Leaf Recognition Based on Elliptical Half Gabor and Maximum Gap Local Line Direction Pattern

XUAN WANG, WEIKANG DU, FANGXIA GUO, AND SIMIN HU
School of Physics and Information Technology, Shaanxi Normal University, Xi’an 710062, China

Corresponding author: Xuan Wang (wxuan@snnu.edu.cn)

This work was supported in part by the Major Program of National Natural Science Foundation of China under Grant 60633020, in part by the National Natural Science Foundation of China under Grant 61703256, and in part by the Natural Science Basic Research Plan in Shaanxi Province of China under Grant 2017JQ6070.

ABSTRACT Plant identification via leaf images is very meaningful to agricultural information. The existing methods were based on one or two kinds of the three distinct characteristics in leaf images including leaf contours, textures and veins. This limits their recognition performance and scope of application. This paper describes a novel counting-based leaf recognition method, which can directly and effectively combine all of the three kinds of significant characteristics in leaf images. In order to obtain the stable and independent local line responses from leaf contour, texture and vein, elliptical half Gabor is introduced and convoluted with the raw grayscale leaf images, and then maximum gap local line direction patterns are extracted from the local line responses and normalized in direction by cyclically right shifting these patterns until the most numerous bit plane with a value of 1 to the left bit. The histogram of the normalized patterns is calculated and regarded as the counting-based local structure descriptor, and support vector machine is utilized as the classifier. Experimental results on three frequently used leaf databases show that the proposed approach yields a better performance in terms of the classification accuracy, applicability and feasibility in comparison with the state of the art methods.

INDEX TERMS Counting-based descriptor, elliptical half-Gabor filters, maximum gap local line direction pattern, plant identification, support vector machine.

I. INTRODUCTION

Automatic plant identification systems are very meaningful to agricultural information and ecological protection. The biological or phytochemical property-based techniques such as morphological anatomy, molecular biology and phytochemistry require complex processing, so they are not suitable for online applications [1], whereas, the plant recognition based on image analysis can extract plant features directly from living plants, and is suitable for online applications. The images from flowers, fruits, roots and leaves can be used for plant recognition, among them, leaf images are the most feasible ones, so plant recognition based on leaf images has attracted more attentions [1].

Popular leaf image recognition works pay attention to shape features [2]–[7]. In these methods, leaf contours are firstly determined by a pre-processing course, and then the curvature scale space [2], [3], inner distance shape context [4]–[6] or multiscale distance matrix [7] approaches are utilized to extract global shape features invariant to position, scale and direction variations. Kumar et al. [2] design a mobile application for plant identification. They extract curvature features from the pre-processed leaf images, and use a nearest neighbor classifier with histogram intersection as the distance metric for classification. Ling and Jacobs [4], [5] propose a shape classification method called inner distance shape context. They sample points along the boundary of a shape, and build a 2D histogram descriptor at each point. This histogram represents the distance and angle from each point to all other points, along a path restricted to lie entirely inside the leaf shape. Bellumeur et al. [6] present an automatic plant identification system using inner distance shape context and a nearest neighbor classifier. Hu et al. [7] construct a leaf image recognition method with the multiscale distance matrix and the nearest neighbor rule with Euclidean distance. These methods yield a good identification performance for the plants with significantly different leaf contours, but they are generally sensitive to the quality of the pre-processing.
results. In practice, there are many different plant species with similar overall contours, and the same kind of plant species possesses leaves with different overall shapes. Hence, their discriminability is not strong enough for all plant recognition applications.

Besides leaf contours, texture and vein are the other two kinds of the most distinct and significant characteristics in leaf images. The contour-based methods extract features only from the boundaries of leaves, and neglect other useful characteristics in the leaf images. In order to alleviate the drawbacks as mentioned above and make them less sensitive to the quality of the pre-processing results, a lot of works extend text analysis techniques to leaf classification. Casanova et al. [8] calculate the energies of the responses for Gabor filters as texture features. Liu et al. [9] propose a leaf classification method using wavelet transforms and support vector machine. In our early work [10], we combine dual-scale decomposition with local binary descriptors for plant leaf recognition. Naresh and Nagendraswamy [11] extract the local texture structure using a modified Local Binary Patterns (LBP) approach. Tang et al. [12] combine Gray Level Co-Occurrence Matrix (GLCM) with LBP for tea leaf classification. Meanwhile, some works focus on the vein characteristics. Fu and Chi [13] combine a thresholding approach with neural network for extracting vein patterns from leaf images. Park et al. [14] classify vein structures with the pattern of end points and branch points. Larese et al. [15] extract vein patterns using hit or miss transform from legume leaves. Compared with the contour-based methods, these texture or vein based methods are less sensitive to the quality of the pre-processing results, and achieve reasonable results in the case of different plant species with similar overall contours or the same kind of plant leaves with different overall shapes, but their classification accuracies are generally lower than the state of the art contour-based methods for the universal plant recognition. Nevertheless, these works adequately demonstrate that texture and vein patterns are very helpful to leaf identification.

Deep learning techniques are also used for leaf identification [16], [17]. Grinblat et al. [18] train a convolutional neural network with vein morphological patterns for leaf identification. Lee et al. [19] design a multiscale fusion convolutional neural network to fuse the features extracted from leaf images with different scales. Hu et al. [20] learn useful leaf features directly from the raw leaf image data using a convolutional neural network, and quantify the learned features based on a deconvolutional network for species identification. They also analyze and justify the subset of features that are most important to describe leaf data via feature visualization techniques, and find that venation structure is a very important feature for identification especially when shape feature alone is inadequate. These methods have reported a promising recognition performance. Moreover, some of them start directly from the raw leaf images and free of the pre-processing course. However, it is well known that deep convolutional neural networks require very large amounts of training data, the number of samples in existing leaf image databases is still far from matching the scale and variety of existing general major databases for images, videos or languages. So, to efficiently train a deep architecture to recognize plant leaf images, much larger datasets are required, preferably with more than a million images and higher category variability [18].

Recently, Zhao et al. [21] find that it is better to “count” the number of certain shape patterns rather than to match the extracted global shape features in a point-wise manner. Based on this idea, they propose a counting-based shape descriptor for identifying plant species. Similarly, Cerutti et al. [22] utilize a sequence-like structured representation to track the spatial information from curvature space and present a smartphone application for identifying plant species. Zhang et al. [23] apply Warshall algorithm [24] to label propagation of leaf contours to obtain a label matrix, and then the matrix is dealt with discriminant neighbors to form an optimum projecting to low dimensionality, the resultant descriptors are classified by the nearest neighbor classifier. Benefitting from the idea of counting, these methods generally outperform over the conventional contour-based methods in terms of classification accuracy and applicability, but they extract counting-based shape features only from the boundaries of leaves, and neglect other useful characteristics such as veins and textures in the leaf images. Furthermore, their sensitivity to the quality of the pre-processing results is in line with the conventional contour-based methods.

In order to take advantage of the three kinds of significant characteristics in leaf images and make leaf recognition free of the pre-processing course, motivated by aforementioned inspirational conclusions about leaf identification including importance of venation structure [18] and advantage of counting the number of local shape patterns over matching global shape features point by point [21]–[24], we propose a novel counting-based local structure descriptor for identifying plant species. In order to fairly combine the three most significant characteristics including shape, venation and main texture in leaf images, we design a new kind of Gabor wavelet namely elliptical half Gabor wavelet to highlight the local dominant orientation information of leaf contours, veins and main textures, and then we present an improved local line direction coding approach named as Maximum Gap Local Line Direction Pattern (MGLLDP) to extract local dominant orientation structure patterns from the elliptical half Gabor wavelet domain. The histogram of the normalized local dominant orientation structure patterns is regarded as the counting-based local structure descriptor, and support vector machine is utilized as the classifier.

The remainder of this paper is organized as follows. Section II provides a brief review of local line direction pattern and support vector machine. The proposed counting-based local structure descriptor is described in Section III, and the experimental results are presented in Section IV. In Section V, the concluding remarks are provided.
II. LOCAL LINE DIRECTION PATTERN AND SUPPORT VECTOR MACHINE

A. LOCAL LINE DIRECTION PATTERNS

As a kind of local image descriptors, Local Binary Pattern (LBP) [25] has been successfully applied to many computer vision applications. Currently, the trend of LBP is to encode the edge gradient information rather than intensity information such as Local Directional Pattern (LDP) [26], Enhanced Local Directional Pattern (ELDP) [27], Gradient Directional Pattern (GDP) [28] and Local Directional Number (LDN) pattern [29]. Since the edge gradient is more stable than the pixel intensity, these local descriptors yield a better recognition performance than original LBP. Recently, Luo et al. [30] proposed a new kind of local descriptors named as Local Line Directional Pattern (LLDP) for palmprint recognition, in which, instead of edge gradients, the local line directional information was encoded. Let \( \{ m_i \} \) denote the line response values obtained with an appropriate line information analysis technology, its LLDP code is derived by [20]

\[
LLDP = \sum_{i=0}^{L} b_i (m_i - m_k) 2^i,
\]

where \( m_k \) is the \( k \)-th minimum directional response. The reason for \( b_i(x) = 1 \) as \( x < 0 \) is that palmprints are dark lines in palmprint images. From the experimental results in [20], the LLDP outperforms over the LBP-like codes based on edge gradients.

B. SUPPORT VECTOR MACHINE

Support vector machine (SVM) is a frequently-used classifier developed by Vapnik [31] based on statistical learning theory. The fundamental principle of SVM can be demonstrated as follows: For two classes of linearly separable problems, as shown in Fig. 1, \( H \) represents the optimal separating hyperplane to be determined, \( H_1 \) and \( H_2 \) are two hyperplanes parallel to \( H \) and no training sample falls between them. According to statistical learning theory, \( H \) should be the separating hyperplane which can separate the two categories with the maximum margin so that structural risk can be minimized. Consequently, the problem of solving \( H \) is equivalent to a constrained optimization problem. Given a training data set of \( N \) points \( \{ x_i, y_i \} \), the task is therefore equal to

\[
\max \left( \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i,j} \lambda_i \lambda_j y_i y_j x_i^T x_j \right)
\]

under the constraints \( \sum_{i=1}^{N} \lambda_i y_i = 0 \) and \( \lambda_i \geq 0 , i = 1, 2, \cdots, N \), where \( \lambda_i \) and \( \lambda_j \) are Lagrange multiplying factors. For linearly non-separable problems, the corresponding form is given by

\[
\max \left( \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i,j} \lambda_i \lambda_j y_i y_j K(x_i, x_j) \right)
\]

where \( K(x_i, x_j) \) is called kernel function selected from the typical functions expressed as follows:

Polynomial kernel of degree \( q \)

\[
k(x_i, x_j) = (x_i^T x_j + 1)^q, \quad q > 0 \tag{4}
\]

Radial basis function kernel:

\[
k(x_i, x_j) = \exp \left( \frac{\|x_i - x_j\|^2}{\sigma^2} \right) \tag{5}
\]

Hyperboloidal tangent kernel:

\[
k(x_i, x_j) = \tanh(\beta x_i^T x_j + \gamma) \tag{6}
\]

In this work, we select the radial basis function with \( \sigma = 6.28 \) as the kernel function for support vector machine.

III. THE PROPOSED METHOD

The block diagram of our method is shown in Fig. 2. The input raw color leaf image is converted to a gray level image \( f(i, j) \), and then it is convoluted with the elliptical half Gabor wavelets to extract the line responses for 12 orientations located in \( (i, j) \). Afterwards, maximum gap local line direction patterns are extracted from these line responses, and normalized in direction by cyclically right shifting these patterns until the most numerous bit plane with a value of 1 to the left bit. The histogram of the normalized patterns is calculated and regarded as the counting-based local structure descriptor, and support vector machine is utilized as the classifier.

A. ELLIPTICAL HALF GABOR WAVELET

Gabor wavelet is one of the most effective tools for texture and orientation analysis due to its useful properties including accurate time-frequency localization, robustness against varying brightness and contrast of images, etc. [32]. The real part of classical 2D-Gabor wavelet is expressed as [33]

\[
G(x, y, \theta, \mu, \sigma) = \frac{1}{2 \pi \sigma^2} \exp \left( -\frac{x^2 + y^2}{2\sigma^2} \right) \cos(2\pi \mu(x \cos \theta + y \sin \theta)) \tag{7}
\]

where \( \mu \) is the radial frequency per unit length, \( \theta \) denotes the orientation of the Gabor function, and \( \sigma \) represents the standard deviation of Gaussian envelope. Apparently, as shown in Fig. 3, the classical 2D-Gabor wavelet has two drawbacks in line response analysis. One is the usage of symmetric Gaussian envelope forming isotropic weight parameters, which results in unwanted distortions to line responses from the pixels far away from direction \( \theta \). The other is that the response
of classical 2D-Gabor wavelet for a certain orientation $\theta$ contains line responses in two directions $\theta$ and $\theta + \pi$, that is to say, it is difficult to obtain the two directional line responses independently using the classical 2D-Gabor wavelet.

In order to overcome the second drawback, Fei et al. [33] proposed a modified 2D-Gabor wavelet for palmprint recognition namely half Gabor wavelet. However, their work contains a crucial mistake in the definition of the half-Gabor filters. In our early work, we present the correct version, which was defined as [34]

$$
r^r_G(x, y, \theta, \mu, \sigma) = \begin{cases} 
G(x, y, \theta, \mu, \sigma) & \text{if } (x \cos \theta + y \sin \theta) \geq -T \\
0 & \text{else}
\end{cases}$$

(8)

and

$$
s^s_G(x, y, \theta, \mu, \sigma) = \begin{cases} 
G(x, y, \theta, \mu, \sigma) & \text{if } (x \cos \theta + y \sin \theta) \leq T \\
0 & \text{else}
\end{cases}$$

(9)

where threshold $T$ is a nonnegative number, in their work, $T$ was set as 2. Obviously, they split the classical 2D-Gabor wavelet into two Gabor filters along the direction perpendicular to $\theta$ with an overlap region of $2T$ width, providing Gabor filters $r^r_G(x, y, \theta, \mu, \sigma)$ for orientation $\theta$ and $s^s_G(x, y, \theta, \mu, \sigma)$ for orientation $\theta + \pi$.

In order to overcome aforementioned two drawbacks synchronously, combining the idea of half Gabor wavelet, we provide an improved 2D-Gabor wavelet namely elliptical half Gabor wavelet, which is defined as

$$
EH^r_G(x, y, \theta, \mu, \sigma, \sigma_l, \sigma_s) = \begin{cases} 
EG(x, y, \theta, \mu, \sigma, \sigma_l, \sigma_s) & \text{if } (x \cos \theta + y \sin \theta) \geq 0 \\
0 & \text{else}
\end{cases}
$$

(10)

and

$$
EH^s_G(x, y, \theta, \mu, \sigma, \sigma_l, \sigma_s) = \begin{cases} 
EG(x, y, \theta, \mu, \sigma, \sigma_l, \sigma_s) & \text{if } (x \cos \theta + y \sin \theta) \leq 0 \\
0 & \text{else}
\end{cases}
$$

(11)

where $EH^r_G(x, y, \theta, \mu, \sigma, \sigma_l, \sigma_s)$ denotes the real part of 2D-Gabor wavelet with an 2D-elliptical Gaussian envelope, which is given by

$$
EG(x, y, \theta, \mu, \sigma, \sigma_l, \sigma_s) = \frac{1}{2\pi \sigma_l \sigma_s} \exp\left(-\frac{(x \cos \theta + y \sin \theta)^2}{2\sigma_l^2} - \frac{(x \cos \theta + y \sin \theta)^2}{2\sigma_s^2}\right) \times \cos(2\pi \mu(x \cos \theta + y \sin \theta))
$$

(12)

where $\sigma_l$ denotes the major axis of the elliptical envelope, and $\sigma_s$ represents its minor axis. Note that the long axis of the elliptical envelope along orientation $\theta$, and the short axis perpendicular to $\theta$. As shown in Fig. 3 and 4, in elliptical half Gabor wavelet, the Gabor filter at $\theta$ direction is split into two half Gabor filters along the direction perpendicular to $\theta$, which can provide two directional line responses for $\theta$ and $\theta + \pi$ independently. Furthermore, compared with the Gabor filter for classical 2D-Gabor wavelet, the elliptical half Gabor filters are more compact in the direction perpendicular to $\theta$, and enlarged along $\theta$ orientation. This means that our elliptical half Gabor wavelet can effectively reduce the
impact on line responses from the pixels deviated from the $\theta$ orientation and enhance the contributions from the concerned pixels along $\theta$.

B. MAXIMUM GAP LOCAL LINE DIRECTION PATTERNS

It has been proved empirically that LLDP defined in Eq.1 outperforms over the LBP-like codes based on edge gradients or pixel intensities [20], but there are still some weaknesses in the original definition of LLDP. The first shortcoming comes from the determination of $m_k$ which was simply set to the third minimum directional response in reference [20] i.e. only the first two dominant orientations were encoded as 1. Since in a real scene image, the number of dominant orientations per pixel is different, for example, as shown in Fig. 5(a), the number of principal directions of the point a, b, c, d is 3,4,0,2, respectively, it is not proper only considering first two dominant orientations for all pixels. This will cause to omit part of principal direction information for pixels with a number of principal directions greater than 2 such as point a, and bring unwanted clutters for pixels with a number of principal directions less than 2 such as point c. The second shortcoming is that in the light of the original definition of LLDP, a non-zero LLDP code will be generated for each pixel even if the pixel is located in smooth region like pixel c. This is unreasonable, and the non-zero LLDPs from pixels located in the smooth region will disturb and degrade overall discriminability of LLDP.

We address the two shortcomings of the original LLDP, and propose an improved LLDP namely Maximum Gap Local Line Direction Pattern (MGLLDP). Given a line response set $\{m_i\} (i = 0, 1, \ldots, L)$, MGLLDP is defined as

$$MGLLDP = \begin{cases} \frac{1}{2} \sum_{i=0}^{L} b(m_i - m_g)2^i & G_x \geq T, \\ 0 & D_g < T \end{cases}$$

where $T$ is a threshold, $D_g$ denotes the maximum gap in $\{m_i\} (i = 0, 1, \ldots, L)$, and $m_g$ represents the superior of the two line responses associated with the maximum gap $D_g$ which can be obtained as follows

1.) Sort $\{m_i\}$, ($i = 0, 1, \ldots, L$) in ascending order.
2.) Calculate the differences between adjacent values of the ascending sequence, a difference sequence $\{D_i\}$, ($i = 0, 1, \ldots, L - 1$) is obtained.
3.) The maximum gap $D_g$ is determined by calculating the maximum of the difference sequence $\{D_i\}$, ($i = 0, 1, \ldots, L - 1$).
4.) The larger one of the two line responses associated with the maximum gap is regarded as $m_g$.

The advantages of MGLLDP over the original LLDP are

1.) the number of principal directions for each pixel center can be adaptively determined, as a result, all principal directions can be properly encoded, no principal direction is neglected and no unwanted clutter is pulled in.
2.) for the pixels located in smooth region, which maximum gaps are less than the threshold $T$, their MGLLDPs equal zero. This can alleviate the distortions from the pixels without principal direction.
3.) the number of principal directions for each pixel center can be adaptively determined, as a result, all principal directions can be properly encoded, no principal direction is neglected and no unwanted clutter is pulled in.
4.) the larger one of the two line responses associated with the maximum gap is regarded as $m_g$.

The coding images for leaf image shown in Fig. 5(a) using the combination of the original Gabor with LLDP and our method based on the elliptical half Gabor and MGLLDP with different threshold are shown in Fig. 5(b), (c), (d), (e) and (f). From these figures, it can be seen that coding image Fig. 5(c) is more clear and distinct than coding image Fig. 5(b). This maybe due to that the elliptical half Gabor can effectively reduce the unwanted distortions from the pixels far away from the direction to be analyzed and the number of principal directions for each pixel center can be adaptively determined, and so no clutter is pulled in during the MGLLDP coding phase. Furthermore, by varying the threshold $T$, one can balance extracting the weak principal direction information with filtering the distortions.

The advantages of MGLLDP over the original LLDP are

1.) the number of principal directions for each pixel center can be adaptively determined, as a result, all principal directions can be properly encoded, no principal direction is neglected and no unwanted clutter is pulled in.
2.) the number of principal directions for each pixel center can be adaptively determined, as a result, all principal directions can be properly encoded, no principal direction is neglected and no unwanted clutter is pulled in.
3.) the number of principal directions for each pixel center can be adaptively determined, as a result, all principal directions can be properly encoded, no principal direction is neglected and no unwanted clutter is pulled in.
4.) the larger one of the two line responses associated with the maximum gap is regarded as $m_g$.

C. THE NORMALIZATION OF MGLLDP IN ORIENTATION AND SCALE

Obviously, MGLLDP is invariant to images’ position change, but relatively sensitive to the change of orientation and scale. In order to improve the robustness of MGLLDP to orientation changes, we provide a directional normalization method for MGLLDP, which is described as
1.) Count the number of 1 for each bitplane in overall MGLLDPs extracted from an image with size of $N \times M$.
2.) cyclically right shift all MGLLDPs until the most numerous bit plane with a value of 1 to the left bit.

The histogram $H(MGLLDP)$ of the normalized MGLLDPs is calculated, in order to make $H(MGLLDP)$ less sensitive to scale change, the histogram is normalized by

$$H'(MGLLDP) = \frac{H(MGLLDP)}{N \times M}$$

(14)

The normalized histogram $H'(MGLLDP)$ is regarded as the counting-based local structure descriptor.

IV. SIMULATION RESULTS AND PERFORMANCE ANALYSIS

In this section, we will conduct experiments on three frequently used leaf databases: Swedish, Flavia and ICL database, in comparison with half Gabor and LLDP version and eight representative methods including two contour-based methods i.e. Inner Distance Shape Context (IDSC) [4] and Multiscale Distance Matrix (MDM) [7], two texture-based methods namely Gabor based method [8] and Wavelet based method [9], two deep learning-based methods named as deep learning on veins [18] and multiscale fusion convolutional neural network [20], two recent counting-based methods i.e. pattern counting approach [21] and Label propagation projection [23].

A. SELECTION OF THRESHOLD T AND PREPARATION OF LEAF IMAGES

In order to balance extracting the weak principal direction information with filtering the distortions, in this subsection, an appropriate threshold is determined via experiment, in which ICL database is considered. The ICL leaf database downloaded from [35] was provided by the Intelligent Computing Laboratory (ICL) of institute of intelligent machines, Chinese academy of sciences, which contains 17,032 leaf images from 221 plant species. In this experiment, 30 plant species are randomly selected from ICL leaf database, the first ten leaf images for each plant species are regarded as training samples and the remainders are used for testing. The threshold T changes from 10 to 90 with 10 increments, the corresponding classification accuracies are determined and summarized in Fig. 6. Based on this figure, threshold T is set to 40 in subsequent experiments.

Since the normalization of MGLLDP in orientation and scale is provided in our method, the method directly works on the corresponding grey-level leaf images, while for the compared methods, the plant leaf images were prepared strictly in accordance with their original references. Notes that some compared methods [8], [9], [21], [23] require an orientation and scale normalization operation in their preparation courses. In addition, the two shape-based methods [4], [7] directly started with the contour of leaf, there is no details about how to obtain the contour in their original papers, and so, for the two methods, the contour was obtained by using the contour extraction approach in [21]. Considering that limited training samples do not effectively highlight the advantages of deep learning, each image in training sets is further augmented to 20 images by the horizontal mirror, rotation (from 0 to 180 with 10 increments) and illumination changes simulated by Gamma correction with Gamma varying from 1.2 to 2.4 with 0.2 increments. All images are scaled to 256 $\times$ 256 pixels for multiscale fusion convolutional neural network. For deep learning on vein method, the hit or miss transform in [15] is used to extract vein morphological patterns, and then a central patch (100 $\times$ 100 pixels) of the resultant image is cropped and the rest of the image is discarded.

B. PERFORMANCE

1) ON SWEDISH DATABASE

The Swedish leaf database contains 1125 leaf images from 15 different Swedish trees with 75 images per species. We randomly select 25 leaf images per species for training and the rest 50 images were used for test. The classification accuracies for the aforementioned methods are listed in Table 1. As can be seen, the classification accuracies of texture-based methods are generally lower than the contour-based methods. Among the contour-based methods, the counting-based methods yield a better performance than the methods matching the extracted contour features in a point-wise manner, and the lower performance of deep learning on vein method may due to that they only consider a central patch (100 $\times$ 100 pixels) of leaf images. Benefiting from the advantages of elliptical half Gabor and MGLLDP, our method yields a 5% increment on the classification accuracy of Half Gabor and LLDP method and outperforms the other eight methods.

2) ON FLAVIA DATABASE

Flavia database downloaded from [36] contains 1907 leaf images of blades without petioles from 32 different species. The first 25 leaf images per species are used as the training set and the other as the test set. The classification accuracies for the aforementioned methods are determined and listed using the contour extraction approach in [21]. Considering that limited training samples do not effectively highlight the advantages of deep learning, each image in training sets is further augmented to 20 images by the horizontal mirror, rotation (from 0 to 180 with 10 increments) and illumination changes simulated by Gamma correction with Gamma varying from 1.2 to 2.4 with 0.2 increments. All images are scaled to 256 $\times$ 256 pixels for multiscale fusion convolutional neural network. For deep learning on vein method, the hit or miss transform in [15] is used to extract vein morphological patterns, and then a central patch (100 $\times$ 100 pixels) of the resultant image is cropped and the rest of the image is discarded.
in Table 2. As can be observed, our method also performs the best among the nine methods, compared with the results on Swedish leaf database, the performance of the texture-based methods increases slightly, while the classification accuracy of the contour-based methods declines within a small range, which indicates that petioles indeed provide useful information for contour-based methods [7], and unwanted distortions for the texture-based methods.

3) ON ICL DATABASE

The ICL leaf database contains 17,032 leaf images from 221 plant species. Three datasets are obtained by carefully selecting the samples with specific characteristics from ICL database. Dataset A consists of all samples of ICL database. The species in dataset B are carefully selected and most of them without vein and texture or with weak veins and textures, part of them are shown in Fig. 7, which contains 3390 leaf images from 44 plant species. Dataset C consists of 2312 leaf images from 30 species, as shown in Fig. 8, most of them with similar overall contours and different veins or textures. For the three datasets, the first 15 leaf images per species are used as the training set and the other as the test set, the classification results are summarized in Table 3.

From Table 3, we have several observations

1.) the classification accuracies of these methods on ICL leaf database are generally lower than the results on the two previous databases. For the deep learning methods, this may be due to that the number of training samples for each species is reduced from 25 to 15. As to the other methods, the reason could be that the samples in the ICL leaf database were captured at more complicated conditions and with more diversified characteristics. Since our method contains the normalized processes for orientation and scale changes, and is invariant to images’ position change, moreover, the elliptical half Gabor wavelet inherits the useful properties of Gabor wavelet including accurate time-frequency localization, robustness against varying brightness and contrast of images, which can effectively enhance the contributions from the concerned pixels, our method can better adapt to various complicated conditions and performs the best among the seven methods.

2.) For the leaf images with weak veins and textures i.e. dataset B, the performance of texture-based methods and deep learning methods has been significantly

| Methods                        | Classification accuracy (%) |
|--------------------------------|-----------------------------|
| Inner Distance Shape Context   | 94.13                       |
| Multiscale Distance Matrix     | 93.60                       |
| Gabor based method             | 85.73                       |
| Wavelet based method           | 91.33                       |
| Pattern counting approach      | 97.07                       |
| Label propagation projection   | 96.27                       |
| Half Gabor and LLDP            | 93.26                       |
| Deep learning on veins         | 94.45                       |
| Convolutional neural network   | 97.83                       |

| Methods                        | Classification accuracy (%) |
|--------------------------------|-----------------------------|
| Inner Distance Shape Context   | 93.22                       |
| Multiscale Distance Matrix     | 90.33                       |
| Gabor based method             | 86.36                       |
| Wavelet based method           | 92.23                       |
| Pattern counting approach      | 95.30                       |
| Label propagation projection   | 94.76                       |
| Half Gabor and LLDP            | 92.37                       |
| Deep learning on veins         | 93.26                       |
| Convolutional neural network   | 95.45                       |
| Our method                     | 97.83                       |

| Methods                        | Classification accuracy (%) |
|--------------------------------|-----------------------------|
| Inner Distance Shape           | 90.75 91.83 73.20          |
| Multiscale Distance            | 88.24 89.41 70.40          |
| Gabor based method             | 84.60 64.32 89.31          |
| Wavelet based method           | 89.45 65.20 93.45          |
| Pattern counting               | 93.42 94.21 87.16          |
| Label propagation              | 91.34 92.42 86.31          |
| Half Gabor and LLDP            | 90.23 89.35 92.68          |
| Deep learning on veins         | 90.46 88.56 92.24          |
| Convolutional neural network   | 93.38 92.57 94.45          |
| Our method                     | 96.76 95.27 97.37          |
redced, while that of contour-based methods has improved slightly. Our method still performs the best among the nine methods, but there have been a slight decline. This indicates that our counting-based local structure descriptor is able to represent contours very well, and lack of texture and vein information has negative impact on its classification capacity.

3.) On dataset C, in which the leaf images are with similar overall contours and different veins or textures, the performance of texture-based methods has increased dramatically, while that of contour-based methods has dropped in different degrees, whereas our method outperforms the other nine methods with a slight rise. This may be due to that our counting-based local structure descriptors combine the contours, veins and textures efficiently, and parts of them from the similar contours have dragged down the rise.

In brief, from the aforementioned experimental results, it can be concluded that our method outperforms the eight representative methods, and can better adapt to various complicated conditions and application situation. It is worth noting that unlike the compared methods [8], [9], [21], [23], which require a preprocessing course and an orientation and scale normalization operation in the preparation course, our method directly starts with the raw grayscale leaf images, so it has wide applicability and good feasibility.

V. CONCLUSION

We have proposed a novel counting-based leaf recognition method based on the elliptical half Gabor wavelet and maximum gap local line direction patterns. The advantages of our methods over the state of the art are (1) direct and effective combination of all three kinds of significant characteristics in leaf images; (2) high adaptability for various complicated conditions and diversified characteristics; (3) high feasibility due to working directly on raw grayscale leaf images without the need for a preprocessing process. The experimental results on three frequently-used benchmark databases demonstrate the advantages of our method over the representative leaf recognition methods.

In future work, we will consider the automatic identification of plant diseases and insect pests using leaves. Another interesting topic would be to solve the problem of how to identify plants via multiple or overlap leaves, and the possible extension applications of elliptical half Gabor and maximum gap local line direction pattern would be included.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers and the editor for their constructive comments to further improve the quality of this paper.

REFERENCES

[1] J. S. Cope, D. Corney, J. Y. Clark, P. Remagnino, and P. Wilkin, “Plant species identification using digital morphometrics: A review,” Expert Syst. Appl., vol. 39, no. 8, pp. 7562–7573, Jan. 2012.

[2] N. Kumar, P. N. Belhumeur, A. Biswas, D. W. Jacobs, W. J. Kress, I. C. Lopez, and J. V. Soares, “Leafsnap: A computer vision system for automatic plant species identification,” in Proc. Euro. Conf. Comput. Vis. Florence, Italy: Springer, 2012, pp. 502–516.

[3] Z. Wang, Z. Chi, and D. Feng, “Shape based leaf image retrieval,” IEEE Proc.-Vis., Image. Signal Process., vol. 150, no. 1, pp. 34–43, 2003.

[4] H. Ling and D. W. Jacobs, “Shape classification using the inner-distance,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 2, pp. 286–299, Feb. 2007.

[5] H. Ling and D. W. Jacobs, “Shape matching for foliage database retrieval,” in Semantic Mining Technologies for Multimedia Databases. Philadelphia, PA, USA: IGI Global, 2009, pp. 100–129.

[6] P. N. Belhumeur, D. Chen, S. Feiner, D. W. Jacobs, W. J. Kress, H. Ling, I. Lopez, R. Ramamonotha, S. Shoorey, and S. White, “Searching the world’s herbaria: A system for visual identification of plant species,” in Proc. Eur. Conf. Comput. Vis. Marseille, France: Springer, 2008, pp. 116–129.

[7] R. Hu, W. Jia, H. Ling, and D. Huang, “Multiscale distance matrix for fast plant leaf recognition,” IEEE Trans. Image Process., vol. 21, no. 11, pp. 4667–4672, Nov. 2012.

[8] D. Casanova, J. J. de Mesquita Sà Junior, and O. M. Bruno, “Plant leaf identification using Gabor wavelets,” Int. J. Imag. Syst. Technol., vol. 19, no. 3, pp. 236–243, Sep. 2009.

[9] J. Liu, S. Zhang, and S. Deng, “A method of plant classification based on wavelet transforms and support vector machines,” in Proc. Int. Conf. Intell. Comput. Shanghai, China: Springer, 2009, pp. 253–260.

[10] X. Wang, J. Liang, and F. Guo, “Feature extraction algorithm based on dual-scale decomposition and local binary descriptors for plant leaf recognition,” Digit. Signal Process., vol. 34, pp. 101–107, Nov. 2014.

[11] Y. G. Naresh and H. S. NagendraSwamy, “Classification of medicinal plants: An approach using modified LBP with symbolic representation,” Neurocomputing, vol. 173, pp. 1789–1797, Jan. 2016.

[12] Z. Tang, Y. Su, M. J. Er, F. Qi, L. Zhang, and J. Zhou, “A local binary pattern based texture descriptors for classification of tea leaves,” Neurocomputing, vol. 168, pp. 1011–1023, Nov. 2015.

[13] H. Fu and Z. Chu, “Combined thresholding and neural network approach for vein pattern extraction from leaf images,” IEEE Proc.-Vis., Image Signal Process., vol. 153, no. 6, pp. 881–892, 2006.

[14] J. Park, E. Hwang, and Y. Nam, “Utilizing venuation features for efficient leaf image retrieval,” J. Syst. Softw., vol. 81, no. 1, pp. 71–82, Jan. 2008.

[15] M. G. LaRese, R. Namias, R. M. Cravioiti, R. M. Arango, C. Gallo, and P. M. Granitto, “Automatic classification of legumes using leaf vein image features,” Pattern Recognit., vol. 47, no. 1, pp. 158–168, Jan. 2014.

[16] Z. Liu, L. Zhu, X.-P. Zhang, X. Zhou, L. Shang, Z.-K. Huang, and Y. Gan, “Hybrid deep learning for plant leaves classification,” in Proc. Int. Conf. Intell. Comput. Fuzhou, China: Springer, 2015, pp. 115–123.

[17] S. H. Lee, C. S. Chan, P. Wilkin, and P. Remagnino, “Deep-plant: Plant identification with convolutional neural networks,” in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2015, pp. 452–456.

[18] G. L. Gritnblat, L. C. Uzal, M. G. LaRese, and P. M. Granitto, “Deep-learning for plant identification using vein morphological patterns,” Comput. Electron. Agricult., vol. 127, pp. 418–424, Sep. 2016.

[19] S. H. Lee, C. S. Chan, S. J. Mayo, and P. Remagnino, “How deep learning extracts and learns leaf features for plant classification,” Pattern Recognit., vol. 71, pp. 1–13, Nov. 2017.

[20] J. Hu, Z. Chen, M. Yang, R. Zhang, and Y. Cui, “A multiscale fusion convolutional neural network for plant leaf recognition,” IEEE Signal Process. Lett., vol. 25, no. 6, pp. 853–857, Jun. 2018.

[21] C. Zhao, S. S. F. Chan, W.-K. Cham, and L. M. Chu, “Plant identification using leaf shapes—A pattern counting approach,” Pattern Recognit., vol. 48, no. 10, pp. 3203–3215, Oct. 2015.

[22] G. Ceroni, T. Lougen, D. Coquin, and A. Vacavant, “Leaf margins as sequences: A structural approach to leaf identification,” Pattern Recognit. Lett., vol. 49, pp. 177–184, Nov. 2014.

[23] S. Zong, X. Lei, T. Dong, and X.-P. Zhang, “Label propagation based supervised locality projection analysis for plant leaf classification,” Pattern Recognit., vol. 46, no. 7, pp. 1891–1897, Jul. 2013.

[24] W. Hong-Wei, S. Hui-Jie, and C. Xiao-hong, “An improved algorithm WARSHALL,” in Proc. Int. Conf. Comput. Sci. Netw. Technol., Dec. 2011, pp. 560–562.

[25] T. Ojala, M. Pietikainen, and T. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 7, pp. 971–987, Jul. 2002.
[26] T. Jabid, “Robust facial expression recognition based on local directional pattern,” *ETRI J.*, vol. 32, no. 5, pp. 784–794, Oct. 2010.
[27] F. Zhong and J. Zhang, “Face recognition with enhanced local directional patterns,” *Neurocomputing*, vol. 119, pp. 375–384, Nov. 2013.
[28] F. Ahmed, “Gradient directional pattern: A robust feature descriptor for facial expression recognition,” *Electron. Lett.*, vol. 48, no. 19, p. 1203, 2012.
[29] A. Ramirez Rivera, J. Rojas Castillo, and O. Oksam Chae, “Local directional number pattern for face analysis: Face and expression recognition,” *IEEE Trans. Image Process.*, vol. 22, no. 5, pp. 1740–1752, May 2013.
[30] Y.-T. Luo, L.-Y. Zhao, B. Zhang, W. Jia, F. Xue, J.-T. Lu, Y.-H. Zhu, and B.-Q. Xu, “Local line directional pattern for palmprint recognition,” *Pattern Recognit.*, vol. 50, pp. 26–44, Feb. 2016.
[31] V. N. Vapnik, *Constructing Learning Algorithms*. New York, NY, USA: Springer, 1995, pp. 119–166, doi: 10.1007/978-1-4757-2440-0_6.
[32] X. Wang, L. Lei, and M. Wang, “Palmprint verification based on 2D–Gabor wavelet and pulse-coupled neural network,” *Knowl.-Based Syst.*, vol. 27, pp. 451–455, Mar. 2012.
[33] L. Fei, Y. Xu, and D. Zhang, “Half-orientation extraction of palmprint features,” *Pattern Recognit. Lett.*, vol. 69, pp. 35–41, Jan. 2016.
[34] X. Wang, W. Du, and F. Guo, “Comments on ‘The definition of half-Gabor filter’ by Lunke Fei [Pattern Recognit. Lett. 69 (2016) 35–41],” *Pattern Recognit. Lett.*, vol. 128, pp. 244–246, Dec. 2019.
[35] *Multiscale Distance Matrix for Fast Plant Leaf Recognition*. Accessed: Aug. 2005. [Online]. Available: http://www.intelengine.cn/English/dataset/index.html
[36] *Flavia Database (Online)*. Accessed: Dec. 2007. [Online]. Available: http://flavia.sourceforge.net/

**XUAN WANG** was born in 1966. He received the B.S. and M.S. degrees in electrical engineering from Shaanxi Normal University, Xi’an, China, in 1983 and 1987, respectively, and the Ph.D. degree from the Key Laboratory of the Ministry of Education for Computer Networks and Information Security. He is currently a Professor and the Head of the School of Physics and Information Technology, Shaanxi Normal University. His research interests include image processing and pattern recognition.

**WEIKANG DU** was born in 1995. He received the B.S. degree in biomedical engineering from Xidian University, Xi’an, China, in 2017. He is currently pursuing the M.S. degree in signal and information processing with Shaanxi Normal University, Xi’an. His research interests include digital image processing and pattern recognition.

**FANGXIA GUO** was born in 1965. She received the M.S. degree in electrical engineering from Shaanxi Normal University. She is currently a Vice Professor with Shaanxi Normal University. Her research interests include image processing and pattern recognition.

**SIMIN HU** was born in 1997. She received the B.S. degree in communication engineering from Xi’an Shiyou University, Xi’an, China, in 2019. She is currently pursuing the M.S. degree with Shaanxi Normal University, Xi’an. Her research interests include image processing and pattern recognition.

* * *