Journal of Advances and Scholarly Researches in Allied Education
Vol. 18, Issue No. 3, April-2021, ISSN 2230-7540

A Study of Classification of Leaf Shape Categorization of two Stages Based on Fusion

Anil Kumar[1*], Dr Ajay Agarwal[2]

[1] Research Scholar, Shri Venkateshwara University Gajraula (U.P)

[2] Professor Department of Physics, Shri Venkateshwara University Gajraula (U.P)

Abstract- The paper work only considered frontal and fresh leaves when creating leaf image databases. In future, CAP-LR can consider and analyze leaves that are wrinkled, occult and dry. Another difficult field of research can be considered in similar ways is color discoloration or discolored leaves. Advanced operations such as parallel workflow processing can be studied to increase the speed of the recognition of the leaf disorder for plant identification. Parallel task processing can be used to group algorithms together during the different recognition stages. This study proposed techniques for improving the operation of plant identification leaf recognition. Positive results from the different experiments show that the model proposals discriminate effectively against the various leaves and identify the right plant to match the image of the inserted leaf. The botanists can therefore safely use this to increase their efficiency in plant recognition and thus save valuable plants in order to improve the quality of human life and life on Earth.A system for improving the leaf image was proposed, called 'Enhanced wavelet-based demoizing with built-in edge enhancement and automatic contrast adjustment algorithm.The 197 method combines the wavelets, CLAHE (contrast adjustment), corner enhances and a relaxed middle filter (noise removal), with a single procedure to increase the visual quality of the leaf image. Texture-based color segmentation technique called 'Enhanced wavelet-based segmentation using the WCF method to extract the leaf image from its background. Five types of features are extracted during extraction: geometrical, texture, colour, fractal and leaf-like. These functions are combined to form the GLFS (Geometrically + Leaf), CLFS (Color + Leaf), TLFS (Texture + Leaf) and FLFS (Fractals + Leaf) functionalities. In addition to shared and merged operators to select optimal feature sets two selection algorithms for feature, the genetic algorithm, and the Kernel main component analysis algorithm have been coupled.

Keywords- Leaf Shape, Fusion, fresh leaves, segmentation, technique, leaf image

- X -

1. INTRODUCTION

Fusion strategy is referred to as the idea of combining the features taken from the patterns using different modes to improve the performance of the classification/cognition system. In several pattern recognition/classification applications this strategy has been thoroughly studied and proved with extensive experiments that fusion strategy approaches outperform non-fusion strategy approaches. In this context, it is possible, in order to improve the classification efficiency of the proposed medicinal plant classification system, to fuse the shape and texture features extracted from the proposed extraction techniques. The form and the texture of both features are considered jointly through fusion strategy in order to identify intrinsically different fluctuations between leaf samples of the same species. The hierarchical clustering of leaves by the feature level fusion belongs to the same species in order to achieve natural groups/clusters within the species. Each cluster is derived from a feature vector of an interval value symbolic data, as described in the above chapter. For example, the multiple representatives for a species based on variations in intra-class due to a merging of form and texture are utilized in order to more effectively represent a particular species. A decision level fusion is considered to be classified as one of the known species based on the majority vote principle. The unknown specimen is classified as one.Multi-stage classification is called classification of models based on characteristics extracted at various levels from different modalities. This idea was studied in a few applications to improve classification accuracy. In this work, we tried to study whether a two-stage classification technique was feasible to improve the accuracy of the classification of medicinal plants. Shape and texture characteristics of plant leaves at levels 1 and 2 respectively are considered. During classification, leaf texture information can be used to resolve the conflict between different species at level 1 due to a similar shape structure at level-2. Several experiments are conducted on different leaf datasets to demonstrate the performance of the proposed two-stage classification. This chapter contains the details of the two-stage classification framework proposed and the results from experiments obtained with the leaf data sets. A thoroughness experiment is carried out on different sheet datasets with all potential combinations of form and texture characteristics extracted from the proposed techniques in order to study the effectiveness of the proposed fusion strategy.

Several researchers have investigated the feasibility for different applications of various fusion strategies. However, in relation to plant classification, only few attempts have been reported. The proposal was to fuse information on a range of different levels (Ross and Jain, 2003) including feature levels, matching score levels and decisions to increase biometric systems performance. Several attempts have been made to fuse shape characteristics of herbs with color and texture, such as (Nam et al., 2008; Anama et al., 2010; Florindo et al., 2010; Kadir et al., 2011; Du et al., 2013; Chaki et al., 2015). However, no attempts to study the fusion of plant leaves and the fusion of the level of decision are reported. This motive has been used in this research work to present a new fusion strategy in which the classification of medicinal plant leaves is examined both at the fusion levels and at the decision-making levels. The overview of the methodology proposed in Fig. 1 is presented.

2. FEATURE LEVEL FUSION FOR CLUSTERING AND REPRESENTATION OF A LEAF

Because of many factors discussed in the previous chapters, leaf samples belong to the same species can be different in form and texture. In order to efficiently represent and classify the leaves, it is very important to capture such changes intraclass. In order to capture variations within the class, the concept of clustering based on the fusion of feature level is examined. A symbolic interval-value vector is derived from the leaf vectors of the cluster. The derived feature vector is able to collect the variations intraclass of leaf samples by form and structure in a cluster and collectively and efficiently represent the leaf samples. A detailed description of the classification based on fusion and the presentation of a leaf is provided under the following sections.

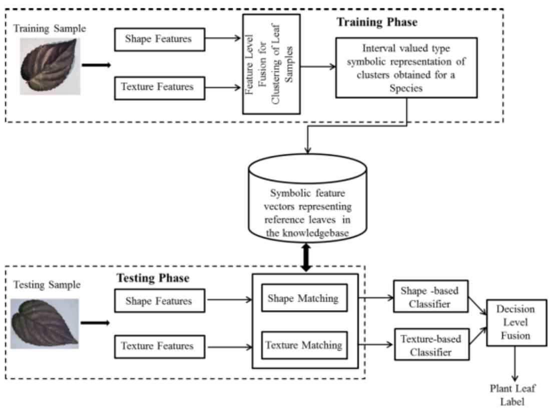

Fig. 1: Overview of the proposed fusion strategy

2.1 Fusion Based Clustering

The feature level fusion of shape and texture information of plant leaf samples is accomplished by simple concatenation of the feature vectors obtained from the shape and texture description techniques discussed in the earlier chapters.

If $\{SF_1, SF_2, SF_3, \ldots, SF_u\}$ and $\{TF_1, TF_2, TF_3, \ldots, TF_v\}$ are the shapes and texture feature vectors describing a particular leaf sample respectively. Then the feature level fusion is performed by simple concatenation of two feature vectors as follows:

$$F = \{SF_1, SF_2, SF_3, \ldots, SF_u, TF_1, TF_2, TF_3, \ldots, TF_v\} \quad (1)$$

In order to capture the joint intra-class variations, we propose to have multiple representatives for a class by grouping similar leaf samples into one group and choose a representative for that group within the class. The grouping is based on the resultant feature vector obtained through the feature level fusion of shape and texture feature vectors and by the use of hierarchical clustering technique

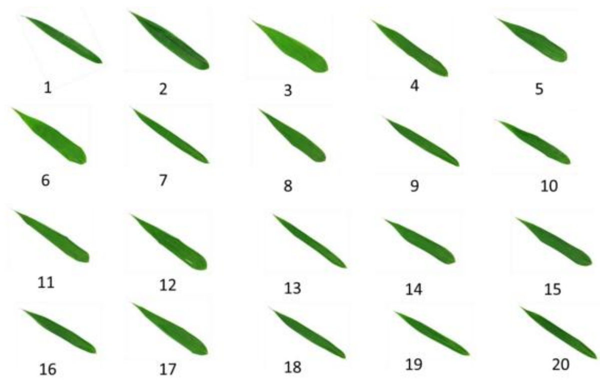

Fig. 2: Sample leaves of a particular plant species

Figures 2 show 20 leaf images of a specific plant species. Fig. 2 shows Fig. 2 clearly shows that both the form and texture of the leaf samples vary. In the case of the hierarchical clustering technique on this plant species we obtained a different number of clusters for (α = 0,1) to identify natural groups in the samples based on form, texture, combined form and textures. The index of samples and the graphs obtained by form, texture, and shape fused characteristics shows in Table 1. In Table 1 it is noted that when shape characteristics are used alone, 7 clusters are formed, 8 classes form, where texture characteristics alone are used and 6 clusters form together when shape and structure characteristics are used. The sample indexes 2, 3, 5, 14, 15 and 20 are also shown in Table 1 when form features alone have been used. The only way to combine indices 2 and 3 is to use texture features alone. However the sample indices 2, 3, 5 are one cluster, while the sample indices 14, 15 and 20 form another cluster when used in combination both shape and texture. This shows that the variations intra-class by form differ from intra-class fluctuations because of texture and is very possible to capture fluctuations in intra-class by form and texture in order to obtain finer and naturally more clusters that contain less inter-class variations. To achieve good classification accuracy, this criterion is essential.

Table 1: Results of clustering due to shape, texture and fusion of shape and texture features

Cluster Index	Sample indices forming the clusters		
	Using Shape Features	Using Texture Features	Using Shape & Texture Features
1	{1, 9}	{1,20}	{1,8,9,12}
2	{2, 3, 5, 14, 15, 20}	{2,3}	{2,3,5}
3	{4, 7, 10, 16}	{4,5}	{4,7,10,16}
4	{6, 19}	{6,16}	{6,19}
5	{8, 12}	{7,8}	{11,13,17,18}
6	{11, 13}	{9,10,13}	{14,15,20}
7	{17, 18}	{11,12,14,15,18}	Nil
8	Nil	{17,19}	Nil

2.2 Leaf Representation

Once the natural clusters of leaf samples in a particular plant species are obtained, the representative vector for each cluster is obtained by aggregating the corresponding features of plant leaf. The reference feature vectors in the knowledgebase representing leaf samples in terms of shape and texture, belonging to jth cluster of ith class is of interval valued type as follows:

$$RS^{(ij)} = \left\{\left[SF_1^{(ij)-}, SF_1^{(ij)+}\right], \left[SF_2^{(ij)-}, SF_2^{(ij)+}\right], \ldots, \left[SF_u^{(ij)-}, SF_u^{(ij)+}\right]\right\} \quad (2)$$

$$RT^{(ij)} = \left\{\left[TF_1^{(ij)-}, TF_1^{(ij)+}\right], \left[TF_2^{(ij)-}, TF_2^{(ij)+}\right], \ldots, \left[TF_v^{(ij)-}, TF_v^{(ij)+}\right]\right\} \quad (3)$$

Likewise, we compute symbolic feature vectors for all the clusters of all the classes. Collection of all the symbolic feature vectors effectively represents leaf samples of all plant species of a dataset in the knowledgebase.

3. LEAF MATCHING AND DECISION LEVEL FUSION FOR CLASSIFICATION

In order to classify a test leaf sample, shape and texture features are extracted as discussed in the previous chapters and the obtained crisp feature vector is compared with the symbolic feature vectors of reference leaf samples in the knowledgebase. Let $TS = \{SF_1, SF_2, SF_3, \ldots, SF_u\}$ and $TF = \{TF_1, TF_2, TF_3, \ldots, TF_v\}$ be the shape and texture feature vectors of dimension u and v respectively, representing the test leaf sample to be classified. The given test leaf sample is compared with all the reference leaf samples stored in the knowledgebase and the matching score is computed.

Two different proximity measurements are studied individually in order to calculate the matching score between test and reference leaf samples on the basis of form and texture. The symbolic difference of chi-square and the symbolic measure as described in the section are used to match texture and formal characteristics respectively. The technique for classifying the K-Nearest Neighbor and fusion of decision level is used to classify the specimen as one of the known species of plant species. We take into account the value 3 for K and determine independently the top 3 plant species indices for form and texture. In order to classify the given test sheet sample the majority voting rule (Kuncheva, 2004) will then be applied.

4. TWO-STAGE CLASSIFICATION

Plant leaves can be classified by means of shape and texture features. The feature of plant leaves is straightforward, easy to carry out and not computer-priced, compared to the feature of plant leaves based on their texture. Since form descriptors only use leaf contour information, lighting conditions and color variations are not restricted. In classification, form descriptors will also be very effective, provided that the shape of the plant leaves differs considerably. Plant leaves are classified in terms of form and are computationally efficient and take less time than texture. In most cases it is therefore desirable to use form-based classifiers. One of the major drawbacks of a forms-based classifier, however, is that it is not able to distinguish between two plant species if their leaf forms are similar, even though their texture or color differs. Therefore, the design classifier must be based upon the leaf texture or color of the plant classification. However, the design of a leaf texture/color-based classificator is costly for computer time and storage of data. For example, it would be more appropriate to devise a two-stage classificator, where the classification based on forms is performed in the first stage and texture-based classifications are performed in the second. In the second stage, only those plant species, which due to their similar forms were not correctly classified in the first stage are considered and thus we are reducing the time needed for computation of texture-based classificators. Some plant species may be present, with a very similar leaf form and texture and cannot be properly classified

Anil Kumar[1]*, Dr Ajay Agarwal[2]

even if two-stage classification is followed. In order to characterize the plant and then deal with the classification tasks, we must consider other methods, such as the flower, the fruit, the bark etc.

As shown in Fig. 3, you can view an overview of the proposed two-stage classification system. From Fig. 3 we can observe that the two-phased classification Framework proposed includes three steps, namely I Classification of the sheet form, (ii) Classification based on shapes and (iii) Classification based upon textures. The following paragraphs describe these steps in detail.

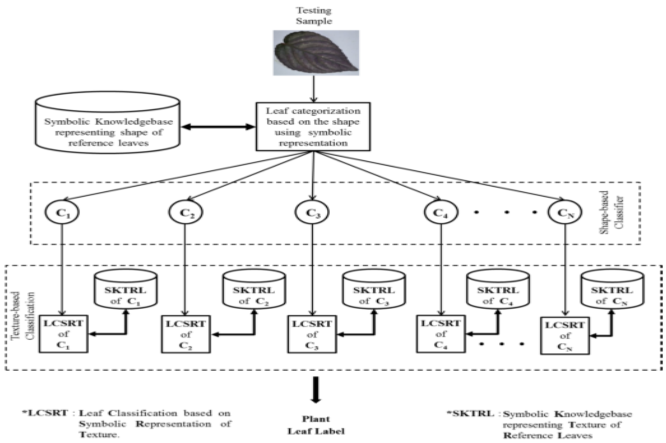

Fig. 3: Overview of the proposed two-stage classification system

4.1 Leaf Shape Categorization

In this step, plant species with similar structures in leaf form are identified as clusters of forms. This task is performed by a supervised classification technique using a form-based classifier. There are enough leaf samples of each species to study the intra-class changes. Subgroups for a particular species are obtained based on the intra-class variations. In the 60:40 ratio, each sub-group selects samples for training and testing. A symbolic vector of the valued interval is derived using training samples that collectively represent the leaf in a subgroup. Using the training samples used to obtain the confusion matrix, multiple representatives are obtained for each plant species by means of its subgroup. The confusion matrix for the whole data set shows the correctly classified plant species and which species are in conflict. All these conflicting plant species are identified and grouped according to their associations. Thus, plant species which are similar in terms of their leaf structure are categorized into one class and in the second phase are further classified as textures. We take a Swedish leaf dataset, which consist of 15 different species each with 75 samples, as example to clearly understand the two-stage classification process. If the natural cluster of the selection of samples and training or testing samples in a ratio of 60:40 is obtained for each species considering these75 samples, we can obtain approximately 30 testing samples and 45 training samples and then represent the cluster. The primary diagonal matrix entries highlighted with the blue color represent the percentage of their respective species of correctly classified leaf samples. The yellow-colored percentage of leaf samples, misclassified with other plant species. The plant species C1 has contradictions to plant species C3, C9 and C12 can be observed in Table 2. Therefore, the C3, C9 and C12 plant species are considered as conflict classes with C1. It can also be seen that there are no conflicts with other classes in plant species C7 and C10. They are therefore treated as separate categories individually.

4.2 Shape-based Classification

Once plant species are categorized based on their leaf shape structure, classification of an unknown leaf sample is accomplished by extracting its shape features and identifying the category to which it belongs to by using the shape-based classifiers as discussed in the previous chapters. The target shape category index for an unknown leaf sample is identified in this step.

Table 2: Confusion Matrix obtained for Swedish leaf dataset using shape-based classifier

	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}	C_{15}
C_1	**0.87**	0.00	**0.03**	0.00	0.00	0.00	0.00	0.00	**0.07**	0.00	0.00	**0.03**	0.00	0.00	0.00
C_2	0.00	**0.91**	0.00	**0.03**	0.00	**0.03**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.03**	0.00
C_3	**0.03**	0.00	**0.84**	0.00	0.00	**0.03**	0.00	0.00	**0.10**	0.00	0.00	0.00	0.00	0.00	0.00
C_4	0.00	0.00	**0.07**	**0.90**	**0.03**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C_5	0.00	0.00	0.00	0.00	**0.84**	**0.03**	0.00	**0.03**	0.00	0.00	0.00	0.00	0.00	0.00	**0.10**
C_6	**0.10**	0.00	0.00	0.00	0.00	**0.87**	0.00	0.00	0.00	0.00	0.00	**0.03**	0.00	0.00	0.00
C_7	0.00	0.00	0.00	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C_8	0.00	0.00	0.00	0.00	**0.03**	0.00	0.00	**0.93**	0.00	0.00	**0.03**	0.00	0.00	0.00	0.00
C_9	**0.03**	0.00	**0.07**	0.00	0.00	0.00	0.00	0.00	**0.87**	0.00	**0.03**	0.00	0.00	0.00	0.00
C_{10}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	0.00	0.00
C_{11}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	0.00
C_{12}	**0.03**	0.00	0.00	0.00	0.00	**0.03**	0.00	**0.03**	**0.03**	0.00	0.00	**0.86**	0.00	0.00	0.00
C_{13}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.97**	0.00	**0.03**
C_{14}	**0.10**	0.00	0.00	**0.06**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.84**	0.00
C_{15}	**0.03**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.03**	0.00	**0.03**	0.00	0.00	0.00	**0.90**

5. CONCLUSION

More than 1.7 million species (humans, plants and algae) of living organisms have been found world-wide, of which plants play a vital role in human life. Plants can exist everywhere and are a vital resource for human well-being. For the development of human society, most plants contain significant information and are regarded as essential resources for human welfare. Plants are widely used because they form the foundation of the food chain, and plants are the source of many medicines. Plants are also crucial to protecting the environment. Even after some innovative developments in the field of botany, many plants still have to be found, identified and used. Unknown plants are treasures waiting to be found, it is well known. Ethno-botanists today unite the regions of the world in search of future medicines and products of agriculture.They explore the functional features and the association of plants within ecosystems to understand the need for plant resource management diversity. 21st century scientists are exploring how genetic diversity and ecologically sensitive issues like population feed and disease management are needed. The identification and classification of plants are two major plant aspects of plant taxonomy that play a vital role in these efforts.

Plant Identification is the determination of the identity of an unknown plant in comparison with previously collected specimen. The process of recognition connects the specimen with a botanical name. Once this connection is established, related details like name and other properties of the plant can be easily obtained.

Plant Classification is the placing of known plants into groups or categories to show some relationship. They use features that can be used to group plants into a known hierarchy.

6. REFERENCES

1. Backes et al., (2011) Backes, A. R., Martinez, A. S., and Bruno, O. M. (2011). Texture analysis using graphs generated by deterministic partially self-avoiding walks. Pattern Recognition, 44(8):1684–1689.
2. Belhumeur et al., (2008) Belhumeur, P. N., Chen, D., Feiner, S., Jacobs, D. W., Kress, W. J., Ling, H., Lopez, I. C., Ramamoorthi, R., Sheorey, S., White, S., and Zhang, L. (2008). Searching the world's herbaria: A system for visual identification of plant species. In ECCV (4), pages 116–129.
3. Belongie et al., (2002) Belongie, S., Malik, J., and Puzicha, J. (2002). Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell., 24(4):509–522.
4. Bhattacharya et al., (2012) Bhattacharya, U., Shridhar, M., Parui, S. K., Sen, P., and Chaudhuri, B. (2012). Offline recognition of handwritten bangla characters: an efficient two-stage approach. Pattern Analysis and Applications, 15(4):445–458.
5. Bisby, F. A. (2000). The quiet revolution: biodiversity informatics and the internet. Science, 289(5488):2309–2312.
6. Bock and Diday, (2000) Bock, H.-H.andDiday, E. (2000). Analysis of symbolic data: exploratory methods for extracting statistical information from complex data. Springer Science & Business Media.
7. Bruno et al., (2008) Bruno, O. M., de Oliveira Plotze, R., Falvo, M., and de Castro, M. (2008). Fractal dimension applied to plant identification. Inf. Sci., 178(12):2722–2733.
8. Casanova et al., (2009) Casanova, D., de MesquitaS´a Junior, J. J., and Bruno, O. M. (2009). Plant leaf identification using gabor wavelets. Int. J. Imaging Systems and Technology, 19(3):236–243.
9. Cerutti et al., (2013) Cerutti, G., Tougne, L., Mille, J., Vacavant, A., and Coquin, D. (2013). Understanding leaves in natural images - a model-based approach for tree species identification. Computer Vision and Image Understanding, 117(10):1482– 1501.
10. Chaki et al., (2015) Chaki, J., Parekh, R., and Bhattacharya, S. (2015). Plant leaf recognition using texture and shape features with neural classifiers. Pattern Recognition Letters, 58:61–68.
11. Cope et al., (2012) Cope, J. S., Corney, D. P. A., Clark, J. Y., Remagnino, P., and Wilkin, P. (2012). Plant species identification using digital morphometrics: A review. Expert Syst. Appl., 39(8):7562–7573.

Corresponding Author

Anil Kumar*

Research Scholar, Shri Venkateshwara University Gajraula (U.P)