Machine Vision and Applications manuscript No.
(will be inserted by the editor)

Finding local leaf vein patterns for legume characterization and classification

Mónica G. Larese · Pablo M. Granitto

Received: date / Accepted: date

Abstract In recent years, the importance of analyzing the effect of genetic variations on the plant phenotypes has raised much attention. In this paper, we describe a procedure which can be useful to discover representative leaf vein patterns for each species or variety under analysis. We consider three legumes, namely red bean, white bean and soybean. Soybean specimens are also divided in three cultivars. In total there are five leaf vein image classes. In order to find the discriminative patterns, we detect SIFT (Self-Invariant Feature Transform) keypoints in the segmented vein images. The Bag of Words model is built using SIFT descriptors, and classification is performed resorting to Support Vector Machines with a Gaussian kernel. Classification accuracies outperform recent results available in the literature and manual classification, showing the advantages of the procedure. The Bag of Words model is useful for vein patterns characterization and provides a means to highlight the most representative patterns for each species and variety.

Keywords Plant phenotyping · Leaf vein characterization · Legume species and varieties classification

1 Introduction

In recent years, the importance of analyzing the effect of genetic variations on the plant phenotypes has raised much attention [17]. Moreover, last year an entire special workshop was organized in conjunction with the European Conference on Computer Vision (ECCV 2014) dedicated exclusively to the research on computer vision applied to plant phenotyping [1]. Plant phenotyping is aimed at the characterization and quantitative description of complex plant traits, including, among many others, leaf characteristics.

From a biological perspective, recent studies in the literature show correlations between venation networks and leaf properties (for example, drought and damage tolerance) [22,23]. Under this hypothesis, it is feasible to think that genetic plant adaptation aimed at enduring the species and varieties to the environmental conditions, are reflected in their leaf vein patterns up to some degree. This occurs even when the leaves present similar appearance in shape, color and size.

In the case of some legumes, such as soybean, the experts are interested in determining if there exist differences in the venation of the different cultivars that they handle. They cannot identify them visually, and aim to investigate them using computational and image analysis tools. The process of identification is difficult to be treated manually since humans find small differences among the cultivars, and highly variable characteristics within each one. The process is also subjective and not repeatable, which is why a computational procedure is needed. Classification algorithms offer a means to automatically separate the classes and highlight relevant features if the differences in venation are present.

In the current literature, leaf feature information has been included in automatic plant classification processes. Classification can be performed as long as the features for each class are separable up to some degree. Different approaches include analyzing the shape

M.G. Larese and P.M. Granitto
CIFASIS, French Argentine International Center for Information and Systems Sciences, UAM (France) / UNR-CONICET (Argentina)
Bv. 27 de Febrero 210 Bis, 2000, Rosario, Argentina
E-mail: {larese, granitto}@cifasis-conicet.gov.ar

[1] Computer Vision Problems on Plant Phenotyping (CVPPP), Zurich, September 12, 2014, http://www.plant-phenotyping.org/CVPPP2014.

of the leaves [12,1,4,7,5], the color and the texture [21] or the combination of the three former traits [10,2,26]. The importance of considering vein information, in conjunction with other features [20,8] or solely [14,16], has also been proposed in recent works. The reported results indicate that even when the leaves share a similar appearance in shape, size, color and texture, as it happens when dealing with several varieties within the same species, some differences in the veins can be established and classification can be attempted, even though it remains as a difficult problem [13,15].

In the work by Larese *et al.* [13], the authors proposed a method based on single classifiers and multiscale morphological vein and areole features to evaluate the automatic recognition of soybean varieties. The species identification and the varieties identification were treated as two separate problems. Varieties identification is challenging since the differences in venation between the classes are not identifiable by humans. However, it is reasonable to expect some venation differences caused by variety adaptation (e.g., draught tolerance). Unlike previous works, the authors showed that it is possible to find some vein differences between the varieties. Resorting to Recursive Feature Elimination (RFE) [11] in combination with several automatic classifiers, they estimated the importance of the traits in order to find the most relevant ones for the characterization of each class. However, the accuracies reported using single classifiers in the recognition were not very high. This may be due to the existence of subtle differences between the cultivars, showing the need to improve the procedure.

More recently [15], the authors focused on the problem of classifying leaves where the classes corresponding to the species and varieties were mixed up in a single classification task. The goal was to perform the leaf identification both for very different leaves, such as from different species, as well as leaves with similar appearance (leaf shape, size, color, texture), such as from different varieties from the same species. The authors proposed to employ hybrid consensus learning to perform classification using multiscale morphological and areole features, outperforming the accuracies of previous works and human classification. Better classification methods could lead to an improved identification of the most informative distinctive features, which could be related to genotype differences.

In this paper, we describe a procedure which can be useful to discover representative leaf vein patterns for each species or variety under analysis. Similarly to previous works in the literature, we consider three legumes, namely red bean, white bean and soybean. Soybean specimens are also divided in three cultivars. In total there are five leaf vein image classes.

Unlike recent works in the literature [13,15] which limited their analysis to a central vein patch from the leaf, we propose to use the entire venation image instead for further processing. We extract local vein features from the whole leaf pursuing to find characteristic patterns for each species/variety, adding potential information from other locations around the center. We still exclude the leaf contour and leaf shape dependence.

In order to find the discriminative patterns, we detect SIFT (Self-Invariant Feature Transform) keypoints in the segmented vein images. Then, we build the Bag of Words (BoW) model using SIFT descriptors, and we perform classification resorting to Support Vector Machines (SVM) [25] with a Gaussian kernel.

SIFT [19] is a state-of-the-art method able to identify and describe image keypoints (saliencies) under a wide range of illumination variations, traslation, scale and rotation, using a localized set of gradient orientation histograms.

As the number of detected SIFT keypoints is different for each image, we use the BoW model [9] to build a fix-sized image descriptor which stores the vein patterns occurrences. In this work we show how it can be useful for vein patterns characterization, and how we used it to highlight the most representative patterns for each species and variety.

Our classification accuracies outperform recent results available in the literature and manual classification, showing the advantages of the whole procedure. Better accuracies indicate that the method is able to find more differences in the vein patterns between the classes. This can be used as an exploratory tool to investigate the nature of these differences in the leaf venation of the different species/varieties.

The contribution of this paper is two-fold. Firstly, the proposed classification framework outperforms the state-of-the-art results on the same dataset, specially for the more difficult (minority) classes, augmenting the benefit over manual classification. Secondly, the method provides a means to detect frequent vein patterns representative for each species and variety under study.

The rest of the paper is organized as follows. In Section 2 we describe the dataset used in this work. The vein segmentation procedure is detailed in Section 3. We explain the feature extraction and classification steps in Sections 4 and 5, respectively. We present and discuss the experimental results as long as the comparison of the performances for the different algorithms in Section 6. Finally, we draw some conclusions in Section 7.

2 Legume image dataset

The image dataset is composed by 866 color leaf images provided by Instituto Nacional de Tecnología Agropecuaria (INTA, Oliveros, Argentina). It consists of 272 images of red bean leaves, 172 images of white bean leaves (*Phaseolus Vulgaris*) and 422 images of soybean leaves (*Glycine max (L) Merr*). The soybean images are divided into three cultivars: 198 from cultivar #1, 176 from cultivar #2 and 48 from cultivar #3. They correspond to the images of the two first foliage leaves (pre-formed in the seed) of each specimen after 12 days of seedling grow. First foliage leaves were selected since their characteristics are less influenced by the environment. The leaves were acquired using a standard flatbed scanner (Hewlett Packard Scanjet-G 3110) at a resolution of 200 pixels per inch, and the images were stored as 24-bit uncompressed TIFF images. The abaxial surfaces of the leaves were scanned since veins appear stronger on this side. All the leaves lay in the same vertical position, thus avoiding significant rotation influences.

Figure 1 shows some typical exemplars from each one of the species and varieties which form the dataset. The reader should notice that the differences between individuals from some of the classes do not compensate for the high variability also present between individuals within the same class. Considering the complete dataset, this application problem is characterized by relatively small differences among different classes and highly variable characteristics within each class.

3 Vein segmentation

We segmented the leaf images following the approach described in the work by Larese *et al.* [16]. We worked with the grayscale image for each leaf. We discarded color information since we are interested in detecting vein patterns associated to vein morphology only. The leaf segmentation is based on the computation of the Unconstrained Hit-or-Miss Transform (UHMT)[24]. The UHMT is an extension of the Hit-or-Miss Transform (HMT) for gray scale images. It extracts all the pixels matching a certain foreground and background neighboring configuration. A composite structuring element $\mathbf{B}$ is employed, which is a disjoint set formed by one structuring element that specifies the foreground configuration, B_{fg}, and one structuring element for the background setting, B_{bg}. The origin of the composite structuring element matches the foreground. For this purpose, we used the same structuring elements described in the work by Larese *et al.* [16].

The UHMT is defined as

$$UHMT_{\mathbf{B}}(Y)(y) = \max\left\{\varepsilon_{B_{fg}}(Y)(y) - \delta_{B_{bg}}(Y)(y), 0\right\}, \tag{1}$$

where Y is a gray scale image with set of pixels y and $\mathbf{B}$ is a composite structuring element. It can be computed as the difference between an erosion with B_{fg}, $\varepsilon_{B_{fg}}(Y)(y)$, and a dilation with B_{bg}, $\delta_{B_{bg}}(Y)(y)$, if $\delta_{B_{bg}}(Y)(y) < \varepsilon_{B_{fg}}(Y)(y)$. Otherwise it equals 0.

We computed the UHMT on five leaf scale images (at 60%, 70%, 80%, 90% and 100% of the original image size). Each UHMT image highlights a different level of vein detail. Next, we added the five UHMTs (resized to the original size) to form the combined UHMT. This combined image highlights both small and large veins. Finally, to obtain the segmented veins we applied contrast enhancement techniques (adaptive histogram equalization) and umbralization (adaptive thresholding). The segmented veins corresponding to the leaves shown in Fig. 1 are depicted in Fig. 2.

The major difference from previous works in the literature [16] is that we did not crop a central vein patch, but used the entire venation image instead for further processing. The reason for this is the advantage provided by SIFT to extract local vein features from the whole leaf region. This decision aims to avoid the location restriction used in previous works, which was limited to finding differences and patterns only in the central part of the leaf, thus ignoring potential information in other locations around the center. In this work we extend the analysis to the whole leaf, while still excluding the leaf contour and keeping the feature independence with respect to the leaf shape, as explained in the next section.

4 Feature extraction

4.1 Scale Invariant Feature Transform (SIFT) for vein saliencies detection and feature extraction

In order to discard keypoints associated to the leaf contour and shape information, we need to skip the keypoints lying near the contour. For this purpose, we create a binary mask by eroding a binary image of the leaf with a square 25×25-pixel sized structuring element (less than approximately the 10% of the image size). This mask is used to limit the keypoints detection to the inner part of the leaf blade.

We use SIFT [19] to perform the venation keypoints detection and feature representation on the segmented vein images, since it is robust to illumination changes, scale, rotation, small shifts and affine transformations

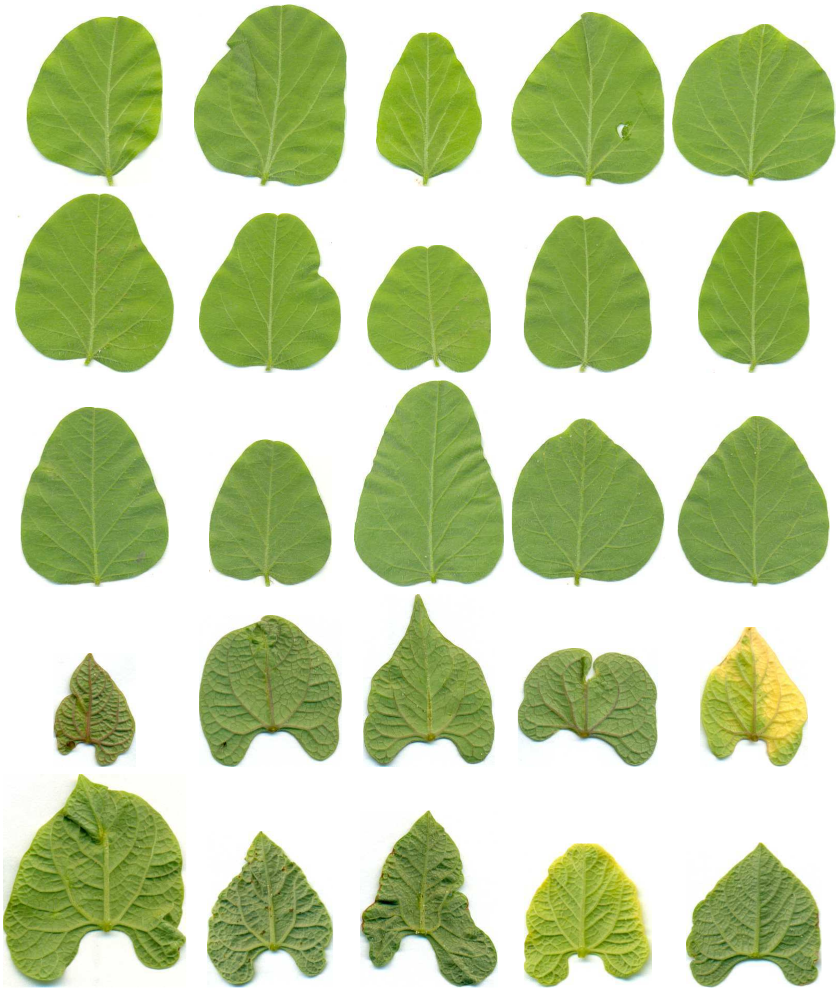


Fig. 1 Sample leaves from each class. First row: Soybean (cultivar #1). Second row: Soybean (cultivar #2). Third row: Soybean (cultivar #3). Fourth row: Red bean. Fifth row: White bean.

by means of the usage of local sets of gradient orientations. SIFT has been successfully applied on binary images in the recent literature [18]. As described in the work by Lowe [19], SIFT is composed of four main steps, as follows.

1. Scale-space extrema detection: The method builds a scale-space by convolving a variable-scale Gaussian with the original input image, and computes the Difference of Gaussian (DoG) as an approximation of the Laplacian of Gaussian filter for blob detection. After that, local extrema are searched in both spatial and scale spaces to detect the potential keypoints.
2. Keypoint localization: At this stage, unreliable keypoints are removed. For this purpose, keypoints having low intensity values and keypoints corresponding to edges are discarded.
3. Orientation assignment: In order to make the descriptor invariant to rotation, for each keypoint the method computes the orientation histogram using the local gradient magnitudes and directions at the points around it. A regular grid of 16×16 points over the region is used for sampling purposes. The points are weighted by the corresponding gradient magnitude and a circular Gaussian function. In this way, the pixels closer to the center of the region (estimated location of the keypoint) have higher weights. The keypoint orientation is selected as the most frequent value from the orientation histogram.
4. Keypoint descriptor: The gradient orientation is computed for each sampled point, and with these values a 4×4 grid of gradient orientation histograms with 8 orientation bins each is constructed. The contribution of the gradient information to the spatial/orientation histogram is smoothly distributed

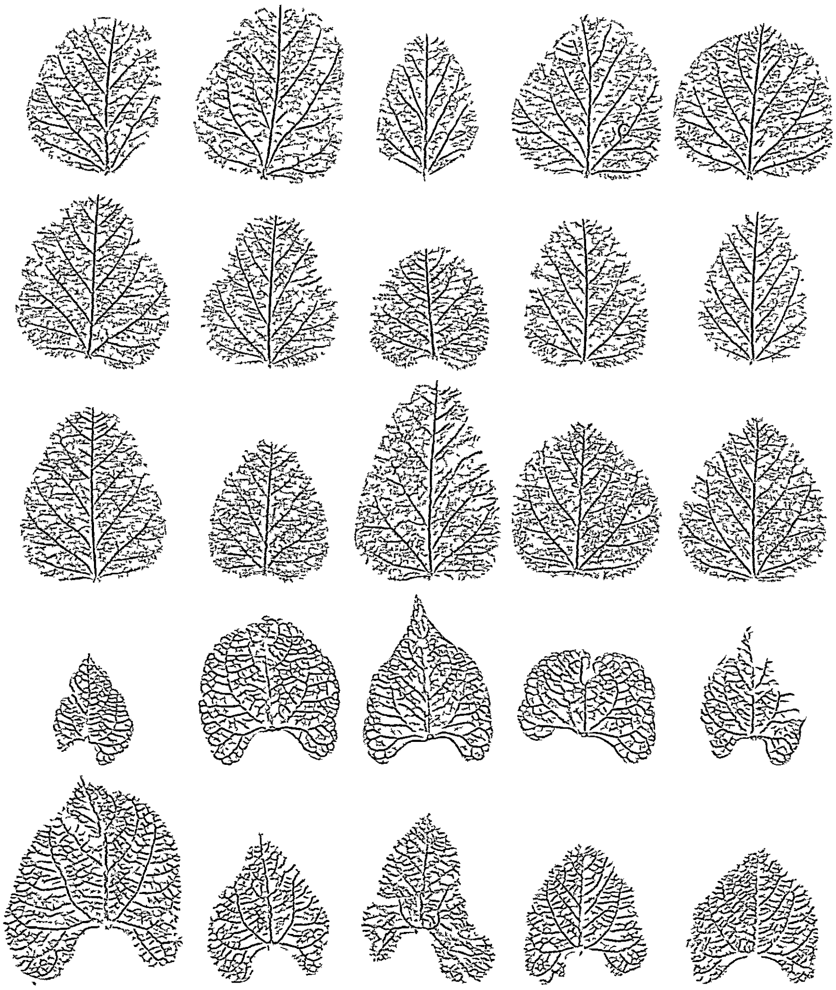


Fig. 2 Sample venation images from each class. First row: Soybean (cultivar #1). Second row: Soybean (cultivar #2). Third row: Soybean (cultivar #3). Fourth row: Red bean. Fifth row: White bean.

among the adjoining bins using trilinear interpolation. Finally, SIFT concatenates the histogram entries as a vector of $4 \times 4 \times 8 = 128$ features. The vector is normalized to unit length in order to correct for illumination changes.

4.2 Building the vein pattern dictionary via a Bag of Words (BoW) model

The number of saliencies detected on each vein image is different, thus the number of SIFT feature vectors is variable for each one. In order to summarize the vein image information the BoW model is employed [9]. For this purpose, the SIFT local descriptors are clustered into a fixed number of clusters (vocabulary words which determine the vocabulary size), and a normalized histogram of occurrences is constructed. In this way, the variable number of SIFT feature vectors per image is summarized by their word counts. The goal is to assign similar SIFT feature vectors to the same cluster, which correspond to local patches which are visually similar.

We constructed the BoW model using the SIFT feature vectors from the training images. Even though we cluster SIFT local descriptors into vocabulary words, there is a direct connection between the feature descriptors and the keypoints or saliencies from where they were extracted. For this reason, in the following text we sometimes refer to the SIFT local patches instead of feature descriptors.

In this work we tried five different vocabulary sizes, namely 10, 50, 100, 200 and 1000 words, and compared the performance obtained in each case. We did not consider dictionary sizes greater than 1000 words (as in object detection applications) since we are using binary images, and thus the pixel configurations are simpler. We used K-means as the clustering algorithm given its simplicity and since it is one of the most popular algorithms for constructing visual vocabularies [6].

5 Classification

Leaf classification was performed resorting to Support Vector Machines (SVMs). This selection is based on the comparably good performance achieved by SVMs against other state-of-the-art automatic classifiers reported in previous works [16,13] on the same application problem.

Support Vector Machines (SVM) [25] is a state-of-the-art classifier which assumes that applying an appropriate nonlinear mapping of the data into a sufficiently high dimensional space, two classes can be separated by an optimum hyperplane. This decision hyperplane is chosen in such a way that the distance between the nearest patterns of different classes (i.e., the margin) is maximized.

Given a dataset $D = \{(\mathbf{x}_i, y_i)\}$, formed by pairs of features-label examples, with $\mathbf{x}_i \in \mathbb{R}^d$, $y_i \in \{-1, 1\}$ and $i = 1, \ldots, n$, consider the case where the training examples can be linearly separated. In this case, the two classes can be separated by one of many possible hyperplanes given by:

$$f(\mathbf{x}_i) = \mathbf{w}^T \mathbf{x}_i + b = 0, \tag{2}$$

where $\mathbf{w} \in \mathbb{R}^d$ and $b \in \mathbb{R}$. A support vector classifier selects the hyperplane which maximizes the margin. This optimization problem can be posed as

$$\min_{\mathbf{w},b} ||\mathbf{w}||, \quad \text{subject to} \quad y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1. \tag{3}$$

If the classes are not completely separable (there is overlap in feature space), some patterns might be allowed to be on the opposite side of the margin by introducing the slack variables $\xi = \{\xi_1, \xi_2, \ldots, \xi_n\}$, and converting the minimization problem in Eq. (3) into:

$$\min_{\mathbf{w},b} \frac{1}{2}||\mathbf{w}||^2 + C \sum_{i=1}^{n} \xi_i,$$

$$\text{subject to} \begin{cases} \xi_i \geq 0; \\ y_i(\mathbf{w}^T \mathbf{x} + b) \geq (1 - \xi_i), \forall i, \end{cases} \tag{4}$$

where C is a regularization constant which controls the trade-off between the complexity of the classifier and the number of allowed misclassifications.

If the decision surface is required to be nonlinear, a kernel function can be used to map the original features into a high dimensional space, where they can be separated by a linear boundary. The kernel κ is related to the transform θ following $\kappa(\mathbf{x}_i, \mathbf{x}_j) = \theta(\mathbf{x}_i)\theta(\mathbf{x}_i)$. In this case, the problem can be stated as $f(\mathbf{x}_i) = \mathbf{w}^T \theta(\mathbf{x}_i) + b$, and an optimization problem similar to Eq. (4) can be derived.

In this work, we considered the Gaussian kernel:

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(\frac{||\mathbf{x}_i - \mathbf{x}_j||^2}{2\sigma^2}\right). \tag{5}$$

Both the standard deviation σ for the Gaussian kernel and the regularization parameter C were optimized using inner validation during the training.

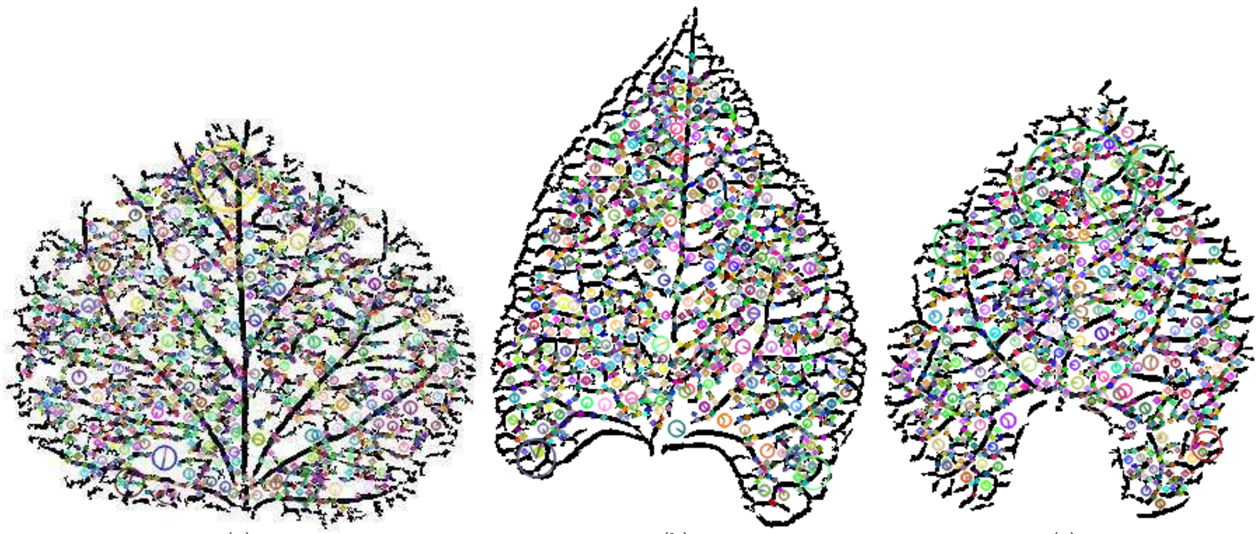
SVM is a binary classifier. In order to extend its use to the present multiclass problem, we used the one-vs-one strategy. In this strategy, $k(k-1)/2$ binary classification problems are formulated between all pairs of the k classes. The final result is obtained using a *max-wins* criterion: the example is preliminary assigned to one of two classes by each binary classifier, the corresponding class adds a vote, and the pattern is finally classified into the class with the maximum number of votes.

For all the experiments, we computed the average classification accuracy after performing 10 runs of 5-fold cross validation.

6 Experimental results

Figure 3 shows some examples of the local patches detected by SIFT for a soybean leaf, a red bean leaf and a white bean leaf. As it can be noticed from this figure, the keypoints (centers of the colored circles) are spreaded all over the leaves except for the border of the blade, thus avoiding the influence of the different leaf shapes. The size of each circle represents the extension of the region of interest according to the Gaussian pyramid in SIFT detection, whereas the radial line inside each circle indicates the orientation of the gradient. Different number of patches are obtained for each image, even for those from the same species/variety.

In Table 1 we report the mean ± standard error of the total accuracies obtained from the 10 runs of 5-fold cross validation, resulting from the usage of the framework based on the combination of SIFT, BoW and SVM, and the five different numbers of words (clusters). The total accuracy per fold is computed as the total number of correctly classified leaves in the fold (from all the classes) divided by the total number of examples in the fold. In addition, the lower part of Table 1 presents previous results reported in the recent literature [15] for reference. These results are based on the extraction of morphological vein and areole features



(a) (b) (c)

Fig. 3 SIFT keypoint detection on sample vein images. (a) Soybean. (b) Red bean. (c) White bean. Each color circle stands for the size of the region of interest corresponding to the keypoint located at its center. The radial line indicates the orientation of the gradient.

(MF) and ensembles classification (HCE and WHCE). MF in combination with SVM is also included.

Table 2 depicts the mean ± standard error of the per class accuracies obtained from the 10 runs of 5-fold cross validation. The per class accuracies per fold are computed as the total number of correctly classified leaves for each class in the fold divided by the total number of leaves from that class in the fold. As before, previous results in the literature are presented for reference. The manual average classification performances and standard errors reported in [15] are also included. In this case, five human experts were asked to manually classify the leaf veins. It is worth noticing that the experts solved two easier problems. Instead of classifying the leaves into 5 classes, they performed the classification into 3 classes, namely red bean, white bean and soybean. In a second independent experiment, they classified only soybean leaves into the three possible cultivars. Manual classification accuracies are a reference baseline that measures how detectable the vein differences are for human beings. The standard error of manual classification reflects the variability across the experts.

It can be noticed from the results in Table 1 that when 100 words or more are used, the mean accuracy is over 90%. Increasing the number of words from 200 to 1000 does not provide any benefit in terms of the mean total accuracy. On the contrary, this accuracy is 1% reduced when using a high-sized dictionary. This is due to the low performance achieved with a dictionary of 1000 words in the recognition of leaves from Soybean #3. This can be observed in Table 2, where there is a consistent improvement of the per class accuracies along with the increase of the dictionary size, reaching its best performance for 1000 words, except for a big drop for Soybean #3 (the minority class). This fact impacts negatively in the total accuracy depicted in Table 1, making 1000 words not appropriate for our problem since the recognition of the minority class is very poor (even lower than for manual classification).

It is also clear from Tables 1 and 2 that the framework used in this paper highly improves the class identification, indicating that better discriminative differences are found in the features corresponding to each species/variety. SIFT + BoW + SVM outperforms the previous results and the manual classification for all the cases when the dictionary size is at least of 100 words big. The only exception is for the identification of Soybean #3 using 1000 words, as previously discussed. Specially, the dictionary with 100 words has a very good performance in detecting this cultivar, while achieving also high accuracies for the rest of the species and soybean varieties.

In Fig. 4 we show the average Bag of Words descriptors for each one of the five classes using a dictionary of 100 words. It is worth noticing that the same procedure can be applied for a different dictionary size (larger or smaller), allowing to increase or reduce the resolution in vein pattern discrimination. We chose 100 words since it represents a relative manageable number of vein pat-

Table 1 Total accuracy (mean$\pm S_E$) for the five-class species and varieties classification problem.

Algorithm	**Total accuracy (mean$\pm S_E$%)**
SIFT + BoW + SVM	
10 words	84.66 ± 0.40
50 words	89.16 ± 0.33
100 words	90.39 ± 0.31
200 words	91.49 ± 0.24
1000 words	90.49 ± 0.23
MF + HCE [15]	72.45 ± 0.42
MF + WHCE [15]	72.97 ± 0.44
MF + SVM [15]	71.05 ± 0.41

Table 2 Per class accuracy (mean$\pm S_E$) for the five-class species and varieties classification problem.

Algorithm	**Per class Accuracy (mean$\pm S_E$%)**				
	RBean (272 images)	WBean (172 images)	SBean#1 (198 images)	SBean#2 (176 images)	SBean#3 (48 images)
SIFT + BoW + SVM					
10 words	77.46 ± 0.95	82.21 ± 0.98	69.26 ± 1.11	43.64 ± 1.11	50.71 ± 1.96
50 words	84.82 ± 0.76	88.01 ± 0.86	72.37 ± 1.02	58.74 ± 1.14	58.11 ± 2.34
100 words	90.60 ± 0.60	93.31 ± 0.65	73.89 ± 0.83	65.59 ± 1.28	51.98 ± 2.29
200 words	93.86 ± 0.46	95.87 ± 0.54	74.34 ± 0.75	68.05 ± 1.32	47.96 ± 2.60
1000 words	97.10 ± 0.27	98.32 ± 0.30	79.28 ± 1.00	74.14 ± 0.93	28.02 ± 1.93
MF + HCE [15]	89.89 ± 0.55	86.97 ± 0.90	59.79 ± 1.05	53.13 ± 1.15	44.56 ± 2.33
MF + WHCE [15]	90.26 ± 0.54	87.26 ± 0.85	61.86 ± 1.07	53.58 ± 1.16	40.60 ± 2.08
MF + SVM [15]	89.86 ± 0.60	82.96 ± 0.91	67.31 ± 1.16	51.41 ± 1.32	9.16 ± 1.37
Manual classification [15]	83.28 ± 3.71	70.82 ± 13.15	44.95 ± 2.00	42.78 ± 5.37	43.98 ± 6.97

terns and a good classification performance at the same time, as previously shown in Tables 1 and 2.

As a reference, we also plot in Fig. 4 two dashed lines at the 10% and 90% of the maximum value, representing thresholds for analyzing the most and least frequent vein patterns occurring at each class. From the figure it can be noticed that the most frequent patterns are the ones corresponding to words 31, 70, 78 and 90 for white bean, 4 and 70 for red bean, and 24 for the three soybean cultivars. Cultivar #3 also has high occurrences of patches from cluster 77. The five classes have few patches from cluster 40, and soybean cultivar #1 also from word 48. This information is summarized in Table 3, where some word examples are included for descriptive purposes. Cluster 40 is not included in the table since it corresponds to an infrequent pattern for all the classes of interest. The local patches shown in Table 3 have different sizes and orientations since SIFT can detect similar configurations invariantly to scale, rotation and traslation.

In order to show in context the characteristic visual words for each class, we exemplify in Figures 5 to 7 portions of white bean, red bean and soybean leaf veins with the different vein patterns from Table 3 highlighted in different colors. We used training images for this purpose. For each leaf class, the most frequent words were selected on the basis of the statistical information provided by the average Bag of Words profiles shown in Fig. 4, *i.e.*, the clusters having a frequency of over 90%. Next, from these clusters we collected some example patches previously extracted by SIFT and corresponding to the same training images. These patches were clustered according to the BoW model during the training phase in order to construct the feature vector (BoW profile) for each image. In Figures 5 to 7 we plot these vein patterns in their true coordinates, colored according to their corresponding cluster.

The information provided in Table 3 and Figures 4 to 7 could be analyzed in conjunction in an attempt to infer the meaning of each cluster, and in this way help to describe in more detail the vein characteristics of each species and variety, similarities and differences between them. As a means to exemplify this idea, and based on the analyzed examples, we could hypothesize that words 24 and 78 seem to be describing characteristics related to inner parts of the veins, such as their width and shape. On the other hand, words 31 and 70 could be more related to configurations between neighboring veins. Word 90 would seem to be associated to vein loop patterns. Word 77 may be related to branching

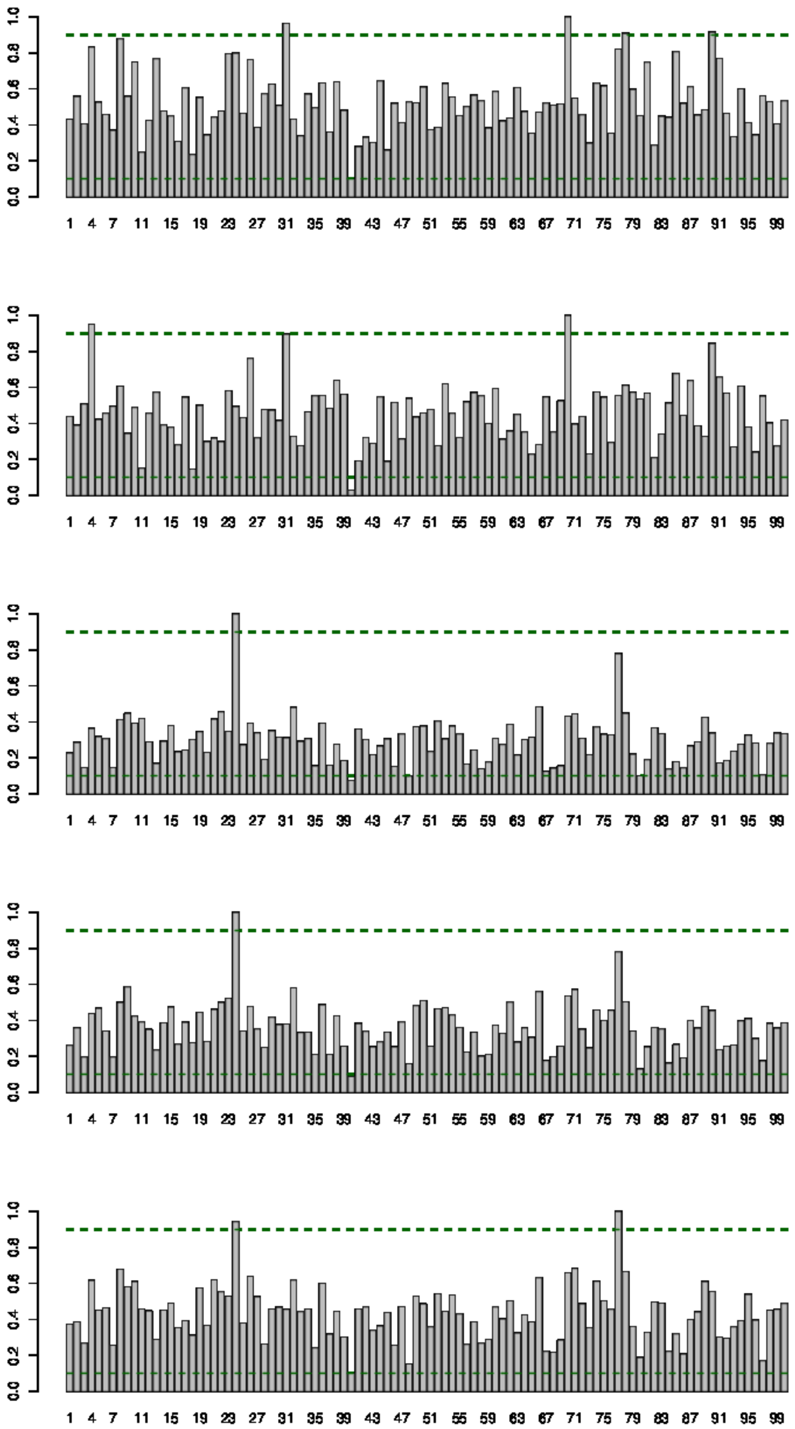


Fig. 4 Average Bag of Words for each one of the five plant species and varieties, including the 10% and 90% of the maximum value (dashed lines) for a dictionary size of 100 words. First row: White bean. Second row: Red bean. Third row: Soybean cultivar #1. Fourth row: Soybean cultivar #2. Fifth row: Soybean cultivar #3.

patterns, while word 4 seems to be describing terminal or border patterns in the veins.

A deeper analysis of the vein patterns is out of the scope of this paper. On the contrary, our goal is to emphasize the potential of the tools employed in this work. The relevance of the proposed framework resides in its potential to help the specialists find local keypoint patterns within the leaf venation. These local patterns are those which better allow to discriminate between the species and varieties. For example, the specialist could be provided with different filters to plot in the context of a selected image all the patterns associated to a certain cluster, or from all the most frequent clusters. The correlation with the location of the patterns inside the leaf could also be studied, for example, in order to see if they are homogeneously distributed or concentrated at certain regions (left/right sides, center/border areas, top/bottom, etc.).

The brief analysis presented here could be extended in many ways. Just to mention a few ideas, different thresholds could be considered in the profiles in Fig. 4 in order to expand the analysis to other words. Moreover, the most discriminative patterns could be described by computing different measures on them in order to better understand their properties. As an exploratory tool, the kind of analysis to be performed depends on the specialists stated objectives.

7 Conclusions

In this paper we show how the combination of SIFT, BoW and SVM can be used to detect relevant vein patterns which characterize different species and varieties. The classification accuracies obtained by means of this procedure outperform previous results in the literature and manual expert classification. Better classification accuracies indicate that the automatic classifier is able to detect more vein differences between the classes. These differences are representations of local patches spreaded all over the blade (except for the leaf contour), extracted by SIFT and represented by BoW.

Characteristic patterns can be detected by analyzing the class profiles obtained by BoW. The most frequent vein patches seem to be related to different configurations of vein patterns, such as loops, branching and vein width and shape. A deeper analysis of the vein patterns is out of the scope of this work, since we aim to show how the employed state-of-the-art tools can be exploited for vein characterization purposes. However, the study can be extended depending on the demand of Biology specialists inquiries.

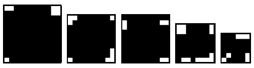
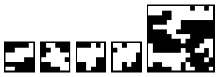
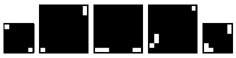
Future work can be directed in several ways. One direction may be to compare the vein patches obtained by SIFT with other state-of-the-art feature detectors and descriptors, such as SURF (Speeded-Up Robust Features) [3]. Another interesting pending task consists in studying the connections or topological relationships between the discovered relevant vein patterns.

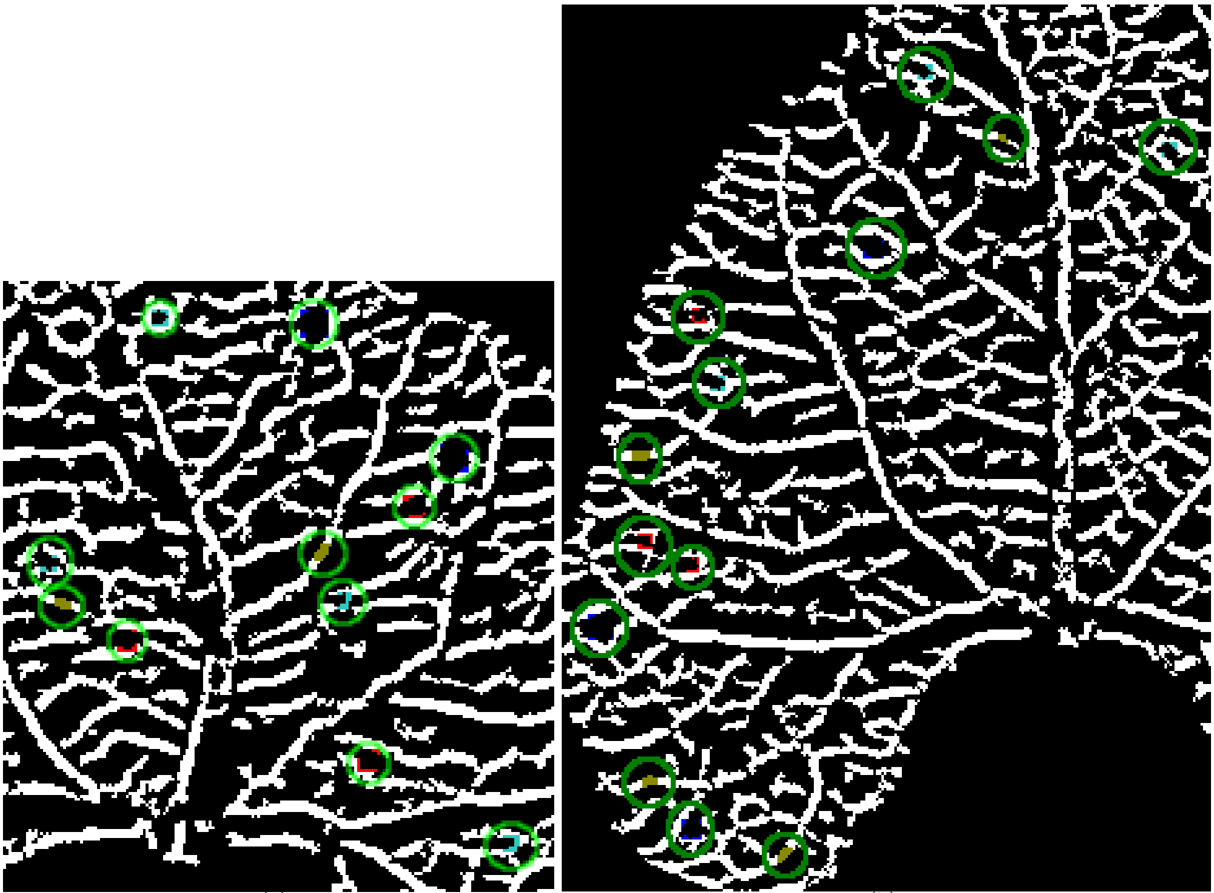
Acknowledgements MGL and PMG acknowledge grant support from ANPCyT PICT 2012-0181. We also acknowledge technical support from R. Craviotto, M. Arango and C. Gallo at Instituto Nacional de Tecnología Agropecuaria (INTA Oliveros, Argentina).

References

1. Agarwal, G., Ling, H., Jacobs, D., Shirdhonkar, S., Kress, W., Russell, R., Belhumeur, P., Dixit, N., Feiner, S., Mahajan, D., Sunkavalli, K., White, S.: First Steps Toward an Electronic Field Guide for Plants. Taxon, Journal of the International Association for Plant Taxonomy **55**, 597–610 (2006)
2. Bama, B.S., Valli, S.M., Raju, S., Kumar, V.A.: Content based leaf image retrieval (CBLIR) using shape, color and texture features. Indian Journal of Computer Science and Engineering **2**(2), 202–211 (2011)
3. Bay, H., Tuytelaars, T., Van Gool, L.: SURF: Speeded Up Robust Features. In: A. Leonardis, H. Bischof, A. Pinz (eds.) Computer Vision – ECCV 2006, *Lecture Notes in Computer Science*, vol. 3951, pp. 404–417. Springer Berlin Heidelberg (2006)
4. Camargo Neto, J., Meyer, G.E., Jones, D.D., Samal, A.K.: Plant species identification using Elliptic Fourier leaf shape analysis. Computers and Electronics in Agriculture **50**, 121–134 (2006)
5. Chaki, J., Parekh, R.: Designing an automated system for plant leaf recognition. International Journal of Advances in Engineering & Technology **2**(1), 149–158 (2012)
6. Chatfield, K., Lempitsky, V., Vedaldi, A., Zisserman, A.: The devil is in the details: an evaluation of recent feature encoding methods. British Machine Vision Conference, (2011)
7. Du, J.X., Wang, X.F., Zhang, G.J.: Leaf shape based plant species recognition. Applied Mathematics and Computation **185**(2), 883–893 (2007). Special Issue on Intelligent Computing Theory and Methodology
8. Du, J.X., Zhai, C.M., Wang, Q.P.: Recognition of plant leaf image based on fractal dimension features. Neurocomputing **116**, 150–156 (2013)
9. Fei-Fei, L., Perona, P.: A Bayesian hierarchical model for learning natural scene categories. In: Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 2, pp. 524–531 vol. 2 (2005)
10. Golzarian, M.R., Frick, R.A.: Classification of images of wheat, ryegrass and brome grass species at early growth stages using principal component analysis. Plant Methods **7:28** (2011)
11. Guyon, I., Weston, S., Barnhill, S., Vapnik, V.: Gene Selection for Cancer Classification Using Support Vector Machines. Machine Learning **46**(1-3), 389–422 (2002)
12. Im, C., Nishida, H., Kunii, T.L.: Recognizing Plant Species by Leaf Shapes-A Case Study of the Acer Family. Pattern Recognition, International Conference on **2**, 1171 (1998)

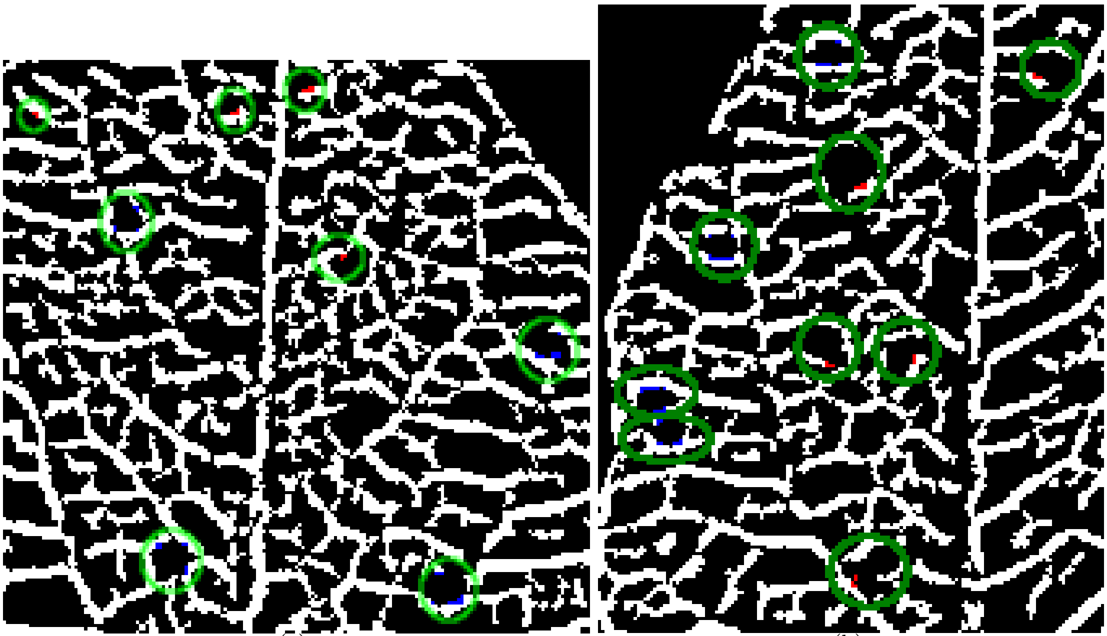
Table 3 Examples of characteristic vein patterns found by the proposed procedure for each species and variety, using a dictionary size of 100 words. ↑ represents that the pattern occurs very frequently; ↓ indicates the almost absence of the pattern. White pixels belong to segmented veins.

White bean	 *Cluster 31* ↑	 *Cluster 70* ↑	 *Cluster 78* ↑	 *Cluster 90* ↑
Red bean	 *Cluster 4* ↑	 *Cluster 70* ↑		
Soybean #1	 *Cluster 24* ↑	 *Cluster 48* ↓		
Soybean #2	 *Cluster 24* ↑			
Soybean #3	 *Cluster 24* ↑	 *Cluster 77* ↑		



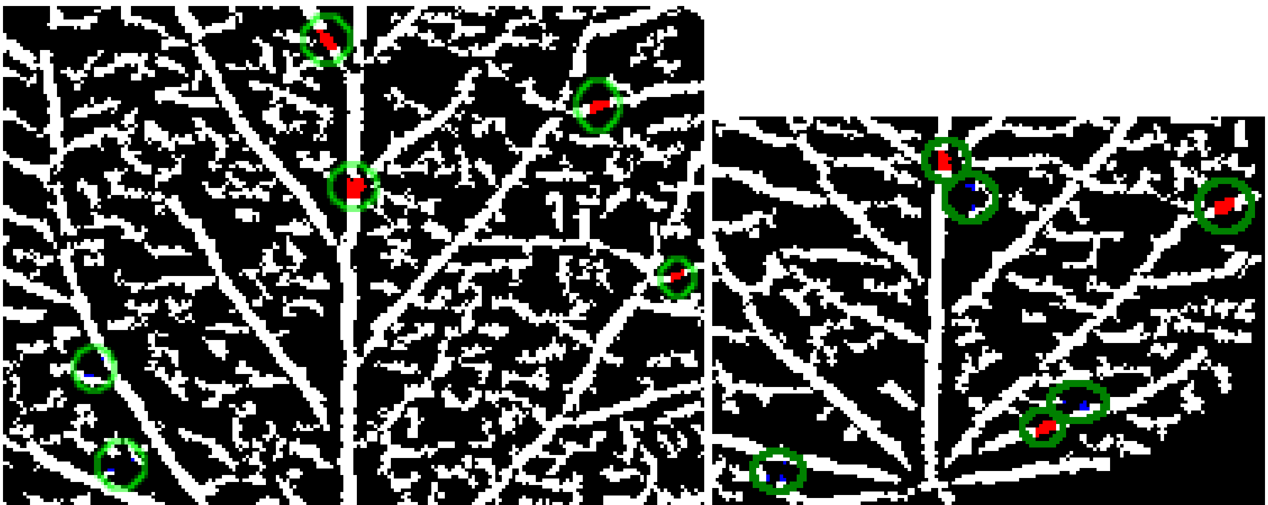
(a) (b)

Fig. 5 Some examples of the most frequent vein patterns found in white bean leaves using a dictionary size of 100 words, shown in leaf vein context. Each pattern is highlighted using a different color: words 31 (red), 70 (blue), 78 (mustard) and 90 (cyan).



(a) (b)

Fig. 6 Some examples of the most frequent vein patterns found in red bean leaves using a dictionary size of 100 words, shown in leaf vein context. Each pattern is highlighted using a different color: words 4 (red) and 70 (blue).



(a) (b)

Fig. 7 Some examples of the most frequent vein patterns found in soybean cultivar #3 leaves using a dictionary size of 100 words, shown in leaf vein context. Each pattern is highlighted using a different color: words 24 (red) and 77 (blue).

13. Larese, M.G., Bayá, A.E., Craviotto, R.M., Arango, M.R., Gallo, C., Granitto, P.M.: Multiscale recognition of legume varieties based on leaf venation images. Expert Systems with Applications **41**(10), 4638–4647 (2014). DOI 10.1016/j.eswa.2014.01.029
14. Larese, M.G., Craviotto, R.M., Arango, M.R., Gallo, C., Granitto, P.M.: Legume Identification by Leaf Vein Images Classification. In: L. Alvarez, M. Mejail, L. Gomez, J. Jacobo (eds.) Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, *Lecture Notes in Computer Science*, vol. 7441, pp. 447–454. Springer Berlin Heidelberg (2012)
15. Larese, M.G., Granitto, P.M.: Hybrid Consensus Learning for Legume Species and Cultivars Classification. In: L. Agapito, M.M. Bronstein, C. Rother (eds.) Computer Vision - ECCV 2014 Workshops, *Lecture Notes in Computer Science*, vol. 8928, pp. 201–214. Springer International Publishing (2015)

16. Larese, M.G., Namías, R., Craviotto, R.M., Arango, M.R., Gallo, C., Granitto, P.M.: Automatic classification of legumes using leaf vein image features. Pattern Recognition **47**(1), 158–168 (2014)
17. Li, L., Zhang, Q., Huang, D.: A Review of Imaging Techniques for Plant Phenotyping. Sensors **14**(11), 20,078–20,111 (2014). DOI 10.3390/s141120078
18. Lin, W.S., Wu, Y.L., Hung, W.C., Tang, C.Y.: A Study of Real-Time Hand Gesture Recognition Using SIFT on Binary Images. In: J.S. Pan, C.N. Yang, C.C. Lin (eds.) Advances in Intelligent Systems and Applications - Volume 2, *Smart Innovation, Systems and Technologies*, vol. 21, pp. 235–246. Springer Berlin Heidelberg (2013)
19. Lowe, D.G.: Distinctive Image Features from Scale-Invariant Keypoints. Int. J. Comput. Vision **60**(2), 91–110 (2004)
20. Park, J., Hwang, E., Nam, Y.: Utilizing venation features for efficient leaf image retrieval. J. Syst. Softw. **81**(1), 71–82 (2008)
21. Pydipati, R., Burks, T.F., Lee, W.S.: Identification of citrus disease using color texture features and discriminant analysis. Computers and Electronics in Agriculture **52**, 49–59 (2006)
22. Sack, L., Dietrich, E.M., Streeter, C.M., Sanchez-Gomez, D., Holbrook, N.M.: Leaf palmate venation and vascular redundancy confer tolerance of hydraulic disruption. PNAS USA **105**, 1567–1572 (2008)
23. Scoffoni, C., Rawls, M., McKown, A.D., Cochard, H., Sack, L.: Decline of leaf hydraulic conductance with dehydration: relationship to leaf size and venation architecture. Plant Physiology **156**, 832–843 (2011)
24. Soille, P.: Morphological Image Analysis: Principles and Applications. Springer-Verlag (1999)
25. Vapnik, V.: The nature of statistical learning theory. Springer-Verlag (1995)
26. Yanikoglu, B., Aptoula, E., Tirkaz, C.: Automatic plant identification from photographs. Machine Vision and Applications **25**(6), 1369–1383 (2014). DOI 10.1007/s00138-014-0612-7