ABSTRACT

This research work has examined the role of various feature extraction techniques which has been utilized for extracting characteristic features from digital leaf images of different plant species. These characteristic features extracted have been further subjected to the process of plant species classification-a process to identify the different plants through the use of techniques available in image processing. This study has rigorously worked with texture based feature extraction methods like GLCM (Gray Level Co-Occurrence Matrix) and Gabor Filter based features extraction method on one end, statistical feature extraction form the digital leaf images on the other end and has gone one step further in fusing two different set of features like statistical and image directionality features together to evaluate the improvement in discrimination power of the different classifiers.

This research work has been motivated by three research questions:

- How to extract relevant and useful features from the digital leaf images of the different plant species?
- How to design a suitable comparison methodology which could make comparative decisions on the class of the new leaf image, on the basis of newly extracted characteristic features from that of the dataset features already in the database?
- To find if the images of the ventral side of the leaves of different plant species can be used for discrimination of the plant species?

To achieve an amicable solution to the above research questions, a new database of ten different plant species from the local vicinity containing both the dorsal and the ventral leaf images was prepared, as there was no dataset with both dorsal and ventral leaf images available on the internet to till date. The whole image dataset was divided into three parts: dorsal image dataset, ventral image dataset and combined dorsal-ventral image dataset. After the process of feature extraction, the features were subjected to the process of discrimination and it has been observed that for all the three different kinds of datasets, the average predictive accuracy that has been

achieved is 89% using Gabor based texture features for dorsal images set, which is fairly higher as compared to the GLCM based texture feature results. The ventral side dataset has proved to be a better discriminator with an average accuracy of about 92.08% with Gabor based texture features. A predictive accuracy value of 88.54% is achievable through the process of fusion. This work has utilized a number of feature selection techniques like: BFS, CFS, Chisq, OneR, Random Forest, ReliefF, Hill Climbing and evolutionary optimization techniques like Particle swarm Optimization (PSO) etc. By using the feature selection techniques, a dataset with relevant features, which is devoid of redundancy of any sort, has been prepared that improves upon the overall predictive accuracy of the plant species. Throughout the entire research work, it has been observed, on the basis of the average predictive accuracy results, that the ventral side is a better alternative as compared to the dorsal side of the leaf image for discrimination purpose using computer vision techniques and feature selection process not only reduces the size of the dataset but also the computational time to discriminate the images gets reduced and the overall predictive accuracy value improves considerably.

Keywords: Plants, Dorsal and ventral leaf images, GLCM, Gabor Filter, Texture, feature selection, PSO, Classification of leaves