Plant Classification Using Artificial Neural Networks

Luciano D. S. Pacifico
Departamento de Computacao - DC
Universidade Federal Rural de Pernambuco - UFRPE
Recife, Pernambuco, Brazil
Email: luciano.pacifico@ufrpe.br, ldsp.ufrpe@gmail.com

Joao F. L. Oliveira
Facul. de Cien. e Tec. de Garanhuns - FACETEG
Universidade de Pernambuco - UPE
Garanhuns, Pernambuco, Brazil
Email: jflo@cin.ufpe.br

***Abstract*—Automatic plant species identification is a difficulty challenge and an interesting area of research for both botanical taxonomy and computer science. From the past few years, some attempts towards the development of automatic plant recognition systems have been proposed, but the performance of such systems is not satisfactory in terms of accuracy, and these systems are also task dependent, since they are strongly influenced by the set of characteristics extracted from plant samples, leading to the problem known as data set bias. In this work, we use a Multi-Layer Perceptron (MLP) artificial neural network trained with Backpropagation algorithm to perform automatic plant classification. To avoid data set bias problem, some plant data sets which use different plant features obtained by different feature extraction processes are employed. We compare MLP algorithm with several supervised learning methods from plant recognition literature using a statistical hypothesis test of type Friedman/Nemenyi test. The obtained results show the potential of MLP algorithm to deal with plant classification in a unbiased fashion.**

I. Introduction

Plants play an important role in Earth, providing valuable resources such as oxygen, and they also play a irreplaceable role in the ecological balance. Some kinds of plants are used as food, clothing, paper, fuel, shelter, and so on. Given the vast range of applications for many plant species, plant identification has become an interesting and challenging task for both science and industry. But even today, identification and classification of unknown plant species are performed manually by expert personnel, making plant identification a time consuming process prone to human failures.

With the evolution of technologies, more and more applications are utilizing the benefits of highly advanced technologies such as Artificial Intelligence and Computer Vision. Automatic plant recognition is the most promising solution towards bridging the botanical taxonomic gap, as it automates the human task of plant identification with the use of Image Processing and Pattern Recognition methods.

Automatic plant recognition systems consist of two main modules: feature extraction module and recognition module [1]. Features typically considered for plant identification are: color, leaf features (such as margin, shape, texture and venation), flower features, seed features and other organs features [2], [3]. For recognition module, the most common classifiers employed for plant recognition are Decision Tree classifier [4], Naive Bayes classifier [4], K-Nearest Neighbors [5], [6], [4], [7], Support Vector Machine [8] and Artificial Neural Networks [1].

Many works on plant identification literature focus on feature extraction techniques [8], [9], [7], leaving the choice of the classier as a secondary task, what may compromise the accuracy obtained by these systems. In general, these works lack the proper evaluation criteria. Many of these works do not even provide any form of statistical analysis to validate the obtained results. Also, it is common to find works that do not furnish any kind of comparison among the proposed classifier and other classifiers from literature (they only present the obtained results for the models being proposed), like in [8], [1], [10]. The performance of such systems is highly dependent on the chosen set of characteristics, which are task or data set dependent, what may lead to data set bias problem [11].

In this work, we extend the analysis performed by Rahmani et al. [4] by employing a Multi-Layer Perceptron Artificial Neural Network trained with Backpropagation algorithm to deal with plant identification task. Artificial Neural Networks (ANNs) are known as universal approximators and computational models with particular characteristics, such as adaptability, capacity of learning by examples and the ability to organize or to generalize data. Also, to avoid data set bias problem, two real-world plant data sets (Iris and Wheat Seeds) obtained from UCI Machine Learning Repository are used, where different plant features are taken into consideration for each data set, and the data sets are obtained from variant feature extraction processes. We compare the performance obtained by MLP algorithm in relation to other four well-established supervised learning methods from Machine Learning and Plant Classification literature (including Support Vector Machine).

The work is organized as follows. Section II presents some related works in the field of plant recognition. After that (Section III), the adopted data sets are presented. The following section (Section IV) presents the selected classifiers employed in the experiments. The experimental results are discussed in Section V. Finally, some conclusions and leads to future works are given (Section VI).

II. Related Works

From the past decades, plant classification has become an area of major interest for researchers, but little efforts have been conducted up to now in the field. One of the first attempts towards the development of a digital plant identification system was proposed by Agarwal et al. [12].

Leafsnap [8] was proposed by Kumar et al. as another plant identification system, where plants were identified based on curvature-based shape features of the leaf by utilizing integral measure to compute functions of the curvature at the boundary. Leafsnap adopts K-Nearest Neighbor algorithm as the classifier.

K-Nearest Neighbors is also adopted as the classifier in Sahay and Chen [7], where local features are extracted using Scale Invariant Feature Transform (SIFT) from plant leaf images.

In [5] and [6], K-Nearest Neighbors variants are employed as the classifier for the proposed plant identification systems. Their classification methods were based on three leaf features: a fine-scale margin feature histogram, a Centroid Contour Distance Curve shape signature, and an interior texture feature histogram. Each feature is represented by a 64-dimensional vector.

Adopting the same data set proposed by [6], Rahmani et al. [4] evaluated the performance of Decision Tree classifier, Naive Bayes classifier and K-Nearest Neighbors (with $k = 3, 4, 5, 6, 7$) for each individual feature vector, for each combination of two feature vectors and for all three feature vectors. The experimental results pointed out that K-Nearest Neighbors with $k = \{3, 4\}$ are able to obtain the better average performances in terms of accuracy in relation to the other comparison classifiers.

Wang et al. [13] proposed a plant leaf recognition method using Intersection Cortical Model (ICM) and Support Vector Machines. The proposed approach used two leaf features represented by the Entropy Sequence (EnS) extracted by ICM and the Center Distance Sequence (CDS).

Jin et al [9] adopted a Sparse Representation-based classifier and four leaf tooth characteristics (leaf-num, leaf-rate, leaf-sharpness and leaf-obliqueness) to perform automatic plant classification.

Some interesting reviews on plant recognition are presented in Cope et al. [2], Sethulekshmi and Sreekumar [14] and Sabu and Sreekumar [3].

III. Data sets

In this section, the adopted data sets are presented. All data sets are real-world problems obtained from UCI Machine Learning Repository [15].

Fisher's Iris Plant [16] is composed of 150 samples equally distributed in three classes, where each class refers to a type of iris plant. Each instance is described by a set of four features: sepal length (in cm), sepal width (in cm), petal length (in cm) and petal width (in cm). The sample values for each feature have been manually collected at the same day and using the same instruments by Edgar Anderson [17].

The Wheat Seed Kernels data set [18] comprises 210 randomly selected samples equally distributed in three varieties of wheat (Kama, Rosa and Canadian). The features have been extracted using a soft X-ray technique to detect the visualization of the internal kernel structure for each sample seed.

The 100 Plant Leaves data set [5] comprises one-hundred species of leaves (problem classes). The data set contains 1600 instances, and for each species, there are sixteen distinct specimens, photographed as a color image on a white background. For each sample, three distinct features have been extracted: a Centroid Contour Curve shape signature (Sha), an interior texture feature histogram (Tex), and a fine-scale margin feature histogram (Mar). Each feature is represented by a 64-dimensional vector. A sample set of the binary plant images from 100 Plant Leaves data set is presented in Figure 1.

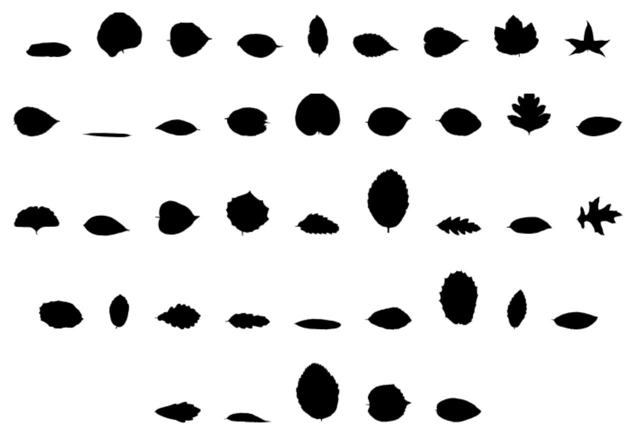

Fig. 1. A sample set of silhouette binary images of some plant specimen from 100 Plant Leaves data set [15], [5]

IV. Classifiers

This section presents a brief description for all classifiers employed in experiments: K-Nearest Neighbors (K-NN), Decision Tree classifier (DT), Naive Bayes classifier (NB), Support Vector Machine (SVM) and the proposed Multi-Layer Perceptron trained with Backpropagation algorithm (MLP-BP).

K-NN its classification process is obtained by the association of unlabeled samples to the majority class among the classes of the closest k training patterns in the feature space. Typically, as a manner to avoid ties in binary classification problems, k value is set to an odd integer, but there can still be ties when k is an odd integer when dealing with multiclass classification problems. The main parameters for K-NN are the number k of training patterns to be considered to classify the new sample and the distance metric.

DT is a method for approximating discrete-valued target functions, in which the learning function is represented by a decision tree [19]. The decision process is performed by splitting the data set according to binary questions performed on one selected feature each time, generating a hierarchical structure represented by a tree with decision nodes and leaf nodes. The splitting process is guided by a measure of satisfaction (the entropy) using a greedy approach to decide which feature will be taken into consideration to divide the data set in a given iteration of the method.

NB is a probabilistic classifier which applies the Bayes' theorem with the strong (naive) assumption that the features describing the object samples to be classified are statistically independent from each other [20]. When a new sample to be classified is presented to the algorithm, its probability to belong in anyone of the problem classes is estimated according to the resemblance of the new pattern to the training samples belonging in that class.

SVMs are based on the concept of decision planes that define decision boundaries. A decision plane is one that separates between a set of objects having different class memberships. Given a set of training examples, each marked as belonging in one or the other of two categories, a SVM training algorithm builds a model that assigns new samples to one category or the other, making it a non-probabilistic binary linear classifier. In addition to performing linear classification, SVMs can efficiently perform a non-linear classification according to the selected kernel function.

A Multi-Layer Perceptron (MLP) is a feedforward ANN composed by an input layer, at least one hidden layer, and an output layer of computational nodes (neurons). MLPs are known to be logistic regression classifiers where the input signals are projected into a space where they become linearly separable. One of the most popular methods for MLP training is the Backpropagation (BP) algorithm [21], [22].

V. Experimental Results

In this section, the experimental results are presented. We compare five different classifiers from plant classification literature and some of their variants: Decision Tree classifier (DT), Naive Bayes classifier (NB), K-Nearest Neighbors (K-NN, with $k = 3, 4$ and 5), Support Vector Machine with RBF (SVM$_{rbf}$) and $Linear$ (SVM$_{linear}$) kernel functions and a Multi-Layer Perceptron trained with Backpropagation algorithm (MLP-BP). We opt not to include the tests with K-NN using $k = 6$ and 7 as in Rahmani et al., since the experimentation performed in [4] showed that the best accuracies for K-NN have been obtained with $k = 3, 4$ and 5. All algorithms are implemented in Python programming language, and all tests have been executed in a computer with an i5-5250U CPU and 8 GB RAM. DT, NB, K-NN, SVM and MLP-BP are implemented using *scikit-learn library* [23], [24]. All algorithms executed with scikit-learn default configurations, except for SMV, where two different kernel functions have been applied. We also used 1000 epochs for MLP-BP training.

For the experiments, three benchmark plant data sets obtained from UCI Machine Learning Repository are employed: Iris, Seeds and 100 Plant Leaves (see Section III). 100 Plant Leaves is decomposed in seven data sets, so we could evaluate the influence of each plant leaf feature (margin, shape and texture) and their combinations on the behavior of the adopted classifiers: Margin, Shape, Texture, Margin and Shape, Margin and Texture, Shape and Texture, and Margin, Shape and Texture [4]. For the experiments, all data sets are divided in training and testing sets, according to the values presented in Table I.

TABLE I
Benchmark Plant Data Sets Description. Attrib.: number of features; Classes: number of classes; Train.: number of training samples; Test.: number of testing samples; Total: total number of data patterns.

Data set	Attrib.	Classes	Train.	Test.	Total
Iris	4	3	105	45	150
Seeds	7	3	145	65	210
Margin	64	100	1100	500	1600
Shape	64	100	1100	500	1600
Texture	64	100	1100	500	1600
Margin and Shape	128	100	1100	500	1600
Margin and Texture	128	100	1100	500	1600
Shape and Texture	128	100	1100	500	1600
Mar., Sha. and Tex.	192	100	1100	500	1600

For the experiments, one hundred independent tests are performed for each data set, and for each test, the data patterns are randomly distributed into training and testing sets, but for each iteration of the experiments, each algorithm is tested using the same training and testing sets.

The evaluation includes an empirical analysis concerning the average accuracy for the testing set and the execution time for each data set. The evaluation also includes a rank system employed through the application of Friedman test [25], [26] to the overall average testing accuracies. The Friedman test is a non-parametric hypothesis test that ranks all algorithms for each data set separately. If the null-hypothesis (all ranks are not significantly different) is rejected, Nemenyi test [27] is adopted as the *post-hoc* test. According to Nemenyi test, the performance of two algorithms are considered significantly different if the corresponding average ranks differ by at least the critical difference

$$CD = q_a \sqrt{\frac{n_{alg}(n_{alg}+1)}{6n_{data}}} \quad (1)$$

where n_{data} represents the number of data sets, n_{alg} represents the number of compared algorithms and q_a are critical values based on a Studentized range statistic divided by $\sqrt{2}$ [28]. Given that the overall average test accuracy is a maximization measure, the best ranked algorithms for the Friedman/Nemenyi test will obtain high ranks. Once our experiments are executed with $n_{data} = 9$ and $n_{alg} = 8$, we have $CD = 3.4998$.

The experimental results are shown in Table II.

From Table II, in an empirical analysis, we can observe that MLP-BP obtained the best average performances for all data sets, being at least as good as the best comparison algorithms. The experiments show the potential of MLP-BP as a classifier for plant recognition independent from the kind of features taken into consideration when modeling the system. For Iris

TABLE II
EXPERIMENTAL RESULTS FOR ALL PLANT CLASSIFICATION DATA SETS. MEAN: AVERAGE ACCURACY FOR THE TEST SET; STD: STANDARD DEVIATION FOR THE TEST SET; TIME: AVERAGE EXECUTION TIME IN SECONDS.

Data set	Metric	Algorithm							
		DT	NB	K-NN$_3$	K-NN$_4$	K-NN$_5$	SVM$_{linear}$	SVM$_{rbf}$	MLP-BP
Iris	Mean	0.9453	0.9562	0.9531	**0.9598**	**0.9602**	0.8258	0.8684	**0.9633**
	Std.	0.0267	0.0248	0.0237	**0.0247**	**0.0258**	0.0799	0.0709	**0.0244**
	Time	0.0006	0.0008	0.0017	0.0017	0.0024	0.0047	0.0059	0.4228
Seeds	Mean	0.9051	0.8965	0.9277	**0.9365**	0.9234	0.9143	0.9022	**0.9366**
	Std.	0.0331	0.0323	0.0243	**0.0231**	0.0280	0.0290	0.0326	**0.0262**
	Time	0.0008	0.0013	0.0036	0.0032	0.0036	0.0058	0.0071	0.6721
Margin	Mean	0.4323	0.7005	0.7269	0.7269	0.7326	0.6858	0.6852	**0.8329**
	Std.	0.0235	0.0225	0.0193	0.0189	0.0198	0.0225	0.0240	**0.0158**
	Time	0.0665	0.0323	0.5081	0.4475	0.8349	0.6627	0.7863	9.9122
Shape	Mean	0.4075	0.5161	0.5571	0.5497	0.5391	0.4060	0.4565	**0.6009**
	Std.	0.0211	0.0223	0.0211	0.0200	0.0215	0.0209	0.0217	**0.0271**
	Time	0.2031	0.0322	0.5079	0.4303	0.7965	1.9469	0.9429	8.2964
Texture	Mean	0.4766	0.6204	0.7363	0.7295	0.7273	0.7133	0.7006	**0.8192**
	Std.	0.0266	0.0220	0.0177	0.0181	0.0187	0.0231	0.0243	**0.0154**
	Time	0.1020	0.0324	0.5081	0.4285	0.8147	0.6546	0.7803	9.5951
Margin and Shape	Mean	0.6324	0.7912	0.9227	0.9158	0.9168	0.8922	0.8622	**0.9538**
	Std.	0.0229	0.0229	0.0109	0.0124	0.0114	0.0167	0.0178	**0.0090**
	Time	0.3080	0.0568	1.3923	1.3960	2.2613	1.1541	1.2942	7.5205
Margin and Texture	Mean	0.5780	0.6488	0.9546	0.9507	0.9488	0.9628	0.9215	**0.9716**
	Std.	0.0252	0.0246	0.0090	0.0097	0.0105	0.0091	0.0163	**0.0067**
	Time	0.1771	0.0541	1.3948	1.3944	2.2700	1.0128	1.235	5.0549
Shape and Texture	Mean	0.6102	0.7395	0.9034	0.8966	0.8929	0.8851	0.8546	**0.9375**
	Std.	0.0253	0.0213	0.0138	0.0145	0.0145	0.0150	0.0188	**0.0107**
	Time	0.3555	0.0543	1.4016	1.3991	2.2565	1.1367	1.2883	7.7685
Margin, Shape and Texture	Mean	0.6668	0.6619	**0.9806**	0.9773	0.9761	**0.9882**	0.9625	**0.9891**
	Std.	0.0247	0.0245	**0.0056**	0.0066	0.0070	**0.0048**	0.0101	**0.0050**
	Time	0.4713	0.0754	2.1261	2.1314	3.4388	1.4676	1.7575	5.2530

and Seeds data sets, the best average accuracy was obtained by MLP-BP, followed by K-NN algorithm (with $k = 4$ and 5).

In relation to 100 Plant Leaves data sets, when considering individual features, we observe that shape, by itself, offers a poor discriminatory power for automatic plant recognition methods. Since many plant species (for example, plant species from the same family) present similar leaf shape, all evaluated algorithms presented low average accuracies when considering only shape feature. The best individual feature (i.e., the feature with high discriminatory power) for 100 Plant Leaves is texture (except for NB, K-NN$_5$ and MLP-BP, where leaf margin was responsible for the best accuracies).

All classifiers presented better average accuracies when the features are combined two-by-two. The best combination of two features for most of the evaluated algorithms have been obtained when margin and texture features have been combined (except for DT algorithm, where the best average accuracy have been obtained by the combination of margin and shape features). For example, MLP-BP was able to achieve an average accuracy of 83.29% when considering leaf margin feature and 81.92% and considering leaf texture feature, but when leaf margin and leaf texture have been combined, the algorithm was able to achieve an average accuracy of 97.16%.

For most of the evaluated algorithms, the best results for 100 Plant Leaves data set have been achieved when all three features have been combined. The only exception is NB, where the best average accuracy have been obtained when only leaf margin and leaf texture have been combined. The obtained results for 100 Plant Leaves problem showed that it is quite important to find the best set of features when dealing with automatic plant recognition, since some features present better discriminatory power than others.

TABLE III
OVERALL EVALUATION: AVERAGE RANKS FOR THE FRIEDMAN/NEMENYI TEST.

Algorithm	Average Rank
DT	132.9561
NB	241.2094
K-NN$_3$	523.9111
K-NN$_4$	504.4428
K-NN$_5$	482.6061
SVM$_{linear}$	366.8772
SVM$_{rbf}$	265.4433
MLP-BP	**686.5539**

Table V presents the average ranks obtained by Friedman/Nemenyi hypothesis test. The Friedman/Nemenyi test shows that MLP-BP obtained the best average performances according to an overall evaluation, in comparison to all other selected algorithms. The second and the third best ranks have been obtained by K-NN$_3$ and K-NN$_4$, respectively. The worst overall performance have been achieved by Decision Tree classifier, followed by Naive Bayes classifier and Support Vector Machine. Figure 2 presents the obtained results for the Friedman/Nemenyi test, from the worst method (on the left side) to the best method (on the right side).

In relation to execution time, we can observe that MLP-BP

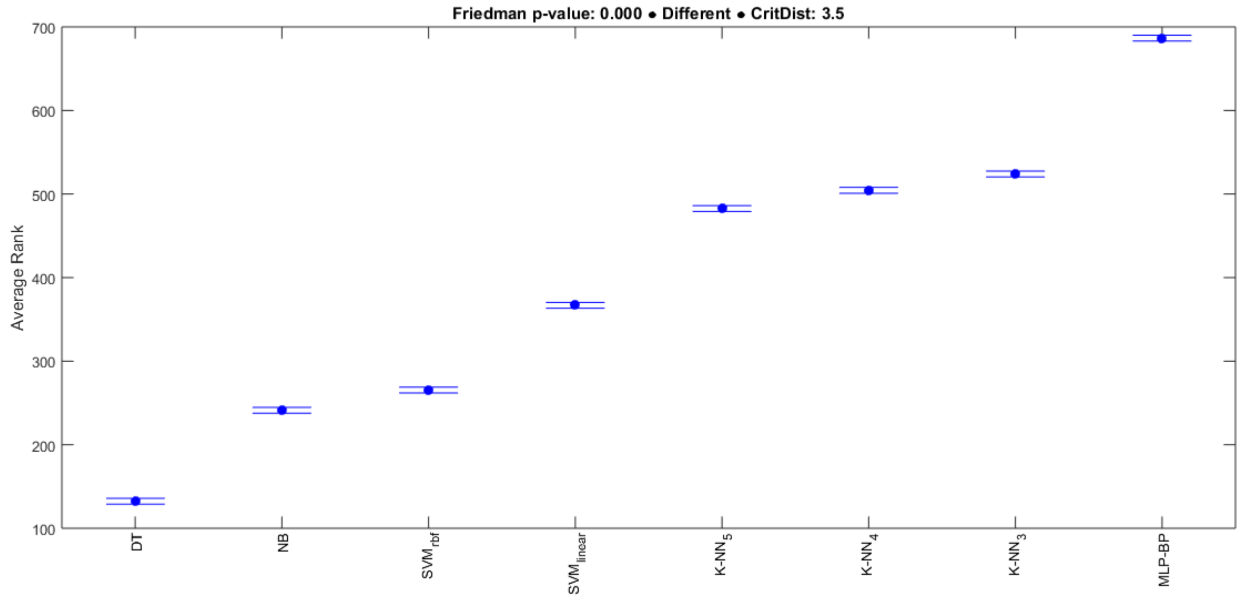

Fig. 2. Average ranks for the Friedman/Nemenyi test.

is slower to train than the selected comparison approaches, what may limit its application in some real-time applications. But if time is not too much critical, the high computational costs demanded by MLP-BP are compensated by the improvements in terms of system accuracy.

VI. Conclusion

In this work, we evaluate the performance of five well-established supervised learning methods when dealing with plant recognition problem: K-Nearest Neighbor, Decision Tree classifier, Naive Bayes classifier, Support Vector Machine and a Multi-Layer Perceptron trained with Backpropagation algorithm. Different configurations for K-Nearest Neighbors algorithm (different k values) and for Support Vector Machines (two different kernel functions - $Linear$ and RBF functions) are compared.

For comparison purposes, three real-world plant data sets obtained from UCI Machine Learning repository are employed: Iris, Wheat Seeds and 100 Plant Leaves. 100 Plant Seeds data set have been divided in seven data sets, so we could test each individual plant leaf feature and all possible combinations of the three features (leaf margin, leaf shape and leaf texture).

The evaluation criterion is based on an empirical analysis complemented by a hypothesis test of type Friedman/Nemenyi test in relation to the average test accuracy obtained by each classifiers for each of the nine adopted data sets.

The experimental results show that MLP-BP is able to achieve the better overall performances in terms of accuracy, outperforming all other algorithms. In an empirical analysis, MLP-BP is able to obtain performances at least as good as the results obtained by the others comparison classifiers in each data set, showing its potential as a classifier to tackle plant identification problems. The Friedman/Nemenyi hypothesis test also showed that MLP-BP obtained the best average rank in an overall evaluation, considering all nine evaluated data sets.

As future works, we intend to extend our study by the inclusion of other plant features extracted automatically using image processing techniques, and by adopting larger data sets. We also intend to evaluate the influence of the new set of features on the behavior of other classifiers from plant recognition literature.

References

[1] W. Rankothge, D. Dissanayake, U. Gunathilaka, S. Gunarathna, C. Mudalige, and R. Thilakumara, "Plant recognition system based on neural networks," in *Advances in Technology and Engineering (ICATE), 2013 International Conference on*. IEEE, 2013, pp. 1–4.

[2] J. S. Cope, D. Corney, J. Y. Clark, P. Remagnino, and P. Wilkin, "Plant species identification using digital morphometrics: A review," *Expert Systems with Applications*, vol. 39, no. 8, pp. 7562–7573, 2012.

[3] A. Sabu and K. Sreekumar, "Literature review of image features and classifiers used in leaf based plant recognition through image analysis approach," in *Inventive Communication and Computational Technologies (ICICCT), 2017 International Conference on*. IEEE, 2017, pp. 145–149.

[4] M. E. Rahmani, A. Amine, and M. R. Hamou, "Plant leaves classification," *ALLDATA 2015*, vol. 82, 2015.

[5] C. Mallah, J. Cope, and J. Orwell, "Plant leaf classification using probabilistic integration of shape, texture and margin features," *Signal Processing, Pattern Recognition and Applications*, vol. 5, no. 1, 2013.

[6] C. D. Mallah and J. Orwell, "Probabilistic classification from a k-nearest-neighbour classifier," *Computational Research*, vol. 1, no. 1, pp. 1–9, 2013.

[7] A. Sahay and M. Chen, "Leaf analysis for plant recognition," in *Software Engineering and Service Science (ICSESS), 2016 7th IEEE International Conference on*. IEEE, 2016, pp. 914–917.

[8] N. Kumar, P. N. Belhumeur, A. Biswas, D. W. Jacobs, W. J. Kress, I. C. Lopez, and J. V. Soares, "Leafsnap: A computer vision system for automatic plant species identification," in *Computer Vision–ECCV 2012*. Springer, 2012, pp. 502–516.

[9] T. Jin, X. Hou, P. Li, and F. Zhou, "A novel method of automatic plant species identification using sparse representation of leaf tooth features," *PloS one*, vol. 10, no. 10, p. e0139482, 2015.

[10] J.-C. Liu, C.-Y. Chiang, and S. Chen, "Image-based plant recognition by fusion of multimodal information," in *Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2016 10th International Conference on*. IEEE, 2016, pp. 5–11.

[11] A. Torralba and A. A. Efros, "Unbiased look at dataset bias," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 1521–1528.

[12] G. Agarwal, P. Belhumeur, S. Feiner, D. Jacobs, W. J. Kress, R. Ramamoorthi, N. A. Bourg, N. Dixit, H. Ling, D. Mahajan *et al.*, "First steps toward an electronic field guide for plants," *Taxon*, vol. 55, no. 3, pp. 597–610, 2006.

[13] Z. Wang, X. Sun, Y. Ma, H. Zhang, Y. Ma, W. Xie, and Y. Zhang, "Plant recognition based on intersecting cortical model," in *Neural Networks (IJCNN), 2014 International Joint Conference on*. IEEE, 2014, pp. 975–980.

[14] A. Sethulekshmi and K. Sreekumar, "Ayurvedic leaf recognition for plant classification," *Journal of Computer Science and Information Technologies*, vol. 5, no. 6, pp. 8061–8066, 2014.

[15] A. Asuncion and D. Newman, "Uci machine learning repository," 2007.

[16] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of human genetics*, vol. 7, no. 2, pp. 179–188, 1936.

[17] E. Anderson, "The irises of the gaspe peninsula," *Bulletin of American Iris Society*, vol. 59, pp. 2–5, 1935.

[18] M. Charytanowicz, J. Niewczas, P. Kulczycki, P. A. Kowalski, S. Łukasik, and S. Żak, "Complete gradient clustering algorithm for features analysis of x-ray images," in *Information technologies in biomedicine*. Springer, 2010, pp. 15–24.

[19] T. M. Mitchell *et al.*, "Machine learning. wcb," 1997.

[20] C. De Stefano, F. Fontanella, and A. S. Di Freca, "A novel naive bayes voting strategy for combining classifiers," in *Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference on*. IEEE, 2012, pp. 467–472.

[21] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," California Univ San Diego La Jolla Inst for Cognitive Science, Tech. Rep., 1985.

[22] S. S. Haykin, *Neural networks: a comprehensive foundation*. Tsinghua University Press, 2001.

[23] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[24] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, and G. Varoquaux, "API design for machine learning software: experiences from the scikit-learn project," in *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, 2013, pp. 108–122.

[25] M. Friedman, "The use of ranks to avoid the assumption of normality implicit in the analysis of variance," *Journal of the american statistical association*, vol. 32, no. 200, pp. 675–701, 1937.

[26] ——, "A comparison of alternative tests of significance for the problem of m rankings," *The Annals of Mathematical Statistics*, vol. 11, no. 1, pp. 86–92, 1940.

[27] P. Nemenyi, "Distribution-free multiple comparisons," in *Biometrics*, vol. 18, no. 2. INTERNATIONAL BIOMETRIC SOC 1441 I ST, NW, SUITE 700, WASHINGTON, DC 20005-2210, 1962, p. 263.

[28] J. Demšar, "Statistical comparisons of classifiers over multiple data sets," *The Journal of Machine Learning Research*, vol. 7, pp. 1–30, 2006.