Complex Identification of Plants from Leaves

Jair Cervantes[1]([⊠]), Farid Garcia Lamont[1],
Lisbeth Rodriguez Mazahua[2], Alfonso Zarco Hidalgo[1],
and José S. Ruiz Castilla[1]

[1] Posgrado e Investigación, UAEMEX (Autonomous University of Mexico
State), 56259 Texcoco, Mexico
{jcervantesc,fgarcial}@uaemex.mx,
jsergioruizc@gmail.com
[2] Division of Research and Postgraduate Studies, Instituto Tecnológico
de Orizaba, Av. Oriente 9, 852. Col. Emiliano Zapata, 94320 Orizaba, Mexico
lisbethr08@gmail.com

Abstract. The automatic identification of plant leaves is a very important current topic of research in vision systems. Several researchers have tried to solve the problem of identification from plant leaves proposing various techniques. The proposed techniques in the literature have obtained excellent results on data sets where the leaves have dissimilar features to each other. However, in cases where the leaves are very similar to each other, the classification accuracy falls significantly. In this paper, we proposed a system to deal with the performance problem of machine learning algorithms where the leaves are very similar. The results obtained show that combination of different features and features selection process can improve the classification accuracy.

Keywords: Plant identification · Vision system · Features selection

1 Introduction

Development of vision algorithms in agronomy has been guide by several researchers to solve many real problems [1–3]. Over the last years, many algorithms to identify plants from leaves have been developed. It is a current challenge that has several applications. The plant identification from leaves is not an easy job, because it involves the solution of different problems, such as: extract the leaf features and select the best features. Moreover, there are a lot of plants on the planet, many of them possess and share one or more properties such as: shape, size, texture, color, even when they belong to different plants.

In the current literature there are many techniques for identifying plants from the leaves. However, there are no systems to automatically identify plants where the leaves are closely related or are very similar to each other. This research has been motivated of this disadvantage. In this research, were used different features selection techniques to get the most discriminative features for each subset of plant leaves.

In the results, performances of several machine learning algorithms are compared using different features to identify plants where the leaves are very similar. The extracted

features allow to identify or classify the leaves. However, sometimes a high number of features introduces noise which affects the performance, i.e. the performance of the identification systems is strongly related to the features. The feature selection algorithms help to reduce the noise introduced in the classifier when the dimensionality in the data set is very large. The dimensionality of data set is reduced by eliminating features with low discriminative power. The general procedure is applied for each data set with the aim of selecting in each subset only the necessary characteristics.

The next section describes the state of the art. The third section describes the feature extraction methods used in this paper. The fourth section shows the proposed method. The fifth section shows the experiment and results. Lastly, the sixth section ends with the conclusions.

2 State of the Art

Plants identification has recently drawn attention from computer sciences. Identify a plant through leaves images is not a trivial job because it requires specialized knowledge. Current identification methods involve advanced algorithms to measure the morphological and texture features of the objects contained in the image. The best way to extract valid features is to get them from the image of the leaves. In the current literature it is shown that the external shape, chromaticity, venation and texture of the leaves give a lot of information to classify them.

Some other researches, have focused on the extraction of features from the leaf, using four important features for classification: Shape [4], texture [5], color [6], and leaf venation [7].

Leaf shape is one of the most important features of plant leaves, and the two basic approaches for these kind of analysis are the ones that are based on contour and region. One of the most employed approaches is to analyze the shape of leaves, extracting geometric characteristics such as size, elongation, ellipse, area, length, diameter, rectangularity, sphericity, eccentricity, etc. [8, 9]. Some authors have added to these basic geometric descriptors, Hu moments and Fourier moments improving performance of the classifiers [10]. The one based on region usually use moment descriptors, which includes Zernike geometric moments and Legendre moments. Some authors use basic descriptors, such as perimeter, area, circularity and elliptical, or invariant descriptors like Hu moments and Fourier descriptors for leaf contour recognition [11]. Methods based on contour, usually use methods based on the leaf curvature. Recently, some systems have been proposed to extract features describing edge variations of the leaf, using descriptors invariant to translation, rotation and size.

Texture of the leaves can be defined as the characteristics that the leaf has on its surface which is manifests as gray scale variations in the image. Texture features include local binary patterns, Gabor filters and gray level co-occurrence matrices, while the shape feature vector is modeled using Hu moment invariants and Fourier descriptors. Other researches had used a combination of geometric and textural, allowing them to use dried, wet or even misshapen leaves [12]. Some author combine both textural information and shape features to identify leaves [13, 14].

On the another hand, the first plant recognition studies used the chromaticity of the plant as an important descriptor to compare images. Very simple descriptors of chromaticity can obtain the average color in the segmented region of the leaf, average gradient in the edge or the similarity of color between two images that can be measured by comparing their color histograms. More complex color descriptors use moments of invariance which are commonly used to obtain geometrical characteristics but incorporating the information of the color variables of the leaves [15]. However, a recurring problem in the leaves of plants is that the chromaticity in the leaves is not static, it is variable with respect to time and commonly with respect to other factors. Other authors consider in addition to chromaticity and form, the texture of the leaf [16] or use combinations of descriptors to improve the classification performance [13, 14].

In other research studies the color is used as a comparison feature of images, since a simple color similarity between two images can be measured by comparing their color histograms [15]. However, a recurring problem is that the chromaticity in the leaves of plants is not static, this is variable with on the time and commonly on the other factors. Although classification approaches such as shape, texture and color are valid, it has not been documented the influence of each type of features in the performance of classification algorithms.

3   Proposed Method

In this Section, the steps of the proposed method are described in detail. After of features extraction the proposed methodology uses different techniques to select the best features of the data set. This step allows to reduce the dimensionality of the data set, reduce the training time and in some cases improve the performance of the system, this due to the elimination of features that introduce noise to the classifier (Fig. 1 and Table 5).
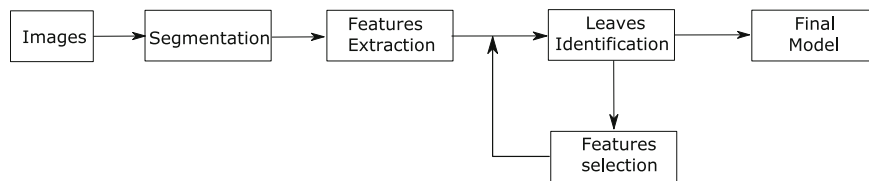
Fig. 1. Proposed methodology diagram

3.1   Segmentation Techniques

Firstly the images are preprocessed and segmented. The leaves images often are surrounded by greenery in the background. However, the images used in the experiments are images in a controlled environment (images with only leaf and white background). In all the experiments Otsu algorithm was used for segmentation. It is worth mentioning that we carry out experiments with different segmentation techniques, however the results are very similar in all cases, this is due to the images have a white

background and the environment is totally controlled. Therefore, it was not necessary to use more powerful techniques to perform the segmentation of the image.

In order to obtain a good segmentation even when there are changes in global brightness conditions, the region of the leave in each image was segmented using the following steps (1) Computation of high-contrast gray scale from optimal linear combination of RGB color components; (2) Estimate optimal border using cumulative moments of zero order and first order (Otsu method). (3) Morphological operations to fill possible gaps in the segmented image. By segmenting the image, the proposed system can use only the region of the leaf, determine its edges and calculate properties by extracting features.

3.2   Features Extractors

Feature extraction is a critical process in any pattern recognition system. The feature extraction has a big influence on the final identification. Feature extraction allows us to represent the image using a set of numerical and/or categorical values. In order to improve the performance the features obtained must be invariant to scaling, rotation and translation, enabling the classifier to recognize objects despite having different size, position and orientation. All these features play an important role in the algorithm performance and allow the classifier to discriminate between different classes in an appropriate manner. In our experiments geometric, chromatic and textural features were obtained.

Geometric Features. The geometric features are one of the most important visual properties used to classify an object. The geometric features provide information on the size and shape of the previously segmented region. Elementary geometric features provide intuitive information of the basic properties of the region to be recognized, such as area of the region, roundness of the leaf, length of the edge of the leaf, elongation defined by the length and width of the leaf, the coordinates x and y of gravity center, rectangularity, projection (on the components x, y), eccentricity, center of gravity (components x, y), Danielson factor, equivalent diameter, axis length (x, y), orientation, solidity, extencion, area convex, filled area, ellipse (variance, orientation, eccentricity, area, major axis, minor axis, ellipse center x, *and*). However, an efficient classification system should be able to recognize leaves regardless of their orientation, location and size, i.e. it must be invariant to scaling, rotation and position.

Moments are commonly used in image recognition, they can recognize these images regardless of their rotation, translation or inversion. Invariant moments were initially introduced by Hu. Other used features were ellipse descriptors, region convexity, Flusser moments ($F_1, \ldots, F_4$), R Moments ($R_1, \ldots, R_{10}$), Fourier descriptors (first 8 descriptors). 57 geometric features were extracted from each image. The geometric feature vector X_g obtained can be represented as:

$$X_g = [X_1, X_2, \ldots, X_{57}] \tag{1}$$

$$X_g = \left[X_{gb}, X_{Hu}, X_F, X_R, X_{DF}\right] \tag{2}$$

where X_{gb} represents the elemental geometric features $(X_{gb} = [X_1, \ldots, X_{28}])$, X_{Hu} represents the Hu invariant features $(X_{Hu} = [X_{29}, \ldots, X_{35}])$, X_F represents the Flusser invariant moments $(X_F = [X_{36}, \ldots, X_{39}])$, X_R represents the invariant moments to changes in illumination $(X_R = [X_{40}, \ldots, X_{49}])$, X_{DF} represents the first 8 Fourier descriptors $(X_{DF} = [X_{50}, \ldots, X_{57}])$.

Textural Features. Textural features provide information on the spatial arrangement of colors or intensities in the image. Extraction algorithms of textural features look for basic repetitive patterns with periodic or random structures in images. These structures are obtained by properties in the image such as roughness, roughness, granulation, fineness, softness, etc. Texture repeats a pattern along a surface, due to which the textures are invariant to displacements, this explains why the visual perception of a texture is independent of the position. In this paper, were used Haralick textural features and the Local Binary Patterns (LBP for its acronym in English - textit Local binary Patterns -) [17]. These features consider the distribution of intensity values in the region, by obtaining the mean and range of the following variables: mean, median, variance, smoothness, bias, Kurtosis, correlation, energy or entropy, contrast, homogeneity, and correlation. 14 textural descriptors of each image were obtained.

In total 219 textural descriptors were obtained from each image. 73 for each color channel. The textural features vector can be represented by:

$$X_t = [X_1, X_2, \ldots, X_{219}] \tag{3}$$

$$X_t = \left[X_{Rlbp}, X_{RH}, X_{Glbp}, X_{GH}, X_{Blbp}, X_{BH} \right] \tag{4}$$

where $X_{Rlbp}, X_{Glbp}, X_{Blbp}$ represents the LBP features obtained in the color channel R, G and B respectively, X_{RH}, X_{GH} and X_{BH} represents the Haralick textural features in the channels R, G and B respectively. A description in detail of LBP descriptors can be found in [17].

Chromatic Features. Chromatic features provide information of the color intensity of a segmented region. These characteristics can be calculated for each intensity channel, for example, red, green, blue, grayscale, hue (Hue), Saturation (Saturation) and intensity (Value), etc. The used features were: standard intensity features, they describe the mean, standard deviation of intensity, first and second derivative in the segmented region, Hu moments with intensity information, Gabor features based on 2D Gabor functions. In experiments 122 characteristics were obtained for each channel. Since the experiments were performed in RGB only 366 chromatic features were used. Chromatic features vector X_c can be defined by:

$$X_c = [X_1, X_2, \ldots, X_{117}] \tag{5}$$

$$X_c = [X_{Re}, X_{RHu}, X_{Ge}, X_{GHu}, X_{Be}, X_{BHu}] \tag{6}$$

where X_{Re}, X_{Ge}, X_{Be} represents the elemental color features in the channels R, G and B respectively, X_{RHu}, X_{GHu} y X_{BHu} represents the Hu invariant color moments in the channels R, G and B respectively.

3.3 Feature Selection

In order to eliminate some features that do not contribute to the classifier performance, in this research several algorithms were used to extract the best combinations of features.

In the proposed algorithm, each set of features per leaf, conforms a vector defined by the number of descriptors or features. The number of features of each data set defines the size of each binary string needed to implement the genetic algorithm. The relationship between each binary string with the feature set, is that 1 is taken as a used feature and 0 as the absence of that feature. The aptitude of each individual is taken from the accuracy obtained by classifying the set corresponding to that chain.

In the proposed method, the individual with better aptitude is taken and it passes intact to the next generation, it was used two-point crosses and mutation probability of 0.08. Figure 2 shows an example of chromosomal chains used and the classifier performance when features labeled 1 are used.
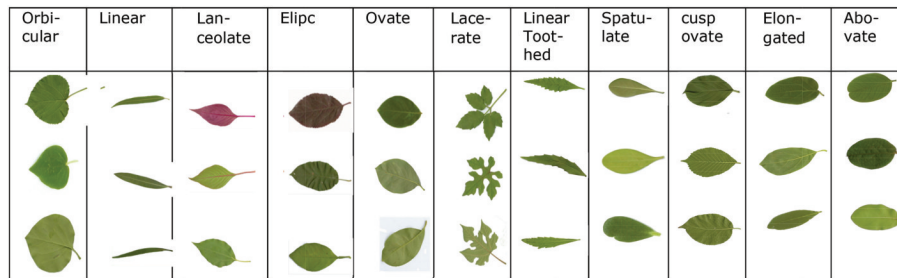
Orbi-cular	Linear	Lan-ceolate	Elipc	Ovate	Lace-rate	Linear Toot-hed	Spatu-late	cusp ovate	Elon-gated	Abo-vate

Fig. 2. Data sets

The dimensionality of data set is an important performance factor. Sometimes inappropriate attributes can affect the performance. Features selection helps to improve the performance of a classifier. This problem has been addressed by several authors, this problem is common in pattern recognition and it is commonly called course of dimensionality. An important factor when reducing characteristics, is to eliminate those that are not important to the classifier or find the combination of attributes that optimizes the performance of the classifier. Feature selection or dimensionality reduction is regularly raised as an optimization problem. In recent years, several algorithms have been used extensively to solve dimensionality problems.

Formally, given a n-dimensional data set, the features selection techniques task is to find a set of attributes in a k-dimensional space that maximizes an optimization criterion, where $k << n$. Obtained patterns are evaluated based on two conditions, dimensionality of the data set and spacing between classes or classification accuracy.

In the experiments were used 3 techniques to reduce the dimensionality of the data sets: Correlation-based feature selection, Information Gain Based Feature Selection and a genetic algorithm. Correlation-based feature selection and Information Gain Based Feature Selection techniques are used only once and then the data sets are trained with

reduced dimensionality. In the case of the genetic algorithm, this process is performed iteratively until the genetic finds the best features in the data set.

3.4   Classification Techniques

In the experiments the results were compared with some classification techniques, logistic regression, Bayesian classifier, the Backpropagation learning algorithm and support vector machine (SVM).

4   Experimental Results

In this section, parameters selection technique is shown, also data normalization and experimental results obtained with the proposed method.

4.1   Data Set

ICL data set, which is a collection of leaves of Hefei University was used in these experiments. The data set contains 16,849 leaf images from 220 species. The images of the leaves were segmented using Otsu's method and the above mentioned features were extracted.

In order to perform an analysis of the behavior of the classifiers on data sets where the leaves of the plants are very similar to each other, we obtained only 11 subsets. The 12 subsets of data are described in Table 1. The selection of sets was made manually according to the similarities and notable differences between classes. Subsets were formed from the leaves that had great similarity but with different class. From the initial 220 leaf species, 11 different subsets were formed with 169 species. The subset with fewer classes contains only 3 and the subset that has more associated classes contains 37.

Table 1. Subsets used in the experiments.

	Leaf shapes	Size	Number of clases	Labels
DS 1	Orbicular	1095	19	1–19
DS 2	Linear	348	8	20–27
DS 3	Lanceolate	962	14	28–41
DS 4	Eliptic	2481	27	42–68
DS 5	Ovate	782	12	69–80
DS 6	Lacerate	186	3	81–83
DS 7	Linear toothed	428	7	84–90
DS 8	Spatulate	832	9	91–99
DS 9	cusp ovate	1248	13	100–112
DS 10	Elongated	1632	20	113–132
DS 11	Abovate	4398	37	133–169
DS 12	Peach leaves	193	6	170–175

Figure 2 shows each subset created with some examples of leaves that were taken to form each group, in addition the reason of similarity that was taken into account to assign each leaf is described. The data set 12 contains only peach leaves. However, these are from 6 different species. As can be seen in the Fig. 3, the similarity between them is very high. The leaves are almost identical and for anyone who is not an expert, the differences could be unnoticed.

Fig. 3. Peach leaves of 6 different species

4.2 Data Normalization

The final feature vector T was stored in a m $\times$ 642 size array containing m images with 642 features. 57 geometric features, 219 textural and 366 chromatic features. All the extracted features were normalized with mean zero and standard deviation equal to 1.

4.3 Parameter Selection

In all used classifiers optimal parameters were obtained by cross-validation and grid search. Cross-validation is a model validation technique for assessing how the results of a statistical analysis will generalize to an independent data set. On the other hand, grid search exhaustively search all parameter combinations obtaining the best parameter combination.

4.4 Results

In the experiments, all data sets were normalized and cross-validation was used with k = 10. Table 2 shows the results obtained with geometric features, textural and chromaticity, as each individual features. DS_i defines the data set. For each classifier used, accuracies obtained with each individual set of characteristics are reported. The metric used to evaluate the performance of the classifier was precision and this is obtained from the classifier hits divided by the total of data set.

In the results, it is not possible to infer that the similarity between leaves significantly affect classifiers, performance of the classifiers that used very similar images of each other and dissimilar, are not contrasting. However, it is possible to appreciate that the textural features are little discriminative for most data sets, except for the set DS 6 and DS_12. One possible reason is that the size of the data set is very small.

Tables 2, 3 and 4 show the results obtained with different combination of feature techniques (chromatic, textural and geometric). In the results it is possible to see an

Table 2. Performance with different features.

	Chromatics				Textural				Geometric			
Subset	Bayes	BP	LR	SVM$_{RBF}$	Bayes	BP	LR	SVM$_{RBF}$	Bayes	BP	LR	SVM$_{RBF}$
DS 1	88.7	94.21	86.38	94.91	36.28	47.78	42.07	56.902	81.89	93.56	84.83	**95.408**
DS 2	81.57	84.50	82.17	88.28	38.87	68.24	70.94	73.053	85.58	90.19	87.64	**92.941**
DS 3	88.22	94.73	89.27	**95.03**	37.18	74.06	71.86	77.518	78.08	89.32	83.63	91.918
DS 4	86.25	94.56	89.61	**95.62**	36.72	67.32	65.91	75.881	74.70	90.74	85.73	93.409
DS 5	93.67	95.56	92.82	95.72	28.30	71.04	68.81	72.841	85.44	95.11	88.33	**97.289**
DS 6	99.37	**100**	98.62	**100**	75.53	91.71	76.53	93.308	93.08	99.37	98.11	99.742
DS 7	96.56	96.70	94.35	97.45	48.25	82.39	81.27	85.082	97.87	**98.75**	96.45	98.733
DS 8	95.85	97.83	96.18	98.56	56.40	84.23	83.85	87.875	95.26	98.31	96.84	**98.941**
DS 9	88.38	90.64	90.38	90.87	38.39	65.77	63.21	69.69	84.59	93.80	87.46	**94.410**
DS 10	86.31	95.85	92.48	**96.36**	39.54	78.89	75.59	82.871	80.21	93.10	86.84	94.182
DS 11	77.62	92.33	88.31	**93.91**	33.72	63.41	62.84	68.161	74.83	88.35	78.03	91.324
DS 12	54.36	62.58	52.48	**67.32**	25.49	36.23	31.43	38.76	51.21	57.28	54.83	64.69

Table 3. Performance with two combined features.

	Chromatic-textural				Chromatic-geometric				Textural-geometric			
Subset	Bayes	BP	LR	SVM$_G$	Bayes	BP	LR	SVM$_G$	Bayes	BP	LR	SVM$_G$
CH 1	87.35	93.42	88.26	94.26	92.37	97.24	93.73	97.59	86.47	93.80	88.15	95.27
CH 2	82.06	84.73	81.82	86.25	87.05	90.58	88.69	92.47	88.12	91.86	89.04	91.42
CH 3	90.41	95.86	90.12	96.42	92.51	96.94	94.12	97.93	85.15	93.04	83.61	94.28
CH 4	87.26	94.00	87.58	95.21	91.42	97.26	91.83	95.82	84.33	92.37	85.45	94.72
CH 5	92.24	95.95	91.34	95.30	96.33	95.11	97.43	96.01	90.21	94.60	91.46	95.08
CH 6	98.74	100	97.75	100	98.74	100	98.76	100	96.22	99.37	97.90	99.37
CH 7	96.86	95.48	96.93	97.28	95.85	96.56	95.26	97.71	96.38	96.53	95.74	98.76
CH 8	95.05	96.13	95.58	96.21	96.84	96.21	95.71	97.82	96.61	96.69	94.52	96.95
CH 9	87.77	91.48	87.04	92.10	94.41	95.07	94.91	95.23	87.55	92.13	86.35	90.17
CH 10	87.74	95.92	88.19	95.33	89.77	95.29	91.29	96.16	90.50	92.62	90.72	91.93
CH 11	80.53	92.56	82.22	93.16	83.81	93.88	85.83	94.92	84.03	92.65	87.18	92.81
CH 12	68.51	66.82	63.94	73.46	69.17	70.81	65.79	76.53	67.06	69.44	64.77	75.06

improvement in classification accuracy compared to the performances obtained with the original data set and compared with the other features selection techniques.

In all tests with different classification techniques, the results with the proposed technique improves the results obtained with other techniques. These results highlight the utility of the proposed method. The obtained results accuracy was improved in all chains using genetic, it is important to note that even though the number of features significantly decreased in all the results the combination of the three types of features is very necessary.

Table 6 shows the general results obtained with the three techniques of feature selection. In the Table ODS represents the results obtained with data set with all features, GA the results obtained with the genetic algorithm, CBFS the results obtained with the Correlation-based feature selection and IGBFS the results obtained with Information Gain Based Feature Selection algorithm.

Table 4. Performance with the three kind of features (chromatic-textural-geometric).

Subset	Bayes	BP	LR	SVM$_G$
DS 1	91.909	96.207	91.21	98.524
DS 2	87.739	90.038	88.41	94.941
DS 3	93.797	96.992	95.36	98.882
DS 4	90.940	96.253	94.71	97.183
DS 5	95.784	98.145	96.18	99.153
DS 6	98.742	100	99.48	100
DS 7	98.535	98.954	98.83	99.062
DS 8	96.354	98.177	97.74	99.828
DS 9	95.196	93.886	92.24	97.893
DS 10	90.027	98.292	93.54	98.531
DS 11	86.482	93.807	88.26	95.917
DS 12	71.52	72.49	65.38	79.45

Table 5. Performance with the proposed algorithm.

Classifier	DS1	DS2	DS3	DS4	DS5	DS6	DS7	DS8	DS9	DS10	DS11	DS12
Bayes	92.07	88.09	93.91	91.02	96.12	98.92	98.72	96.62	95.84	91.68	89.22	75.64
BP	96.36	91.27	97.25	96.38	98.45	100	98.95	98.22	94.68	98.6	95.87	76.02
LR	92.51	88.73	96.09	95.19	96.73	99.86	98.91	97.78	93.51	94.22	90.26	73.057
SVM$_G$	98.97	96.16	99.31	98.21	99.83	100	99.62	99.93	98.51	99.03	97.18	83.97

Table 6. Performance using the feature selection techniques

Classifier	ODS	GA	CBFS	IGBFS
Bayes	93.22	94.35	93.11	93.22
BP	96.43	97.54	96.34	96.43
SVM	98.17	98.86	98.17	98.17
LR	94.17	94.63	94.09	94.12

5 Conclusions

In this paper, a feature selection algorithm is proposed to improve the performance of classifiers for identifying plants that are very similar. The proposed method helps to improve the performance of the classifiers removing attributes that introduce noise. The experiments obtained show that the proposed method generates notable results by eliminating attributes that do not provide information. The main advantage of the proposed method is its ease of implementation and ease of use on small and medium size data sets. Features reduction is important to improve the response time it takes for the system to recognize a new leaf. Several issues could be considered as future works. First, the algorithm can be used for plants identification of species from leaves, however, there are leaves of different species that have a very high degree of similarity,

it is necessary to add another type of characteristics such as leaf venation that were not included in this research. Second, in the results of this research only images of the front of the leaves were used. However, it is possible that the back of the leaves provides more important information than the front of the leaf. A study of this would be very important for this research.

References

1. Huang, Y., Lan, Y., Hoffmann, W.C.: Use of airborne multi-spectral imagery for area wide pest management. Agric. Eng. Int. CIGR Ejournal Manuscr. IT **07**(010), 1–14 (2008)
2. Singh, V., Misra, A.K.: Detection of plant leaf diseases using image segmentation and soft computing techniques. Inf. Process. Agric. **4**(1), 41–49 (2017). ISSN:2214-3173
3. Ferentinos, K.P.: Deep learning models for plant disease detection and diagnosis. Comput. Electron. Agric. **145**, 311–318 (2018). ISSN:01681699
4. Du, J.X., Wang, X.F., Zhang, G.J.: Leaf shape based plant species recognition. Appl. Math. Comput. **185**(2), 883–893 (2007)
5. Sampallo, G.: Reconocimiento de tipos de hojas. Inteligencia Artificial. Rev. Iberoam. Intel. Artif. **7**(21), 55–62 (2003)
6. Cerutti, G., Tougne, L., Mille, J., Vacavant, A., Coquin, D.: Understanding leaves in natural images - a model-based approach for tree species identification. Comput. Vis. Image Underst. **117**(10), 1482–1501 (2013)
7. Larese, M.G., Namías, R., Craviotto, R.M., Arango, M.R., Gallo, C., Granitto, P.M.: Automatic classification of legumes using leaf vein image features. Pattern Recognit. **47**(1), 158–168 (2014)
8. Chaki, J., Parekh, R.: Designing an automated system for plant leaf recognition. Int. J. Adv. Eng. Technol. **2**(1), 149–158 (2012)
9. Park, J.-S., Kim, T.Y.: Shape-Based Image Retrieval Using Invariant Features. In: Aizawa, K., Nakamura, Y., Satoh, S. (eds.) PCM 2004. LNCS, vol. 3332, pp. 146–153. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-30542-2_19
10. Kumar N., Belhumeur P.N., Biswas A.: Leafsnap: a computer vision system for automatic plant species identification. In: Proceedings of the ECCV 2012, pp. 502–516 (2012)
11. Novotny, P., Suk, T.: Leaf recognition of woody species in Central Europe. Biosyst. Eng. **115**(4), 444–452 (2013)
12. Husin, Z., Shakaff, A.Y.M., Aziz, A.H.A., Farook, R.S.M., Jaafar, M.N., Hashim, U., Harun, A.: Embedded portable device for herb leaves recognition using image processing techniques and neural network algorithm. Comput. Electron. Agric. **89**, 18–29 (2012)
13. Liu, N., Kan, J.-m.: Improved deep belief networks and multi-feature fusion for leaf identification. Neurocomputing **216**, 460–467 (2016). ISSN:0925-2312
14. VijayaLakshmi, B., Mohan, V.: Kernel-based PSO and FRVM: an automatic plant leaf type detection using texture, shape, and color features. Comput. Electron. Agric. **125**, 99–112 (2016). ISSN:0168-1699
15. Tico, M., Haverinen, T., Kuosmanen, P.: A method of color histogram creation for image retrieval. In: Proceedings of the Nordic Signal Processing Symposium (NORSIG-2000), Kolmarden, Sweden, pp. 157–160 (2000)

16. Cope, J., Corney, D., Clark, J., Remagnino, P., Wilkin, P.: Plant species identification using digital morphometrics: a review. Expert Syst. Appl. **39**(8), 7562–7573 (2012)
17. He, D.C., Wang, L.: Texture unit, texture spectrum, and texture analysis. IEEE Trans. Geosci. Remote Sens. **28**, 509–512 (1990)