A Fast Leaf Recognition Algorithm based on SVM Classifier and High Dimensional Feature Vector

Cecilia Di Ruberto and Lorenzo Putzu

Department of Mathematics and Computer Science, University of Cagliari, via Ospedale 72, 09124 Cagliari, Italy

Keywords: Image Analysis, Feature Extraction, Leaf Recognition, Plant Classification, Support Vector Machine.

Abstract: Plants are fundamental for human beings, so it's very important to catalog and preserve all the plants species. Identifying an unknown plant species is not a simple task. Automatic image processing techniques based on leaves recognition can help to find the best features useful for plant representation and classification. Many methods present in literature use only a small and complex set of features, often extracted from the binary images or the boundary of the leaf. In this work we propose a leaf recognition method which uses a new features set that incorporates shape, color and texture features. A total of 138 features are extracted and used for training a SVM model. The method has been tested on Flavia dataset (Wu et al., 2007), showing excellent performance both in terms of accuracy that often reaches 100%, and in terms of speed, less than a second to process and extract features from an image.

1 INTRODUCTION

Hundreds of thousands of plants exist on earth, which share a very close relationship to human beings and they play a very important role for human life in many areas, such as food, medical science, industry and environment. For this reason it is very important to be able to catalog and preserve all species of existing plants, many of which are endangered just because of human activity and pollution. Traditionally the cataloging activities of an unknown plant species requires the collection of a sample of the plant, which must be done dry before being analyzed and processed by specialists for comparison with other samples present in the herbaria. A faster way to catalog the plants is based on leaves analysis which, unlike flowers and fruits, are present for most of the year and provide comprehensive morphological characteristics for classification. Morphological analysis of the leaf, although more rapid, can not be a simple task especially for a non-expert user, as it needs for accurate observations of the leaf itself and the comparison with other samples of the same species, requiring prior knowledge or the ability to navigate through complex databases. For this reasons it's important to develop an automatic method in order to extract the best leaves' features and able to catalog a leaf from an unknown plant species. Plant species classification based on plant leaves images has become a very

important area of active research and many methods for leaves recognition are present in literature. Typically these methods use features extracted from the binary images or the boundary of the leaf. Some researchers used part of those features only, while others tried different and more complex approaches (Gao et al., 2010a). In addition only few works make use of color and texture information, such as first and second order statistics (Kadir et al., 2011). However, in almost all works in literature has been used a fairly narrow set of feature, or in some cases techniques like the PCA were used to reduce this set of features (Wu et al., 2006). In this work we propose a plant classification system based on leaf recognition which makes use of a new and larger features set, that incorporates shape, color and texture features, extractable from the leaves images easily and quickly, in order to ensure not only a remarkable accuracy but also speed in cataloging an unknown leaf species.

The rest of this paper is organized as follows. Section 2 describes how the image is pre-processed from the input image. Section 3 illustrates the features we propose. Section 4 describes the classification step. The experimental results are presented in Section 5 and conclusions in Section 6. Each phase of the method is analyzed in detail and compared with other approaches present in literature. The whole process can be schematized as showed in Figure 1.
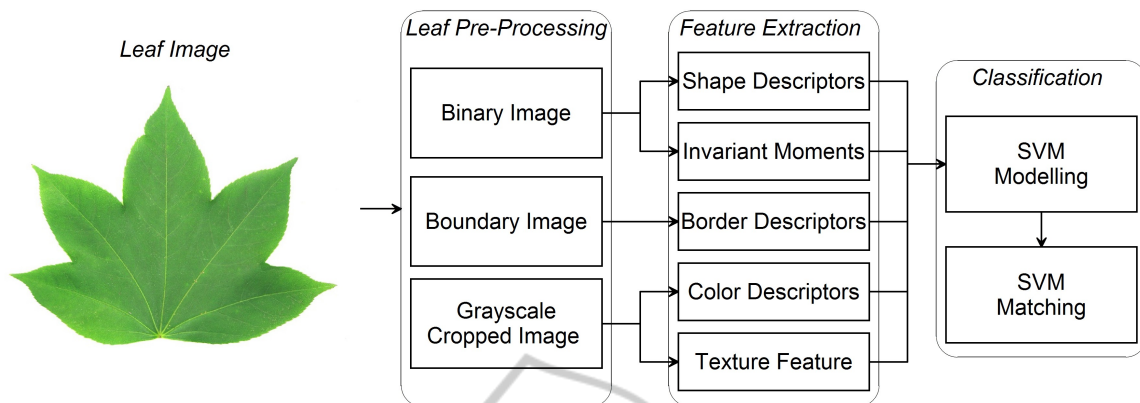
Leaf Image

Figure 1: Proposed method diagram.

2 IMAGE PRE-PROCESSING

The identification of leaves from Flavia dataset images is not a difficult task, in fact all the images present a white and almost uniform background. First of all the image is scaled in order to make the subsequent stage faster and then is converted from RGB to gray scale. To make faster also the segmentation step we use a predetermined threshold, like that used in (Wu et al., 2007), but with a value of 0.92, which from the experiments carried out we observed that is optimal for all the images present in this dataset. Then a complement image is calculated in order to obtain a white leaf on a dark background. However, this threshold operation does not produce a binary image cleaned in all its parts, due to the presence of dark particles in the background or light particles inside the leaf, such as the vein. Therefore an image cleaning is performed through an area opening operation, that removes all the objects in the image except one, that presents the maximum size, which is the leaf. After that an hole filling operation is performed in order to fill all the holes inside the leaf created by the threshold operation. At this point it is possible to use the binary image just obtained to crop the background from the original image in gray scale, useful for the color and texture feature extraction. The last operation required before the feature extraction step is the leaf boundary extraction, which is obtained by a simple removal operation, that removes from the binary images all those pixel belonging to the interior of the object, which have as neighbors only white pixels. At this point we have all the images (shown in Fig. 2) necessary for the extraction of our set of features. Unlike other works in literature, our method does not present a vein extraction step, as this approach seems very linked to the quality of the segmentation and therefore susceptible to many errors, but it also seems very expensive

in terms of time, given that this operation consists of repeated operations of opening followed by as many threshold operations, according to the approach proposed in (Wu et al., 2006), or by repeated operations of edge detection exploiting different threshold values, according to the approach proposed in (Du et al., 2006).

3 FEATURE EXTRACTION

Speaking about feature extraction in this context means to transform the images into data, then extract information reflecting the visual patterns to which botanists refer to, but at the same time it is necessary to extract the descriptors that are most relevant to the subsequent classification process. As said earlier, methods of feature extraction from leaf images present in literature, typically, use features extracted from the binary images or the boundary of the leaf, considering leaf external shape features the most important information for classification. For example, Wang (Wang et al., 2003) represented the shape of leaf with a centroid-contour distances curve. Im (Im et al., 1998) and Du (Du et al., 2006) used polygonal approximation to recognize plant species of Acer family. Lee (Lee and Chen, 2006) and Arribas (Arribas et al., 2011) instead proposed a region-based classification method for leaves. Other methods combine the previous shape features with other features that involve the analysis of many parts of the leaves, such as edge, vein and skeleton considering also leaf internal shape features. For example, Wu (Wu et al., 2006) combined geometric features and vein features. Cheng (Cheng et al., 2007) used a method that combines leaf shape, leaf apex, leaf base, leaf margin, leaf phyllotaxy, leaf vein and leaf size. Pauwels (Pauwels
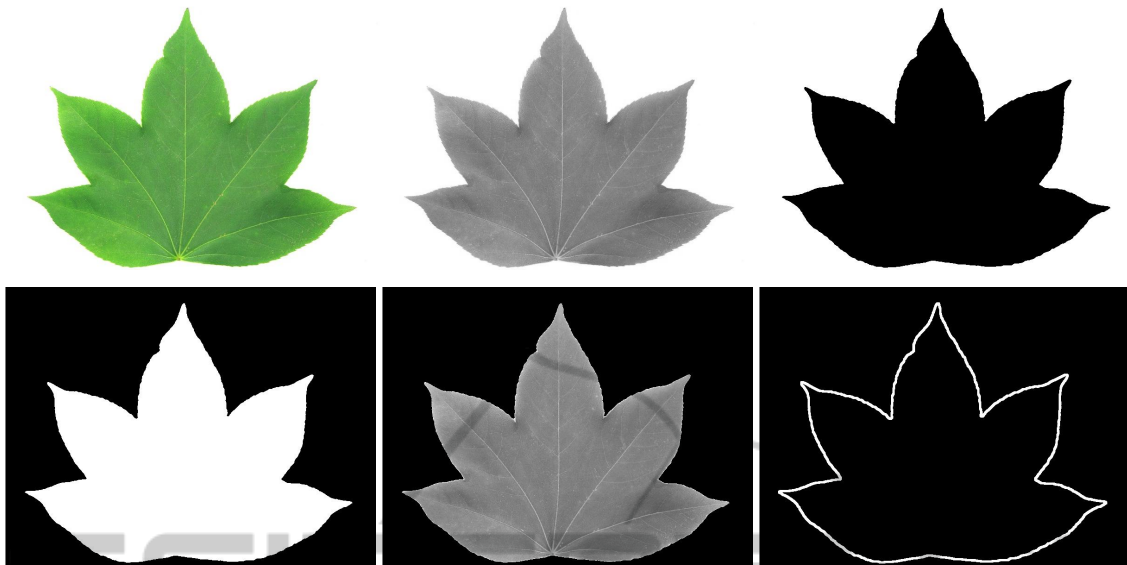
Figure 2: Leaf image pre-processing: RGB scaled image, gray level image, binary image, complement binary image after cleaning, gray scale cropped image, boundary image.

et al., 2009) used several shape features, including indentation depth which is the distance from leaf contour to the convex hull. Gao (Gao et al., 2010a), (Gao et al., 2010b) used shape features in combination with contour angles. Some researchers used part of those features only, while others tried different and more complex approaches, for example the invariant moments proposed by Hu (Hu, 1962) have been used by Zulkifli (Zulkifli et al., 2011). Zhang (Zhang and Lei, 2011) proposed the use of Polar Fourier Transform. Chaki (Chaki and Parekh, 2011) analyzed plant leaf images using shape modeling techniques based on the invariant moments and the centroid-radii model. Wu (Wu et al., 2007) identified plants species using shape features such as aspect ratio, leaf dent, leaf vein and invariant moment. In addition only few works make use of color and texture information, such as first and second order statistics. Man (Man et al., 2008) involved color moments based on mean, standard deviation and skewness. Ehsanirad (Ehsanirad, 2010) used textural features extracted from the gray-level co-occurrence matrices. Zhang (Zhang and Zhang, 2008) retrieved the leaf database using features of color, texture and shape. Machado (Machado et al., 2013) investigated the use of partial differential equations and fractal dimension theory to describe the texture present in leaves. In a similar way to that proposed by Kadir (Kadir et al., 2011) who realized a leaf classification using shape, color and texture feature, from the images generated previously we have extracted different types of descriptors: external shape features, regional shape features, invariants moments, color features and texture features. Differently from

Kadir (Kadir et al., 2011) our method involves the use of a large number of features but easy and quick to extract, in order to guarantee a high level of performance. So, starting from the boundary image we have extracted 36 descriptors of the centroid radii. This was done by calculating the centroid and after the distance of each boundary point from the centroid itself. Maximum and minimum distance are then used to rotate the contour, in such a way that the first centroid radii descriptor represents for all the leaves the maximum distance. At this point the distances are sampled to 36 values, in order to preserve a good approximation of the originals radii, preserving a radius at every 10 degrees interval. Finally, the sampled radii are normalized using again the maximum distance value, in order to obtain values between 0 and 1.

Then, starting from the binary image containing only the leaf, we have extracted 7 Invariant Moments proposed by Hu (Hu, 1962). Again from the binary image we have extracted some classical shape descriptors such as area, perimeter, convex area, convex perimeter, minimum bounding box width and height (showed in Fig. 3). These classical measures with the maximum and minimum distance (often called excircle radius and incircle radius, showed in Fig. 3), calculated previously, are combined in order to calculate 8 regional shape descriptors that are elongation, rectangularity, solidity, eccentricity, compactness, roundness, convexity and sphericity. We have chosen these shape descriptors for their ability to discriminate various shapes and therefore different leaves species, preserving, due to their calculation (for formulas refer to Gonzalez (Gonzalez et al., 2004)), values always be-
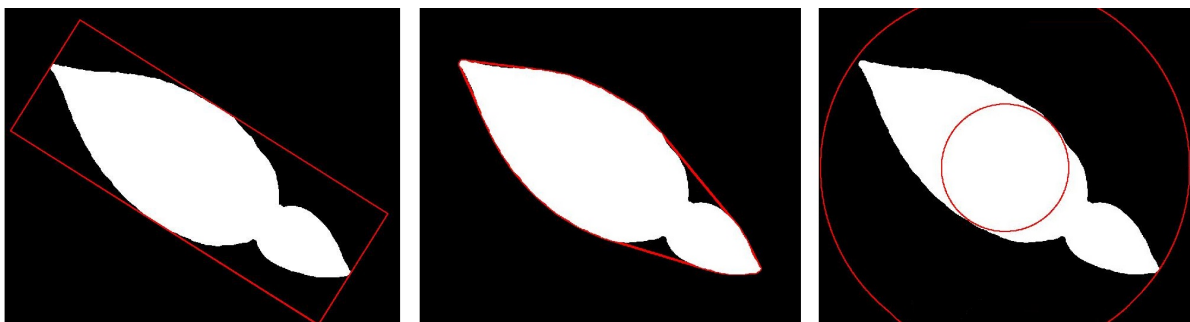
Figure 3: Leaf minimum bounding box, convex hull, excircle and incircle.

tween 0 and 1. Also, these descriptors together with invariant moments are invariant to the operations of rotation, translation and scaling.

The main disadvantage of these and others shape features is that they are very susceptible to errors in segmentation. For this reason we have used these descriptors together with regional descriptors less susceptible to errors. Among these there are the color descriptors which delineate the gray level color distribution of the images. So, starting from the gray scale containing the cropped leaf, we have extracted 7 color descriptors that are mean, standard deviation, smoothness, skewness, kurtosis, uniformity and entropy. Often, however, the descriptors based only on histogram have some drawbacks as they do not give information on the mutual position of the pixels. Some objects, in particular leaves, have in fact a repeating pattern as the primary visual characteristic and so it is necessary to consider not only the intensity distribution but also the positions of the pixels having a similar gray level, as showed by Haralick in (Haralick et al., 1973). Then we have also considered the descriptors applied to the gray level co-occurrence matrix (GLCM) calculated starting from the images containing the cropped leaf in gray scale. These descriptors are autocorrelation, contrast, correlation, cluster prominence, cluster shade, dissimilarity, energy, entropy, homogeneity, maximum probability, variance, sum average, sum variance, sum entropy, difference variance, difference entropy, information measure of correlation1, information measure of correlation2, inverse difference normalized and inverse difference moment normalized. These characteristics have been calculated for angles of 0, 45, 90 and 135 degrees and distance 1, so we have extracted 80 texture features. The total number of extracted features is then 138: 36 boundary descriptors, 7 Hu invariant moments, 8 shape descriptors, 7 color descriptors and 80 texture descriptors.

4 CLASSIFICATION

Various types of classifiers have been used for leaves recognition. In particular Neural Networks have been attracted researchers in area pattern recognition because of their power to learn from training dataset. For example, Chaki (Chaki and Parekh, 2011) presented a scheme for automated detection using several Neural Network classifiers, while Zulkifli (Zulkifli et al., 2011) used a General Regression Neural Network. Probabilistic NN is another neural network that has been used by several authors, such as Wu (Wu et al., 2006) and Huang (Lin and Peng, 2008). Du (Du et al., 2007) instead defined a classification method called the move median centers (MMC) hypersphere classifier. In some cases, before the classification step, various authors also used classical dimensionality reduction approaches in order to find a low-dimensional subspace in a high-dimensional input space by linear or no linear transformation. Among them, Principal Component Analysis PCA used by Wu (Wu et al., 2006) and Kernel Principle Component Analysis KPCA used by Valliammal (Valliammal and Geethalakshmi, 2011). On the contrary, in our work, the idea is to use the entire feature space for the training of a SVM model, both because a high number of feature guarantees the creation of a model with Support Vector able to achieve an optimal separation among the various classes, but especially because the process of reducing the number of features using linear transformation techniques, such as PCA, or by feature selection techniques are time consuming. The SVM classifier is trained following the one vs rest approach, which is preferred to the one vs one approach, since the latter is very expensive in terms of time and resources, necessary both for the creation of an higher number of models and for the classification of new instances that requires the comparison with all the models previously created. Since the performance of the SVM, while
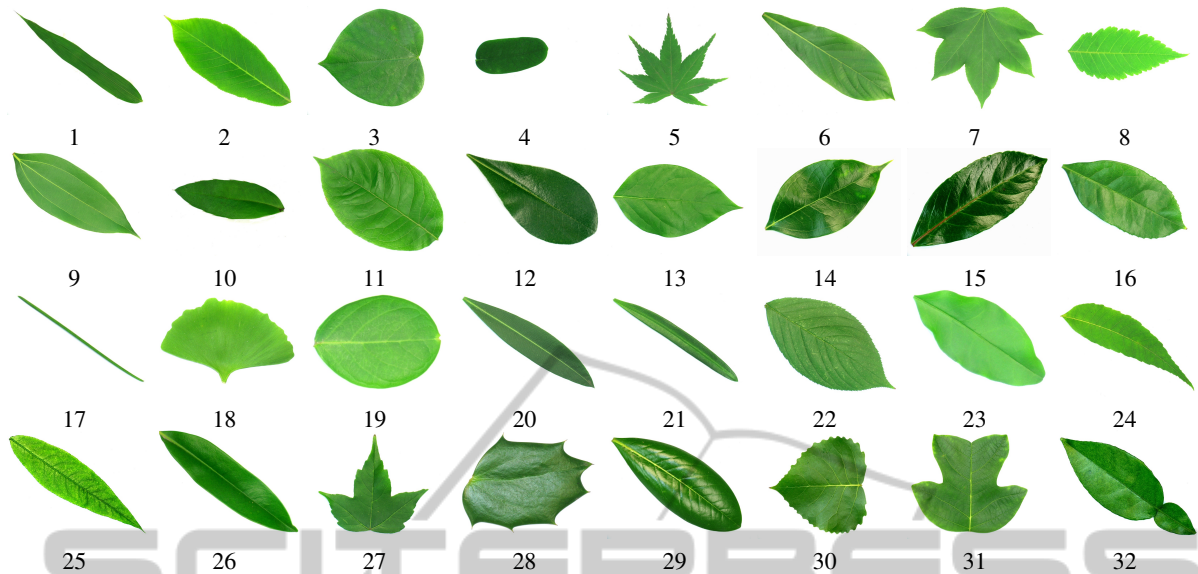
Figure 4: All leaves species in the Flavia dataset.

adopted with different kernel functions and different parameters, are quite different, the classifier has been modeled with different kernels and have been chosen only those with better performances. For each kernel functions the parameters have been tuned through optimization techniques in order to find the maximum accuracy value.

5 EXPERIMENTAL RESULTS

The proposed method has been finally tested on Flavia dataset, shared by Wu et al. (Wu et al., 2007) for academic research purpose, that contains a total of 1907 leaf images belonging to 32 plant species, showed in Fig. 4. In Table 1 are then listed the names of the plants and the number of samples for each species, which on average is approximately 60. Given the considerable size of the dataset we have decided to perform the validation of our method using a stratified sampling, that allows us to maintain in the training and test sets the same proportion between the classes present in the original dataset. The performances of the classification models were then evaluated by calculating the accuracy, which gives us a good indication of the performance since it considers each class of equal importance. The accuracy value has also been calculated individually for each plant species, as showed in Table 1, in which we have also listed the most common recognition mistakes to indicate with which other leaves classes are frequently confused. In order to assess the robustness of our algorithm, we have tested the classifier with different percent-

ages of training and test set. The first test was performed using the largest possible number of samples for the training, therefore using only one sample for each species for the test, so the training set size is 1875 and the test set size is 32. The following tests have been made using a decreasing number of samples for the training and an increasing number for the test, up to a percentage of 50%. So, in the last test we have used almost the same number of samples for the training and test sets, which in this case have a size of 953 and 954 respectively. The experimental results in Table 2 show the excellent performance of the proposed system in all tests performed. Despite each kernel function used in the creation of the models leads often to a precision of 100%, the highest performances are related undoubtedly to the Linear kernel and the RBF kernel that uses a Gaussian radial basis function. In particular, it can be noted that, also varying significantly the percentage of samples used for training and for the test, in both cases the system is able to effectively classify the plants species, while maintaining a consistent high accuracy value and a very low standard deviation value, which goes to stabilize even more with the use of a greater number of samples for testing. The efficacy of the feature set used can be seen from Table 3, which shows the performance of the system using different feature subsets among those listed previously. For this test we have used a sample percentage for training set and test set, respectively, equal to 80% and 20%.

From the results shown it is possible to observe how each feature subset added leads to a greater level of accuracy and contributes greatly to the stability of

Table 1: Details about leaf name and number, mean accuracy for each species and common recognition mistakes.

Common name	Label	Species Samples	Accuracy	Common Mistake
Pubescent Bamboo	1	59	98.2%	20
Chinese Horse Chestnut	2	63	98.4%	6
Chinese Redbud	3	72	98.2%	31
True Indigo	4	73	100%	no.
Japanese Maple	5	56	100%	no.
Nanmu	6	62	98%	2
Castor Aralia	7	52	100%	no.
Goldenrain Tree	8	59	100%	no.
Chinese Cinnamon	9	55	100%	no.
Anhui Barberry	10	65	98.2%	4
Big-fruited Holly	11	50	100%	no.
Japanese Cheesewood	12	63	98%	24
Wintersweet	13	52	98%	11
Camphortree	14	65	100%	no.
Japanese Viburnum	15	60	100%	no.
Sweet Osmanthus	16	56	100%	no.
Deodar	17	77	100%	no.
Ginkgo Maidenhair Tree	18	62	100%	no.
Crape Myrtle	19	61	100%	no.
Oleander	20	66	98.3%	21
Yew Plum Pine	21	60	98%	20
Japanese Flowering Cherry	22	55	100%	no.
Glossy Privet	23	55	100%	no.
Chinese Toon	24	65	98.2%	2
Peach	25	54	98%	10
Ford Woodlotus	26	52	94%	1, 20
Trident Maple	27	53	100%	no.
Beale's Barberry	28	55	100%	no.
Southern Magnolia	29	57	100%	no.
Canadian Poplar	30	64	98%	3
Chinese Tulip Tree	31	53	100%	no.
Tangerine	32	56	100%	no.

Table 2: Experimental results with different sampling percent: for each test accuracy (acc) and standard deviation (SD) are reported.

	Linear SVM		Quadratic SVM		RBF SVM	
Sampling perc.	acc	S D	acc	S D	acc	S D
1 s	0,994	0,013	0,977	0,028	0,994	0,013
10%	0,99	0,006	0,976	0,013	0,989	0,007
20%	0,99	0,004	0,97	0,007	0,988	0,006
30%	0,984	0,004	0,963	0,008	0,981	0,006
40%	0,983	0,004	0,957	0,006	0,982	0,004
50%	0,971	0,004	0,949	0,009	0,974	0,004

the system by reducing the standard deviation value. Furthermore, one can observe the remarkable contribution of color and texture features, which in most of the works reported in the literature have not been taken into account, but which obviously are very useful in the discrimination of different leaves species, in particular the feature of texture with regard to the veins of the leaves. The Table 4 compares the results

obtained through our method with the results obtained by various authors who have proposed a method of leaf recognition tested on the same dataset Flavia. Wu (Wu et al., 2007) used 12 shape and vein features, reduced to 5 principal components through PCA in order to train a PNN classifier. Also Singh (Singh et al., 2010) used shape and vein features but for the training of an SVM model arranged in a binary tree struc-

Table 3: Experimental results for different feature sets: for each test accuracy (acc) and standard deviation (SD) are reported.

	Linear SVM		RBF SVM	
Feature Type	acc	S D	acc	S D
Shape	0,723	0,014	0,869	0,017
Shape+Inv.mom.	0,738	0,018	0,874	0,009
Shape+Inv.mom.+Cent.radii	0,784	0,017	0,907	0,012
Shape+Inv.mom.+Cent.radii+Color	0,918	0,015	0,97	0,009
Shape+Inv.mom.+Cent.radii+Color+Texture	0,99	0,004	0,988	0,006

Table 4: Comparison between our method and other methods present in literature.

Method	Accuracy
Wu (Wu et al., 2007)	93%
Singh (Singh et al., 2010)	96%
Kadir (Kadir et al., 2011)	94,7%
Kulkarni (Kulkarni et al., 2013)	93,8%
Lee (Lee and Hong, 2013)	97,2%
Our method	99%

ture. Kadir (Kadir et al., 2011) extracted shape, vein, color and texture feature and Polar Fourier Transform to train a PNN classifier. Kulkarni (Kulkarni et al., 2013) tested Zernike moments combined with shape, color and texture feature. Finally, Lee (Lee and Hong, 2013) used shape, vein and contour feature and FFT to recognize leaf species with an iterative score priority system. From the results it can be seen how the proposed method leads to greater accuracy than that achieved by other methods tested on Flavia dataset. The tests carried out have also shown the excellent performance in terms of speed of the proposed algorithm. Despite the speed of a system of leaf recognition depends on many factors, such as the size and resolution of the image, the size of the leaf and the complexity of texture and last but not least the configuration of the computer (our computers was configured respectively with a processor Intel(R) Core(TM) i7 CPU, 960@3.20 GHz, RAM 4.00 GB size and with a processor AMD Phenom(TM) II x6 1090T, RAM 4.00 GB size), in any case, we are able to extract the feature vector from a leaf sample in less than a second. Even, the SVM classifier with the one vs. rest approach has proven itself sufficiently fast, being able to create the model in the worst case, that is using the RBF kernel, in about 12 seconds, while the predictor step is performed instantaneously. So, we are able to classify an unknown leaf species in less then a second.

6 CONCLUSIONS

In this work we have proposed an innovative method for the completely automatic identification and classi-

fication of plant species based on leaf recognition, in order to provide an automated procedure as support for plant cataloging and preserving. Starting from an original leaf image we can process it entirely, in a fully automatic way without need for manual intervention by the user. The proposed method makes use of a new and larger features set, that incorporates shape, color and texture features, extractable from the leaves images easily and quickly. This set of features is used entirely to train a SVM classifier with one vs rest approach. The experimental results demonstrate that the proposed leaf recognition system has excellent performance, both in terms of accuracy and in terms of speed, less than a second to catalog an unknown leaf species. The accuracy often reaches 100% and on average is around 99%, so it is better than the methods reported in literature. Further developments could affect the feature extraction phase, in order to extract vein features easy to detect and more discriminating than the features shown in the works reported in literature. Thereafter, it will be important to test the proposed method on other available datasets that present greater variations in the type of images acquired, both with regard to the dimensions, the resolution and the quality of the images themselves. In this way we would be able to assess the quality of the proposed method and make the changes necessary to maintain a high level of performance and speed. Finally we will be able to focus in the development of a real application for PC, which allows to catalog images of leaves acquired with a digital camera, and a mobile application that allows to catalog the plants directly in their natural habitat. This will require the addition of new phases in the algorithm, in particular for the segmentation of the images with the presence of both uniform and non-uniform backgrounds.

ACKNOWLEDGEMENTS

REFERENCES

Arribas, J. I., Snchez-Ferrero, G. V., Ruiz-Ruiz, G., and Gmez-Gil, J. (2011). Leaf classification in sunflower crops by computer vision and neural networks. *Computers and Electronics in Agriculture*, 78(1):9 – 18.

Chaki, J. and Parekh, R. (2011). Plant leaf recognition using shape based features and neural network classifiers. *International Journal of Advanced Computer Science and Applications*, 2(10):41 – 47.

Cheng, S.-C., Jhou, J.-J., and Liou, B.-H. (2007). Pda plant search system based on the characteristics of leaves using fuzzy function. In *New Trends in Applied Artificial Intelligence*, volume 4570 of *Lecture Notes in Computer Science*, pages 834–844. Springer.

Du, J.-X., Huang, D.-S., Wang, X.-F., and Gu, X. (2006). Computer-aided plant species identification based on leaf shape matching technique. *Transactions of the Institute of Measurement and Control*, 28(3):275–285.

Du, J.-X., Wang, X.-F., and Zhang, G.-J. (2007). Leaf shape based plant species recognition. *Applied Mathematics and Computation*, 185(2):883 – 893.

Ehsanirad, A. (2010). Plant classification based on leaf recognition. *International Journal of Computer Science and Information Security*, 8(4):78–81.

Gao, L., Lin, X., Zhao, W., Chen, S., and Huang, H. (2010a). An algorithm of excising leafstalk while keeping its main body intact for leaf recognition. In *Image and Signal Processing (CISP), 2010 3rd International Congress on*, volume 6, pages 2732–2736.

Gao, L., Lin, X., Zhong, M., and Zeng, J. (2010b). A neural network classifier based on prior evolution and iterative approximation used for leaf recognition. In *Natural Computation (ICNC), 2010 Sixth International Conference on*, volume 2, pages 1038–1043.

Gonzalez, R. C., Woods, R. E., and Eddins, S. L. (2004). *Digital Image Processing Using MATLAB*. Pearson Prentice Hall Pearson Education, New Jersey, USA, 1st edition.

Haralick, R., Shanmugam, K., and Dinstein, I. (1973). Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, SMC-3(6):610–621.

Hu, M.-K. (1962). Visual pattern recognition by moment invariants. *Information Theory, IRE Transactions on*, 8(2):179–187.

Im, C., Nishida, H., and Kunii, T. (1998). Recognizing plant species by leaf shapes-a case study of the acer family. In *14th International Conference on Pattern Recognition*, volume 2, pages 1171–1173.

Kadir, A., Nugroho, L. E., Susanto, A., and Santosa, P. I. (2011). A comparative experiment of several shape methods in recognizing plants. *International Journal of Computer Science & Information Technology*, 3(3):256–263.

Kulkarni, A. H., Rai, D. H. M., Jahagirdar, D. K. A., and Upparamani, P. S. (2013). A leaf recognition technique for plant classification using rbpnn and zernike moments. *International Journal of Advanced Research in Computer and Communication Engineering*, 2(1):82–93.

Lee, C.-L. and Chen, S.-Y. (2006). Classification of leaf images. *International Journal of Imaging Systems and Technology*, 16(1):15–23.

Lee, K. B. and Hong, K. S. (2013). An implementation of leaf recognition system using leaf vein and shape. *International Journal of Bio-Science and Bio-Technology*, 5(2):57–65.

Lin, H. and Peng, H. (2008). Machine recognition for broad-leaved trees based on synthetic features of leaves using probabilistic neural network. In *Computer Science and Software Engineering, 2008 International Conference on*, volume 4, pages 871–877.

Machado, B. B., Casanova, D., Gonalves, W. N., and Bruno, O. M. (2013). Partial differential equations and fractal analysis to plant leaf identification. *Journal of Physics: Conference Series*, 410(1).

Man, Q.-K., Zheng, C.-H., Wang, X.-F., and Lin, F.-Y. (2008). Recognition of plant leaves using support vector machine. In *Advanced Intelligent Computing Theories and Applications. With Aspects of Contemporary Intelligent Computing Techniques*, volume 15 of *Communications in Computer and Information Science*, pages 192–199. Springer Berlin Heidelberg.

Pauwels, E. J., de Zeeuw, P. M., and Ranguelova, E. (2009). Computer-assisted tree taxonomy by automated image recognition. *Engineering Applications of Artificial Intelligence*, 22(1):26–31.

Singh, K., Gupta, I., and Gupta, S. (2010). Svm-bdt pnn and fourier moment technique for classification of leaf shape. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 3(4):67–78.

Valliammal, N. and Geethalakshmi, S. N. (2011). Hybrid image segmentation algorithm for leaf recognition and characterization. In *International Conference on Process Automation, Control and Computing*, pages 1–6.

Wang, Z., Chi, Z., and Feng, D. (2003). Shape based leaf image retrieval. *Vision, Image and Signal Processing, IEEE Proceedings*, 150(1):34–43.

Wu, Q., Zhou, C., and Wang, C. (2006). Feature extraction and automatic recognition of plant leaf using artificial neural network. *Avances en Ciencias de la Computacion*, pages 5–16.

Wu, S., Bao, F., Xu, E., Wang, Y.-X., Chang, Y.-F., and Xiang, Q.-L. (2007). A leaf recognition algorithm for plant classification using probabilistic neural network. In *IEEE International Symposium on Signal Processing and Information Technology*, pages 11–16.

Zhang, S. and Lei, Y.-K. (2011). Modified locally linear discriminant embedding for plant leaf recognition. *Neurocomputing*, 74(1415):2284 – 2290.

Zhang, X. and Zhang, F. (2008). Images features extraction

of tobacco leaves. In *Congress on Image and Signal Processing (CISP)*, volume 2, pages 773–776.

Zulkifli, Z., Saad, P., and Mohtar, I. (2011). Plant leaf identification using moment invariants amp; general regression neural network. In *11th International Conference on Hybrid Intelligent Systems (HIS)*, pages 430–435.