Journal of King Saud University – Computer and Information Sciences xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

Journal of King Saud University – Computer and Information Sciences

journal homepage: www.sciencedirect.com

Plant leaf classification using multiple descriptors: A hierarchical approach

Jyotismita Chaki [a,*], Ranjan Parekh [b], Samar Bhattacharya [c]

[a] School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, Orissa, India
[b] School of Education Technology, Jadavpur University, Kolkata, India
[c] Dept. of Electrical Engineering, Jadavpur University, Kolkata, India

ARTICLE INFO

Article history:
Received 6 October 2017
Revised 11 January 2018
Accepted 22 January 2018
Available online xxxx

Keywords:
Hierarchical architecture
Feature based Shape Selection Template
Heterogeneous plant leaf
Shape
Texture
Non-green leaf

ABSTRACT

The present work proposes another path for classification of plant species from digital leaf images. Plant leaves can have an assortment of unmistakable elements like green and non-green hue, simple and compound shape and distinctive vein designed surfaces, a solitary arrangement of elements may not be sufficiently adequate for a viable classification of heterogeneous plant sorts. A hierarchical architectural design is proposed where numerous components are joined together for a more powerful and strong classification of the visual data. The study likewise incorporates the arrangements of customization of the feature extraction modules and classifiers for best execution. The database itself is sectioned in light of conspicuous components by visual discriminators, as this enhances proficiency. As new layers can be added to the current system to take into account up to this point obscure leaves with new qualities, the design likewise provides options of adaptability. Another Feature based Shape Selection Template (FSST) is proposed for the choice of shape features for various sorts of leaves. Broad examinations are completed on two openly accessible databases including green, non-green, simple and compound leaves with variations in shape, size and designs about exhibit the advantages of the proposed strategy over best in class procedures.

1. Introduction

Plants are the one of the basic component of the earth responsible for protecting the World's environment. They give sustenance, protect, medicines, fuel and keep up a sound breathable climate. Be that as it may, as of late an ever-increasing number of plants are at the very edge of termination because of ceaseless de-forestation. Thus, to monitor the plants, assembling a plant database for speedy and effective grouping and classification is an essential stride. The vast majority of these systems depend on extraction of visual components like hue, texture and shape and their portrayals as information models for correlations and classification. Albeit different parts of a plant like blossom, bud, natural product, seed, root can be utilized for distinguishing, leaf based classification is the most widely recognized and viable approach.

A number of visual features, data modeling techniques and classifiers have been proposed for plant leaf classification. The Manifold learning based dimensionality reduction algorithm is used (Zhang et al., 2016) in plant leaf recognition as the algorithms can select a subset of effective and efficient discriminative features in the leaf images. For plant leaf recognition a dimensionality reduction method based on local discriminative tangent space alignment (LDTSA) is used where the manifold learning based dimensionality reduction algorithm is applied to reduce the size of the neighborhood matrix generated from the within class neighborhood and between class neighborhood is estimated. Deng et al. (2016) focuses on the spectral classification of weeds and crops for detecting the weeds in crop fields. The Principal Component Analysis (PCA) is used to determine the characteristic wavelengths (CW). Anjomshoae and Rahim (2016) used a template-based method for overlapping rubber tree leaf identification. Initially, the key point based feature extraction method is adopted using the Scale Invariant Feature Transform (SIFT). The steps used in the SIFT method is finding the scale space extreme by the difference of Gaussian, key point localization by principal curvature,

* Corresponding author.
E-mail address: jyotismita.chakifcs@kiit.ac.in (J. Chaki).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.jksuci.2018.01.007
1319-1578/

Please cite this article in press as: Chaki, J., et al. Plant leaf classification using multiple descriptors: A hierarchical approach. Journal of King Saud University – Computer and Information Sciences (2018), https://doi.org/10.1016/j.jksuci.2018.01.007

orientation assignment by gradient directions and key point descriptor. An automated identification of plant species using leaf shape descriptor used by Salve et al. (2016a,b) addresses the automatic classification of plants and simplifies taxonomic classification process. In this research work, the authors use Zernike moments (ZM) and Histogram of Oriented Gradient (HOG) method as a shape descriptor. Scharr et al. (2016) compares several leaf segmentation solutions on a unique and first-of-its-kind dataset containing images from typical phenotyping experiments. Four methods are presented: three segment leaves by processing the distance transform in an unsupervised fashion and the other via optimal template selection and Chamfer matching. De Souzaa et al. (2016) uses the simulated annealing, differential evolution and particle swarm optimization methods, which is based on the silhouette measure, to achieve the set of optimal parameters for leaf shape characterization. The combination of texture features and shape features is used by Liu and Kan (2016) for the identification of plant leaf. Texture features are derived from local binary patterns, Gabor filters and gray level co-occurrence matrix while shape feature vector is modeled using Hu Moment invariants and Fourier descriptors. Modified Local binary patterns (MLBP) approach is used by Naresh and Nagendraswamy (2016) for classification of plant leaves based on texture features. Here mean and standard deviation of the pixels is considered instead of considering a hard threshold like normal LBP. The Angle View Projection (AVP) is used by Prasad et al. (2016) for plant leaf identification. The AVP shape profile curve (a set of four shapelets) is extracted from the leaf images. The 1-D Discrete Cosine Transform (DCT) compactness is applied over the 1-D AVP shape curve to extract the features of the leaf image. Caoa et al. (2016) uses R-angle for the leaf shape characterization. R-angle describes the curvature of the contour by measuring the angle between the intersections of the shape contour with a circle of radius R centered at points sampled around the contour. Varying the parameter R of the proposed R-angle describes the notation of scale, which indicates a coarse-to-fine description of the local curvature. Visual parameters include length, width, area, perimeter is used by Sakai et al. (1996) and leaf contour shape is used by Wang et al. (2000) for the classification of plant leaves. Different data modeling techniques used include orthogonal discriminant projection (Zhang et al., 2016), the focus is on the spectral classification of weeds and crops for detecting the weeds in crop fields. The Principal Component Analysis (PCA) is used to determine the characteristic wavelengths (Deng et al., 2016), multidimensional embedding sequence similarity (Fotopoulou et al., 2013), fuzzy logic (Wang and Feng, 2002), Fourier descriptors (Yang and Wang, 2012), Recognizing leaf images based on Ring Projection Wavelet Fractal Feature is used by Wang et al. (2010), Zernike moments is used by Kadir et al. (2012) to build foliage plant identification systems. Zernike moments were combined with other features: geometric features, color moments and gray-level co-occurrence matrix (GLCM). The geometric features include aspect ratio, circularity, irregularity, solidity, convexity and two types of vein features are used. The vein features is constructed by using the morphological opening operation. The color moment features include the mean, standard deviation, skewness and kurtosis. After that the GLCM based features are extracted which include the energy, contrast, local homogeneity pair of pixels, entropy and correlation, In Salve et al. (2016a,b) an automated identification of plant species using leaf shape descriptor addresses the automatic classification of plants and simplifies taxonomic classification process. In this research uses Zernike moments (ZM) and Histogram of Oriented Gradient (HOG) method as a shape descriptor, local binary descriptors (LBD) uses (Wang et al., 2014; Le et al., 2014) used kernel descriptor (KDES) based plant leaf identification. Before the feature extraction, the leaf images are segmented using the Watershed algorithm. After that the images are converted to the grayscale image. To extract the KDES feature, first the patch level features are extracted from the leaf images. Here three types of kernels are considered: gradient, local binary pattern and color. After extracting the patch level features, the K-means algorithm is applied to build the dictionary. After that the features are extracted from the image level using the spatial pyramid matching throughout several layers, (Markos et al., 2015) uses morphological characterization, a variety of classifiers and comparison metrics have been used viz. neural networks (Aakif and Khan, 2015; Kumar et al., 2012) describes the mobile app for identifying plant species using automatic visual recognition. First of all a binary leaf/non-leaf classifier is applied to all inputs. After that color based segmentation is applied to segment the leaf from an un-textured background. After the segmentation, the stems are removed from the binary images by the top-hat morphological operation. After the preprocessing the features are extracted by the curvature of the leaf's contour over multiple scales. The histograms of the curvature values at each scale are computed and those histograms are concatenated to form the histograms of curvature over scale feature. In Kalyoncu and Toygar (2015) the leaf image recognition is done using geometric features, multiscale distance matrix and moment invariant. Before the feature extraction segmentation of the leaf image is done by the simple adaptive threshold method over the blue channel. For the feature extraction moment invariant, convexity, perimeter ratio, multiscale distance matrix, average margin distance and the average margin peak height, peak height variance, average peak distance and peak distance variance are extracted from the contour of the leaf image. Chaki et al. (2015) uses Shape based modeling scheme based on curvelet transform and invariant moments and texture based modeling scheme based on Gabor filter and Gray Level Co-occurrence Matrix(GLCM) with neuro-fuzzy classifier.

In the vast majority of the works inspected, a solitary arrangement of elements and classifiers have been utilized to segregate between leaf classes in a given dataset of plant leaf pictures. Such an approach functions admirably when the leaf classes in the dataset are for the most part homogeneous and could be separated by a single arrangement of components. However, in Nature, plant leaves can have for all intents and purposes unbounded sorts of varieties in geometric arrangements, form shapes, hue, texture examples and additionally could be basic, compound, twisted or even fragmented (parts of the leaf missing). To take into account such heterogeneous varieties, a solitary approach is normally insufficient, rather a hierarchical architectural approach is essential where each layer utilizes a particular visual characteristic and is connected to an arrangement of custom classifiers. Results from various layers can in this manner be combined together for a more complete comprehension of the leaf features which prompts a powerful classification plot. Additionally, such designs can have arrangements for versatility, by adding new layers to take into account new leaves with various attributes.

This paper establishes the framework of such a hierarchical architecture approach for plant leaf classification and exhibits its advantages over a solitary included plan. The arrangement of the paper is as per the following area: segment 2 diagrams the proposed approach with discussions on feature calculation and order plans, segment 3 provides points of experimentation and results, segment 4 analyzes the proposed approach opposite other contemporary methodologies, while segment 5 mentions the general conclusions and degrees for future research.

2. Proposed approach

A block diagram depicting major functional blocks and data flow pathways of the proposed approach is shown in Fig. 1.

Please cite this article in press as: Chaki, J., et al. Plant leaf classification using multiple descriptors: A hierarchical approach. Journal of King Saud University – Computer and Information Sciences (2018), https://doi.org/10.1016/j.jksuci.2018.01.007