessment of Oryza sativa SSP indica (rice) using computer vision," *Int. J. Innov. Res. Comput. Commun. Eng.*, vol. 1, no. 4, pp. 1107–1115, 2013.
[31] Y. Freund, R. E. Schapire, and N. Abe, "A short introduction to boosting," *J.-Jpn. Soc. Artif. Intell.*, vol. 14, no. 5, pp. 771–780, 1999.
[32] C. Ananthi, A. Periasamy, and S. Muruganand, "Pattern recognition of medicinal leaves using image processing techniques," *J. Nano Sci. Nano Technol.*, vol. 2, no. 2, pp. 214–218, 2014.
[33] A. Ehsanirad and Y. H. S. Kumar, "Leaf recognition for plant classification using GLCM and PCA methods," *Oriental J. Comput. Sci. Technol.*, vol. 3, no. 1, pp. 31–36, 2010.
[34] A. Olsen, S. Han, B. Calvert, P. Ridd, and O. Kenny, "*In situ* leaf classification using histograms of oriented gradients," in *Proc. Int. Conf. Digit. Image Comput., Techn. Appl. (DICTA)*, Nov. 2015, pp. 1–8.

MUNISH KUMAR received the master's degree in computer science and engineering and the Ph.D. degree from Thapar University, Patiala, India, in 2008 and 2015, respectively. He started his career as an Assistant Professor in computer application at Jaito Centre of Punjabi University, Patiala. He is currently an Assistant Professor with the Department of Computational Sciences, Maharaja Ranjit Singh Punjab Technical University, Bathinda, India. His research interests include character recognition, computer vision, and pattern recognition. He is a Professional Member of IEEE.

SURBHI GUPTA received the B.Tech. degree in computer science and engineering from Punjab Technical University, in 2002, and the master's degree in merit from Punjab Agricultural University, Punjab, India, in 2004. She is awarded a Ph.D. degree in CSE from IKG Punjab Technical University, India, in 2018. She has 15 years of experience in teaching and research. She is currently an Associate Professor with the Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India. She has published 37 articles in reputed international journals and conferences. She has 84 citations for her articles at Google scholar. Her research interests include artificial intelligence, soft computing, machine learning, and digital image processing.

XIAO-ZHI GAO received the B.Sc. and M.Sc. degrees from the Harbin Institute of Technology, China, in 1993 and 1996, respectively, and the D.Sc. (Tech.) degree from the Helsinki University of Technology (now Aalto University), Finland, in 1999. He has been a Professor with the University of Eastern Finland, Finland, since 2018. He is currently a Guest Professor with the Harbin Institute of Technology, Beijing Normal University, and Shanghai Maritime University, China. He has published more than 400 technical articles in refereed journals and international conferences. His current Google Scholar H-index is 30. His research interests are nature-inspired computing methods with their applications in optimization, data mining, machine learning, control, signal processing, and industrial electronics.

AMITOJ SINGH received the master's degree from the Department of Computer Science, Punjabi University, Patiala, and the Ph.D. degree in computer science from the Department of Computer Science, Punjabi University, Patiala. He is currently an Assistant Professor with the Department of Computational Sciences, Maharaja Ranjit Singh Punjab Technical University, Bathinda, India. He has more than 12 more than teaching and research experience. He has published more than 20 research articles and four patents. His research interest includes speech recognition.