A Leaf Recognition Approach to Plant Classification Using Machine Learning

Conference Paper · July 2018

DOI: 10.1109/NAECON.2018.8556785

3 authors:

Redha Ali
University of Dayton

Russell Hardie
University of Dayton

Almabrok Essa
Cleveland State University

A Leaf Recognition Approach to Plant Classification Using Machine Learning

Redha Ali, Russell Hardie
Department of Electrical and Computer Engineering
University of Dayton
300 College Park, Dayton, Ohio 45469
{almahdir1, rhardie1}@udayton.edu

Almabrok Essa
Department of Electrical Engineering and Computer Science
Cleveland State University
2121 Euclid Ave, Cleveland, OH 44115
a.essa@csuohio.edu

***Abstract*—The identification of plants is a very important component of workflows in plant ecological research. This paper presents an automated leaf recognition method for plant identification. The proposed technique is simple and computationally efficient. It is based on a combination of two types of texture features, named Bag-of-features (BOF) and Local Binary Pattern (LBP). These features are utilized as inputs to a decision-making model that is based on a multiclass Support Vector Machine (SVM) classifier. The introduced method is evaluated on a publicly available leaf image database. The experimental results demonstrate that our proposed method is the highly efficient technique for plant recognition.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

Automated plant recognition is an important research topic in machine learning and computer vision. Various research studies have been conducted to address the problems of the plant classification. Several approaches are presented in [1]. Before the invention of digital cameras and computerized systems, people were utilizing their knowledge and skill to distinguish between the different types of medical plants. The risk of using the wrong plant for medicine extraction rises with the lack of experience, and that can cause the fatal error that can cause the death of some patients. The presence of digital devices, artificial intelligence, and machine vision has encouraged the botanists and computer scientists to develop automatic systems for plant classification or recognition [2].

Image processing and computer vision are still an active in the research field with many practical applications [5], [6]. In recent years, numerous researches have shown that the shapes of plant leaves are important features. The leaves provide rich information for the computer-aided plant classification systems that are based on the leaf shape [7], [8]. The texture features of the plant leaf are also significant features used to classify plants [9]. Hence, many methods of feature extraction have been proposed for shape and texture features such as the shape context with dynamic programming (SD+DP) [22], orthogonal locally discriminant spline embedding (OLDSE) introduced by Lei et al. [23], Fourier descriptor for shape classification [22], and MEW [24]. Hu et al. [25] introduced the multi-scale matrix distance matrix(MMD) for plant recognition. In 2007, a shape classification technique using the inner distance shape context with dynamic programming (ISDC) was proposed in [26]. There are several simple algorithms described in [4], [29]–[31]. In this paper, we describe our method for the plant identification using a novel leaf recognition approach. Our approach has two main stages, as shown in Fig. 1.. The two stages are, 1) Feature extraction and 2) classification. The feature extraction strategy uses a combination of two different texture techniques, bag-of-features (BOF) and local binary pattern (LBP). The motivation for combining BOF and LBP is that this combination inherits the advantages of BOF, which are aggregates local components to form a global histogram characterization. In addition, it inherits the computational efficiency of LBP and avoids the limitations of LBP as well. First, we extract the image features using the BOF algorithm. This BOF method is based on vector quantization of affine invariant descriptors of image patches. The key points can be found using a speeded-up robust features (SURF) detector which also provides excellent scale-invariant features. Then, we iteratively group the descriptors into k clusters by using the k-means clustering technique.

Each cluster center can be viewed as a visual word (or feature), then the k-means is employed to achieve low-dimensional histograms and to evaluate the similarity of these histograms. After that, the histograms are accumulated into a feature vector for each image. Second, the LBP technique is used to extract the image features based on the leaf textures and form another histogram for each image. Then these two histograms are concatenated to form the final feature vector. This final feature set is fed to a decision-making model based on a multiclass support vector machine (SVM) classifier.

The remainder of this report is organized as follows: Section II introduces the use of BOF and LBP for feature extraction. We provide a brief description of the dataset used in this research in Section III. Experimental results are presented in Section IV. Finally, we offer conclusions in Section VI.

II. METHODOLOGY

In this section, we present featuresof two extraction methods, namely BOF and LBP. It is based on two different descriptors that are commonly used for image recognition

978-1-5386-6557-2/18/$31.00 ©2018 IEEE

tasks, which are speeded-up robust features (SURF) and LBP technique.

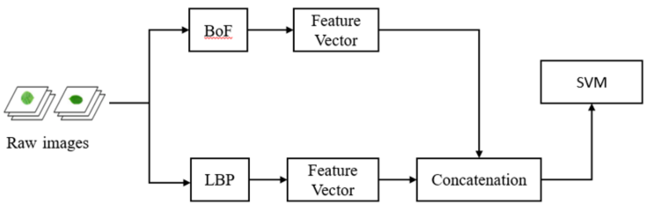

Fig. 1. Block Diagram of the Process of Leaf Recognition Approach

A. BOF Approach

Recently, the BOF model has received significant attention in the computer vision community because of the simplicity, robustness, and good practical performance [10]. BOF model is a representation used in natural language processing and information retrieval which represent an image as an orderless collection of local features [11], [12]. Since images do not have discrete words, we treat the images like a collection of visual words derived from key points. In order to extract the BOF feature from the image data set we have to involve the following steps:

- automatically detects points of interest in the image.
- computes local descriptors around each interest point.
- quantizes or clusters the descriptors to construct the visual vocabulary using unsupervised learning process.
- finds the occurrences of each image to construct a histogram of feature frequencies (or the BOF feature).

Fig. 2 illustrated these four steps to extract the BOF feature from images.

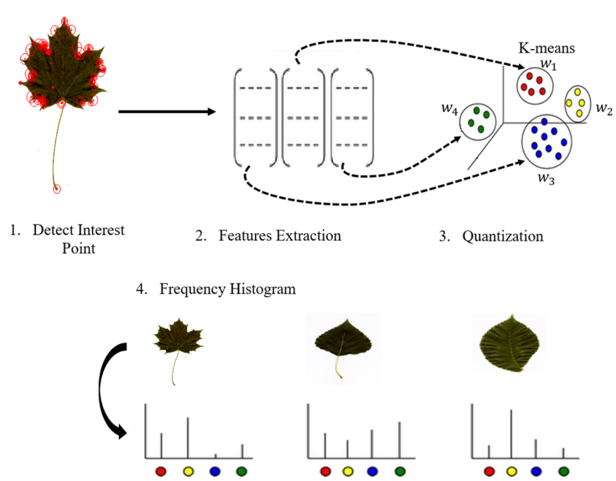

Fig. 2. Four steps for constructing the bag-of-feature for image representation.

1) Interest Point Detection: The first step of the BOF methodology is the interest point or regions detection. Harris Corner detector probably is the most widely used detector which was proposed back in 1988 [13]. However, Harris Corners are not scale-invariant or rotation-invariant interest point detectors [14]. Several well-known scale-invariant interest point detectors have been proposed in [15] which maximize the entropy within the salient regions that operate across feature space and scale. However, these detectors fall outside the scope of our research.

Since detector feature points is an important step, we use SURF detector which is a Hessian-based blob detector to find interest points. Mikolajczyk et al. [16], they compare six types of well-known detectors, and they conclude that the Hessian-based detector performs best. SURF uses the Hessian matrix because of its excellent performance in computation time and accuracy. SURF is more stable and reputable than their Harris-based counterparts [14]. The Fast-Hessian detector is described in detail in the SURF paper [14].

2) Local Descriptors: The goal of a descriptor is to produce a unique and robust description of the feature; a descriptor can be created based on the area surrounding the region point. The Scale Invariant Feature Transform (SIFT) descriptor is one of the most widely used descriptors [17]. It combines crudely localized information and the distribution of gradient in the detected regions. SIFT utilizes a cascading filtering strategy to identify the interest points in the images by constructing a cascade of the Difference of Gaussians (DoG). The feature descriptors are created based on the area surrounding the interest points after the interest points localized. The descriptor dimensionality of the SIFT is 128. These descriptors should carry the essential distinct information for their corresponding interest points.

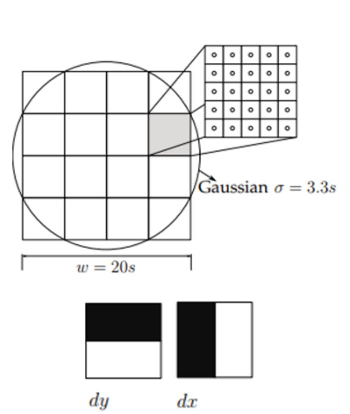

Fig. 3. Shows a $20s$ window is split up into 4×4 sub-regions.The wavelets response is black and white areas corresponds to a weight -1 and 1 for the Haar kernels with a filter size of 2s

The SURF descriptor uses Haar wavelet responses and can be calculated efficiently with entire images. SIFT employs a different scheme for descriptors which is based on the Hough transformation. The SURF descriptors are robust to rotations and an upright version, which is referred to as 'upright SURF' $(U-SURF)$. The SURF size window is $20s$. Fig. 3 shows the example of square regions. The region area is split up regularly into 44 sub-regions. The x denotes the values of a wavelet response in the horizontal direction, and y denotes the vertical direction of the Haar wavelet response in the (filter size $2s$). The wavelet response is referred to as dx and dy respectively.

$$v = \{\sum dx, \sum \mid dx \mid, \sum dy, \sum \mid dy \mid\} \qquad (1)$$

Note that the variable v is a feature vector. The wavelet responses dx and dy are computed at 5×5 sampled point and then summed up over each sub-region to create the feature vector. The interest area is weighted with a Gaussian ($\sigma = 3.3s$) and centered at the interesting point. This provides robustness for deformations and translations [18]. The descriptor for a interest point is the 16 vectors for all 4×4 sub-regions concatenated.

3) Visual Word Generation/Vector Quantization: After the interest points are detected, and their features are extracted with the SURF descriptor, the final step of the BOF model is vector quantization. In general, the unsupervised learning algorithm such as K-means is employed to find the closest cluster for each data point and construct the codebook.

First, the K-means initialize a set of k clusters; and all data points will be vector quantized against these clusters. The clusters are updated by computing the sum of squared Euclidean distances between the set of data points and their nearest cluster centers. These steps are iterated until the mean clusters do not change anymore. Then, the visual-word vector of each image comprises the presence or absence features of each visual word in the image. Finally, to construct the final features vector (or a histogram of word frequencies) of the BOF model, we determined the occurrences in the image of each word in the codebook. From the codebook, we produce a histogram per image. The generated histogram has as many bins as visual words. Each bin yields the number of clusters assigned to this visual word. Then the assignments are aggregated over the image to obtain the image representation [19].

B. LBP

The original LBP introduced in [20] is a powerful and efficient illumination invariant texture primitive. The histogram of the LBP computed over the sub-regions to obtain the texture description. Divide the input image into local regions, usually 16x16 pixels. The LBP value of each pixel in the local areas of the input image is computed with respect to its neighborhood. A histogram of each local sub-region is calculated and then concatenated to construct the final LBP features vector.

C. SVM

After we constructed the final feature set, we fed the feature set to a decision-making model based SVM classifier for plant identification. SVM is one of a discriminative type of classification technique based on the structural risk minimization principle. A separating hyperplane well defines the boundaries for each class. SVM creates the best separating hyperplane that separates the training data by a maximal margin and classifies the test set based on the plane it falls under [3]. The support vectors are the features points that are nearest to the separating hyperplane. Several types of the kernel can be used for the final decision. in this paper, we mainly concentrate on SVM with a Gaussian kernel function.

III. Swedish Leaf Dataset

The introduced technique is trained and tested on one of the publically available datasets. The Swedish leaf dataset [21] contains leaves from 15 tree classes shown in Fig 4, with 75 images per species. All images are 24-bit RGB uncompressed with different dimensions. The dataset divided into 70% for training and 30% are used for testing.

Fig. 4. Shows the fifteen samples from the Swedish Leaf dataset, one image per species.

IV. EXPERIMENTAL RESULTS

In this section, we show the experimental results using the introduced technique for classifying the testing dataset of the Swedish Leaf. In order to demonstrate the performance of our approach, a comparison was conducted with several competitive and state-of-the-art methods. Including Bag of Contour Fragments (BCF) introduced in [27] for shape classification, and feature extraction based on dual-output pulse-coupled neural network (DPCNN) and bag of words (BOW) is proposed in [28]. We randomly divided the dataset into two portions. The first part is used for training and the second part is held for the testing process. For comparison, we take 53 examples per species for a training set and a testing set consisting of the remaining 22 images.

Fig. 5 shows the confusion matrices obtained utilizing the SVM classifier with a Gaussian Kernel, and it shows that the accuracy of our proposed method is 99.4%. The results in Table I show that our approach consistently outperforms all these existing methods which obtain the highest accuracy, 99.4%.

CONCLUSION

The combination of BOF and LBP produces powerful results, and has been shown to contain complementary information. By combining BOF and LBP, this combination inherits the advantages of the bag of feature which aggregates the local components to form the global histogram characterization. The combination of local and non-local texture features significantly outperforms some of the state-of-the-art algorithms. The BOF is fixed length vector irrespective of the number of

Method	Average accuracy
SC+DP [22]	88.12%
OLDSE [23]	89.16%
Fourier [22]	89.36%
MMD [25]	93.6%
ISDC [26]	94.13%
MEW [24]	96.53%
BCF [27]	96.56%
DPCNN+BOW [28]	97.63%
Our method	**99.4%**

TABLE I
THE AVERAGE ACCURACY COMPARISON WITH PUBLISHED RESULTS FOR SEVERAL COMPETITIVE ALGORITHMS ON SWEDISH DATASET

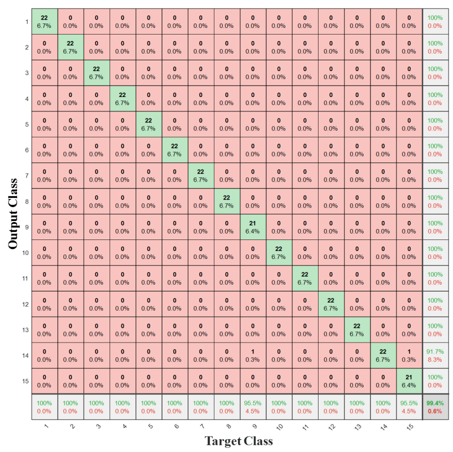

Fig. 5. Shows the Confusion Matrix obtained for testing set of the Swedish Leaf dataset using SVM classifier

detection. Furthermore, our results show that combining BOF and LBP outperforms a much more complex method, such as DPCNN.

REFERENCES

[1] Khairnar, Khushal, and Rahul Dagade. ”Disease Detection and Diagnosis on Plant using Image Processing–A Review.” International Journal of Computer Applications 108, no. 13 (2014): 36-38.
[2] Şekeroğlu, Boran, and Yücel İnan. “Leaves recognition system using a neural network.” Procedia Computer Science 102 (2016): 578-582.
[3] Narayanan, Barath Narayanan, Ouboti Djaneye-Boundjou, and Temesguen M. Kebede. ”Performance analysis of machine learning and pattern recognition algorithms for malware classification.” In Aerospace and Electronics Conference (NAECON) and Ohio Innovation Summit (OIS), 2016 IEEE National, pp. 338-342. IEEE, 2016.
[4] I. Aldaouab, M. Daniels, and K. Hallinan, “Microgrid cost optimization for a mixed-use building,” 2017 IEEE Texas Power and Energy Conference (TPEC), 2017.
[5] Ali, Redha A., and Russell C. Hardie. ”Recursive non-local means filter for video denoising.” EURASIP Journal on Image and Video Processing 2017, no. 1 (2017): 29.
[6] Almahdi, Redha, and Russell C. Hardie. ”Recursive non-local means filter for video denoising with Poisson-Gaussian noise.” In Aerospace and Electronics Conference (NAECON) and Ohio Innovation Summit (OIS), 2016 IEEE National, pp. 318-322. IEEE, 2016.
[7] Sakai, N., S. Yonekawa, A. Matsuzaki, and H. Morishima. ”Two-dimensional image analysis of the shape of rice and its application to separating varieties.” Journal of Food Engineering 27, no. 4 (1996): 397-407.
[8] Neto, João Camargo, George E. Meyer, David D. Jones, and Ashok K. Samal. ”Plant species identification using Elliptic Fourier leaf shape analysis.” Computers and electronics in agriculture 50, no. 2 (2006): 121-134.
[9] Zhang, Shanwen, Yingke Lei, Tianbao Dong, and Xiao-Ping Zhang. ”Label propagation based supervised locality projection analysis for plant leaf classification.” Pattern Recognition 46, no. 7 (2013): 1891-1897.
[10] Nowak, Eric, Frédéric Jurie, and Bill Triggs. ”Sampling strategies for bag-of-features image classification.” In European conference on computer vision, pp. 490-503. Springer, Berlin, Heidelberg, 2006.
[11] Lazebnik, Svetlana, Cordelia Schmid, and Jean Ponce. ”Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories.” In null, pp. 2169-2178. IEEE, 2006.
[12] Wang, Jinjun, Jianchao Yang, Kai Yu, Fengjun Lv, Thomas Huang, and Yihong Gong. ”Locality-constrained linear coding for image classification.” In Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, pp. 3360-3367. IEEE, 2010.
[13] Harris, Chris, and Mike Stephens. ”A combined corner and edge detector.” In Alvey vision conference, vol. 15, no. 50, pp. 10-5244. 1988.
[14] Bay, Herbert, Tinne Tuytelaars, and Luc Van Gool. ”Surf: Speeded up robust features.” In European conference on computer vision, pp. 404-417. Springer, Berlin, Heidelberg, 2006.
[15] Kadir, Timor, and Michael Brady. ”Saliency, scale and image description.” International Journal of Computer Vision 45, no. 2 (2001): 83-105.
[16] Mikolajczyk, Krystian, Tinne Tuytelaars, Cordelia Schmid, Andrew Zisserman, Jiri Matas, Frederik Schaffalitzky, Timor Kadir, and Luc Van Gool. ”A comparison of affine region detectors.” International journal of computer vision 65, no. 1-2 (2005): 43-72.
[17] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, 2004.
[18] Pedersen, Jacob Toft. ”Study group SURF: Feature detection and description.” Department of Computer Science, Aarhus University (2011).
[19] Bosch, Anna, Xavier Muñoz, and Robert Martí. ”Which is the best way to organize/classify images by content?.” Image and vision computing 25, no. 6 (2007): 778-791.
[20] Ojala, Timo, Matti Pietikäinen, and David Harwood. ”A comparative study of texture measures with classification based on featured distributions.” Pattern recognition 29, no. 1 (1996): 51-59.
[21] Oskar J. O. Söderkvist, “Computer vision classifcation of leaves from swedish trees,” Master’s Thesis, Linkoping University, 2001.
[22] Ling HB, Jacobs DW(2007) Shape classification using the inner-distance. IEEE Trans Pattern AnalMach Intell 29(2):286–299
[23] Lei YK, Zou JW, Dong TB et al (2014) Orthogonal locally discriminant spline embedding for plant leaf recognition. Comput Vis Image Underst 119:116–126
[24] Kadir A, Nugroho LE, Susanto A et al (2012) Experiments of zernike moments for leaf identification. J Theor Appl Inform Technol 41(1):82–93
[25] Hu RX, JiaW, Ling HB et al (2012) Multiscale distance matrix for fast plant leaf recognition. IEEE Trans Image Process 21(11):4667–4672
[26] Ling, Haibin, and David W. Jacobs. ”Shape classification using the inner-distance.” IEEE transactions on pattern analysis and machine intelligence 29, no. 2 (2007): 286-299.
[27] Wang ZB, Wang S, Zhu Y et al (2016) Review of image fusion based on pulse-coupled neural network. Arch Comput Methods Eng 23(4):659–671
[28] Wang, Zhaobin, et al. ”Leaf Recognition Based on DPCNN and BOW.” Neural Processing Letters 47.1 (2018): 99-115.
[29] I. Aldaouab, M. Daniels, “Renewable Energy Dispatch Control Algorithms for a Mixed-use Building,” 2017 IEEE Green Energy and Smart Systems Conference IGESSC 2017.
[30] Aldaouab, Ibrahim, and Malcolm Daniels. ”Microgrid battery and thermal storage for improved renewable penetration and curtailment.” Energy and Sustainability Conference (IESC), 2017 International. IEEE, 2017.
[31] I. Aldaouab, M. Daniels, R. Ordóñez “Model Predictive Control Energy Dispatch to Optimize Renewable Penetration for a Microgrid with Battery and Thermal Storage,” 2018 IEEE Texas Power and Energy Conference (TPEC), 2018.