ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/228441194

Improved Classification of Pollen Texture Images Using SVM and MLP

Article · January 2003

CITATIONS
9

READS
70

4 authors, including:

Pilar Carrión
University of Vigo
18 PUBLICATIONS **150** CITATIONS
SEE PROFILE

J.F Galvez
University of Vigo
25 PUBLICATIONS **162** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

DETEPRE View project

Bioinformatics View project

All content following this page was uploaded by J.F Galvez on 29 May 2014.

The user has requested enhancement of the downloaded file.

IMPROVED CLASSIFICATION OF POLLEN TEXTURE IMAGES USING SVM AND MLP

M. Fernández-Delgado
Dpto. de Electrónica e Computación
Universidade de Santiago de Compostela
15782 Santiago de Compostela (Spain)
delga@dec.usc.es

P. Carrión, E. Cernadas, J.F. Gálvez
Dpto. de Informática, E.S.E.I.
Universidade de Vigo
Campus Universitario As Lagoas s/n
32004 Ourense (Spain)
pcarrion@uvigo.es, cernadas@ei.uvigo.es

Pilar Sá-Otero
Dpto. de Bioloxía Vexetal e Ciencias do Solo
Facultad de Ciencias, Universidade de Vigo
Campus Universitario As Lagoas
32004 Ourense (Spain)

ABSTRACT
Humans are interested in the determination of the geographical origin of honeybee pollen due to its nutritional value and therapeutical benefits. This task is currently being developed in a manual way using images from optical microscopy. We have proposed [1, 2] an automatic system for pollen identification, based on its texture classification using a minimum distance classifier. In the present paper, we explore the use of more sophisticated classifiers to improve the classification stage. Specifically, we apply several well-known classifiers, KNN, Support Vector Machine and Multi-Layer Perceptron, in order to increase the classification rate on this problem.

KEY WORDS
Image analysis, Texture classification, Pollen loads, Support Vector Machine, Multi-Layer Perceptron.

1 Introduction

Humans use products coming from the hive such as honey, royal jelly or apicultural pollen for different purposes. They are interested on honeybee pollen as nutritional complement due to its high content in amino acids, vitamin, etc. Their consumption in Spain and other countries is relatively recent, but in the last years it has become economically very important. Nowadays, the apitherapy also studies their curative characteristics.

Corbicula pollen is the essential feeding for the hive. The worker bees collect the pollen in the flower, form small balls, stick them to the corbiculas of their back legs and carry them to the hive. The pollen loads collected come from plants placed in the surrounding of the hive. So, this local flora influences the palynological composition of pollen loads. The presence of a specific combination of pollen types in a sample indicates its geographical origin. The current method to determine the floral origin is analyzing the pollen by an optical microscopy and then counting the number of pollen of vegetal species [3]. This procedure is tedious and requires expert personal. Industries are interested in the development of methodologies, which can systematically be applied on a chain of production, to classify pollen loads.

Investigations about the composition of pollen loads collected by honeybee have demonstrated [4] that bees form pollen loads with monospecific pollen grains, i.e., pollen grains of only one plant specie. Each pollen plant has special physical characteristics such as colour, size, shape, texture, etc. [5]. In particular, texture has qualitatively been described by some authors using stereoscopic microscope as thick, medium and thin texture [6, 7]. Hence, Li et al. [8] identify pollen grains microtextures using neural networks.

Differences in visual texture of pollen loads are due to the microtexture associated to pollen grain of each vegetable specie, which is related to the structure of its exine and the nature of pollenkit that covers it. So, pollen loads texture must be characteristic of each plant specie. Additionally, the collection process of honeybee pollen and/or the geographical place of collection (characterized by its atmospherical conditions) have an effect on the final appearance of gray-level image intensities (see figure 2). Nevertheless, pallynologists believe that spatial structure of image is retained.

We have investigated the feasibility of the application of computer vision to determine in a fast and precise way the floral origin of apicola pollen marked in the northwest of Spain (Galicia). It will be useful to be able to detect fraud in both product composition and geographical origin of them. The idea is innovative and the only related works in the literature have been done by our group. Firstly, we briefly outline the entire process of honeybee pollen classification [1, 2]. This approach uses a minimum distance classifier based on the Mahalanobis distance. However, in

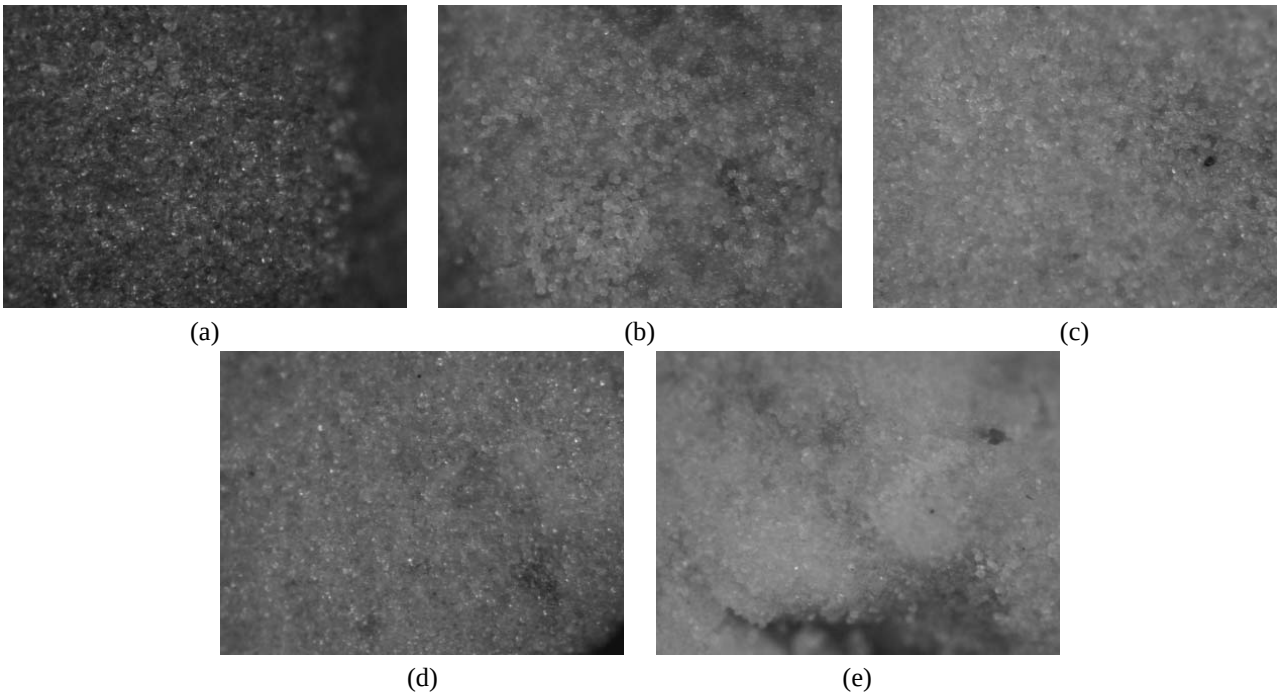


Figure 1. Digital images of *Cytisus* (a), *Castanea* (b), *Quercus* (c), *Rubus* (d) and *Raphanus* (e) pollen loads.

the present work we improve this classification stage by using more sophisticated classifiers, which are described in section 3. A comparison among classification results for the five most abundant plant species of our region are provided and discussed in section 4. Finally, we present the main conclusions at which we arrived and the future work.

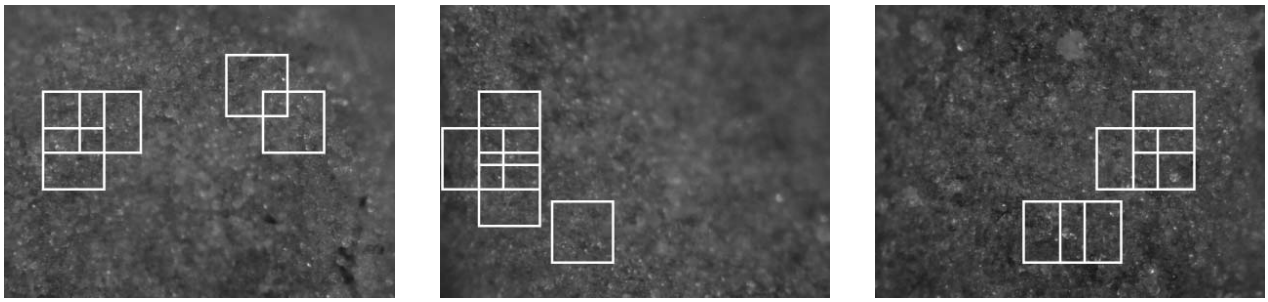
2 System Overview

The system is composed of the following four modules: image adquisition, pre-processing, texture feature generation and classification [1, 2]. Images acquisition of pollen loads is carried out using the infrastructure of Biology Lab (a Nikon SMZ800 magnifying glass connected to a general purpose digital camera Nikon Coolpix E950). Afterwards, images are transferred to the PC by a serial cable. Corbicula pollen is digitized at spatial resolution of 480 point per mm, yielding an image of 1600×1200 pixels. Figure 1 shows images of the five most abundant plant species in the northwest of Spain (Galicia).

As it can be observed in figure 1, pollen images present random blurred areas due to corbicula pollen is not flat over the acquisition area. Then, the pre-processing step tries to extract *non-blurring* regions of interest (ROIs) in the original image. Some metrics have been proposed to quantify the level of blurring in the image [2]. The best performance was achieved using the *Edge Strength (ES)* meassure. It is the mean value of an edge map, which is computed applying an edge operator to the original image (we use Sobel operator). ES will be higher when the image presents many sharp edges, i.e. images are less blurred. Pollen images are sweeping out left to right and top to down and it is taken overlapping regions of $P \times P$ pixels (a shifting of 150 pixels is used and P is fixed to 256 pixels). Next, ES measure is computed on every ROIs of an image and a set of ROIs, which have the highest values of ES, is chosen. Figure 2 shows an example of the ROIs selected on pollen images, where overlays of the boundary of regions are overlapped to the original image.

A texture feature extractor computes image properties on those ROIs that will be used for pollen load classification [2]. We used first-order and second-order statistical features, wavelet packet signatures and filtering features to discriminate among the five plant species mentioned. In particular, the following texture feature vectors were tested [2]: Haralick's coefficients (HC) (7 features) [9, 10], Grey Level Run Length Statistics (GLRLS) (5 features) [11, 12], Neighboring Gray Level Dependence Statistics (NGLDS) (5 features) [12], First-order statistics (FOS) (11 measures) [13], and energy and entropy features computed for three levels of decomposed wavelet packets (feature vectors WE and WH respectively with 12 features) [14, 15] using Daubechies wavelet of filter length 20 (D_{20}) [16].

Alternative methods to exploit spatial dependencies are based on computation of statistical measures on filtered images [11]. Laws [17] proposes the filtering of image with the three basic convolution vectors: $V_A = [1, 2, 1]$, $V_E = [-1, 0, 1]$ and $V_S = [-1, 2, -1]$, for kernel size $L = 3$, and then the computation of energy statistic. If we consider V_A, V_E and V_S as elementary vectors, higher dimensional



(a) (b) (c)

Figure 2. ROIs extracted from an image of *Rubus* pollen load of the hive of Lobios (a), Pontevedra (b) and Viana (c).

vectors can be easily built from the elementary masks. In all cases, mutual multiplication of these vectors for each order L, considering the first term as a column vector and the second one as a row vector, yields 2-D masks of size $L \times L$.

What statistics and what neighbourhood (L) are suitable for solving our texture discrimination problem is a critical decision. It is basically due to the lack of intuitive understanding that humans have about texture parameters. This implies that many texture features are suggested in the literature and the only way to choose the best one for a specific application is by experimental testing.

We compute for each image the following first order statistics: variance (μ_2), 3^{rd} (μ_3) and 4^{th} (μ_4) central moments, energy (m_2) and entropy ($Ent.$). Several texture feature vectors, LLT(L) (*Local Linear Transformations*), are computed as a function of filtering neighbourhood ($L = 3, 5, 7, \ldots$) [2].

The classifier uses these features to assign the pollen load to a specific plant species. A minimum distance classifier (MDC) was used. Let M be the number of classes, L the number of corbicula pollens and J the number of regions of interest extracted from each corbicula pollen. Lets also $\mathbf{x}_{lj} = [x_{lj1}, x_{lj2}, ..., x_{ljN}]$ be the feature vector of N elements that identifies uniquely the ROI j of image l. The metric used to measure the similarity between a query case and the mean class prototypes is the Mahalanobis distance to each class i, D_i, defined as:

$$D_i = (\mathbf{x}_{lj} - \mathbf{m}_i)^T \Sigma^{-1} (\mathbf{x}_{lj} - \mathbf{m}_i) \quad i = 1, ..., M \quad (1)$$

Where Σ is the covariance matrix in the training set and $\mathbf{m}_i$ is the mean class prototype for each class i. Prototype class is calculated taking the mean vector on the training set. We assume the same covariance matrix for all classes. The training set is performed using $L - 1$ pollen loads and the test is carried out using the excluded one (leave-one-image-out approach). If this is correctly classified a hit is counted. This is repeated L times, each time excluding a different pollen load. The class of excluded one is derived by majority voting among the set of J ROIs extracted from each pollen load image.

To improve rate classification, an important issue is how to select "good" features. Among several approaches tested for pollen classification [1], such as PCA or Fisher discriminant, the best results were achieved using Floating Search Method (FSM) (Pudil *et al* [18]), where system sensitivity, using the MDC based on Mahalanobis distance, is used as discrimination criterion. The main drawback of normal FSM is that it can be trapped in cycles. In this case, the searching process is stopped.

We tested this approach using MDC to discriminate the five most abundant pollen species in the North-West of Spain, collected in two different places. The classification rate was 73%. Unfortunately, when increasing variability of the problem (through increasing geographical places of collection), system performance decreased until it was no useful for pollen load classification. In order to overcome this problem, in this paper we explore the ability of more sophisticated classifiers to obtain an acceptable classification rate for this problem.

3 Classification

This section briefly describes the techniques that we have used to improve the classification rates obtained by the MDC. Specifically, we have used the K-Nearest Neighbor classifier (KNN), the Multi-Layer Perceptron (MLP) and the Support Vector Machine (SVM), which are three of the most used standard classifiers.

3.1 Classification using KNN

K-Nearest Neighbors ([19], p. 182) is an well-known classification method which assigns each input pattern to the class of the majority of its k nearest neighbours. As with MDC (see section 2), we used the Mahalanobis distance to select the k nearest neighbours, with values for k ranging from 3 to 20. The main drawback of this method is the need for storing the whole training set and for calculating distances to all its patterns.

3.2 Classification using MLP

Multi-Layer Perceptron [20] is a very popular neural network, and it is an interesting alternative to MDC because its ability to classify non-separable pattern sets using its supervised off-line training. We used MLP via the Torch library (`http://www.torch.ch`) [21], a free, modular and efficient machine-learning library writen in C++. We tried MLP using 1, 2 and 3 hidden layers, but we found that 3 layers were necessary to reach an acceptable training error, using 25-35 units per layer, which reflects the high complexity of this data set. We tried sigmoid activation function for the hidden units, but we reached better results using hyperbolic tangent: $out_i = tanh(\sum_{j=1}^{n} w_{ji} x_j)$, being w_{ji} the j-th connection weight and x_j the j-th input to the hidden unit i. Learning rate was 0.01 and number of training iterations was 1000.

3.3 Classification using SVM

Support Vector Machine is a classification algorithm based on optimization theory and initially developped by Vapnik and coworkers [22]. It efficiently trains a linear learning machine in a kernel-induced feature space, while its generalization capability is guaranteed and overtraining is avoided. SVM uses results from the statistical learning and optimization theories in order to maximize its generalization capability for test examples. These properties suggested us that SVM could improve performance on this pollen image classification problem.

The objective of SVM is to learn the function $f : \mathbf{R}^N \rightarrow \{\pm 1\}$ defined by a training set $(\mathbf{x}_i, y_i) \in \mathbf{R}^N \times \{\pm 1\}, i = 1, \ldots L$, where N is the dimension of the input pattern, L is the size of the training set and y_i is the desired output for training pattern $\mathbf{x}_i$. For classification problems, function f is the discriminator for class k, with $f(\mathbf{x}_i) = 1$ for input patterns belonging to class k, and $f(\mathbf{x}_i) = -1$ otherwise.

Learning theory under SVM uses the Cover's theorem [23], which states that, given a linearly nonseparable set of patterns, there is a nonlinear mapping $\Phi(\mathbf{x})$ from $\mathbf{R}^N$ to a higher dimensional space in which the set of patterns is linearly separable. In this high-dimensional space, SVM selects the linear discriminant $f(\mathbf{x}) = sgn(\mathbf{w}\mathbf{x} + b)$ associated with the hyperplane (defined by vector $\mathbf{w}$) which best separates the training samples. Statistical learning theory [24] states that hyperplane with highest margin on its nearest training samples must be selected in order to minimize the error on the training set while preventing overfitting. This is the *maximal margin classifier*, which correctly classifies all the training samples. The margin is inversely proportional to the norm $\|\mathbf{w}\|$ of the vector defining the hyperplane, and therefore SVM learning tries to minimize $\|\mathbf{w}\|^2$ subject to the constraints $y_i f(\Phi(\mathbf{x}_i)) \geq 1, i = 1, \ldots, L$ using the Lagrange multipliers method. However, for the real-life problems, where noise and outliers must be tolerated, the so-called *soft-margin classifier* is used, which allows some training errors by introducing the slack variables $\xi_i = max(0, 1 - y_i f(\mathbf{x}_i))$. For this case, and using the 1-norm ([25], p. 104), the function to minimize is:

$$\frac{1}{2}\|w\|^2 + C \sum_{i=1}^{L} \xi_i, \quad C > 0 \tag{2}$$

Subject to the following constraints:

$$y_i f(\mathbf{x}_i) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, \ldots, L \tag{3}$$

Nonlinear mapping $\Phi(\mathbf{x})$ from the original feature space to the high-dimensional one is introduced by replacing the inner product $\mathbf{w}\mathbf{x}$ in the linear discriminant by a kernel $K(\mathbf{x}, \mathbf{y})$ that matches conditions of Mercer's theorem and, therefore, it can be written as $K(\mathbf{x}, \mathbf{y}) = \Phi(\mathbf{x})\Phi(\mathbf{y})$. The optimization problem can be stated, in the dual form ([25], p. 108), as find $\alpha_i, i = 1, \ldots, L$ (Lagrange multipliers) to maximize:

$$\sum_{i=1}^{L} \alpha_i - \frac{1}{2} \sum_{i=1}^{L} \sum_{j=1}^{L} \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \tag{4}$$

subject to (Karush-Kuhn-Tucker (KKT) conditions):

$$\sum_{i=1}^{L} \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq C, i = 1, \ldots, L \tag{5}$$

Finally, the discriminating function for the SVM can be written as:

$$f(\mathbf{x}) = sgn\left[\sum_{i \in SV} \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b \right] \tag{6}$$

Where SV is the set if *support vectors*, the training vectors for which $\alpha_i \neq 0$. Therefore, just $L_s = card(SV) < L$ training vectors are necessary to compute $f(\mathbf{x})$. Support vectors are the closest patterns to the separating hyperplane, the most difficult to classify, and $\{\Phi(\mathbf{x}_i), i \in SV\}$ defines the optimal hyperplane in the high-dimensional space. Specific techniques have been developped to solve this large-scale quadratic optimization problem, being Sequential Minimal Optimization (SMO) [26] one of the most widely used.

SVM has been applied to many application fields, and also to texture classification, as in [27], where the gray-level of raw pixels were used. However, for the pollen image classification we used the texture feature vectors (see section 2) as inputs to the SVM. We think that this approach is more efficient than use of raw pixels, because dimensions of input vectors are lower. Implementation of SVM was using the Torch library [21]. We used one SVM for each class ("class-against-the-others" approach), being each pattern assigned to the SVM with highest output. SVM parameters were the default ones: $C = 100$ (see eq. 5) and

		Classification method			
Feature vector	N	MDC	KNN	MLP	SVM
FOS	11	47	48	53	53
HC	7	52	61	60	71
NGLDS	5	42	53	57	59
GLRLS	5	40	52	51	51
SF	28	62	65	**70**	75
WE	12	59	63	65	65
WH	12	59	65	61	71
LLT(3)	45	**67**	63	68	74
LLT(5)	125	62	65	68	74
LLT(7)	180	**67**	**67**	69	**76**

N: Number of texture features
SF: union of FOS, HC, NGLDS and GLRLS

Table 1. Percentage of correct pollen loads classification using minimum distance classifier (MDC), KNN, MLP and SVM.

$end\_accuracy = 0.01$ (tolerance level in the evaluation of the KKT conditions). We tried polynomic $K(\mathbf{x}, \mathbf{y}) = (a\mathbf{xy}+b)^c$, dot product $a\mathbf{xy}$ and hyperbolic $tanh(a\mathbf{xy}+b)$ kernels, but the best performance was obtained with gaussian kernel: $K(\mathbf{x}, \mathbf{y}) = exp(-\|\mathbf{x}-\mathbf{y}\|^2/N\sigma^2)$ (division by N makes the range of values for parameter σ independent on N and, therefore, on the feature set). This parameter was tuned to get the best performance, obtaining better rates for values in the range $0.5 \leq \sigma \leq 2$.

4 Results and discussion

Classifier performance was tested using a dataset of 200 pollen loads collected in three distant places (Viana, Lobios and Pontevedra). There were 60 pollen loads of species *Rubus*, *Cytisus* and *Castanea* (20 from each place), and 40 pollen loads of *Quercus* and *Raphanus* (20 from Lobios and 20 from Viana). Afterwards, 5 non-blurring ROIs of 256×256 pixels were extracted from the pollen image, yielding a dataset of 1300 ROIs or patterns. For each pattern, the texture feature vectors mentioned in section 2 were computed. To apply KNN, MLP and SVM, we used the leave-one-image-out approach, with $J = 5$ ROI's per image. Each input feature was normalized by substracting to each component its average over the training set and dividing by its standard deviation. The classification rate obtained for all texture vectors (rows) and classifiers (columns) are shown in table 1.

SVM reached 76% classification rate on LLT(7) set (using gaussian kernel with $\sigma = 2.0$), and it is clearly the best classifier. Its main drawback was the high number L_s of support vectors: for feature set LLT(7), the number of support vectors (including all the classes) was $L_s = 886$, which represents 68% of the whole patterns set ($L = 1300$ patterns). However, this is better than using KNN, which requires storing the whole training set. MLP

Desired	Observed Class				
Class	*R*	*C*	*Q*	*Ra*	*Ca*
R	**47**	8	3	0	2
C	9	**50**	0	1	0
Q	1	1	**24**	10	4
Ra	0	0	5	**34**	1
Ca	3	2	6	3	**46**
Se(%)	78	83	60	85	76
PP(%)	78	82	63	70	86

Table 2. Confusion matrix for the best classification rate, obtained using SVM on LLT(7) (76% performance, see table 1). Se and PP are also shown for each class. Class labels are: R = *Rubus*, C = *Cytisus*, Q = *Quercus*, Ra = *Raphanus*, Ca = *Castanea*.

was only slightly better than KNN or MDC, and the difference among them was not significant. MLP attained its higher rate (70%) on SF set, using 3 hidden layers with 35 units each one. Finally, KNN and MDC rates were similar, reaching both the same higher rate (67%) on LLT(7) set.

Table 2 displays the confusion matrix associated to the best results (SVM on LLT(7)). It also shows sensitivity ($Se = TP/(TP + FN)$) and positive prediction ($PP = TP/(TP + FP)$) for each class (TP = True Positives; FN = False Negatives; FP = False Positives): Se and PP reached 85% for some classes and they were always over 60%.

5 Conclusion and future work

We proposed a methodology to identify honeybee pollen. It consists of three steps: ROIs extraction, texture feature generation and classification. This paper is mainly focused on the improvement of the classification stage. The combination of SVM classifier and LLT(7) texture vector achieved the best performance (76%) to discriminate among the five most abundant plant species from three geographical places in North-West of Spain. System behaviour is quite uniform for every class. Although these results are very promising, additional work with a wider set of plant species and geographical location of the hives will be needed in order to apply the proposed methodology to on-line pollen industrial analysis.

Future work also includes comparison with other classification techniques, applied on these texture feature vectors or on the gray-level of the raw pixels. However, results suggest that feature sets are hard enough to classify using a single classifier, and therefore we think that the use of combination schemes may improve significantly the performance rate obtained by the individual classifiers.

Acknowledgment

This investigation was supported by the Spanish CICyT project TIC2000-0873-C02 and by the Xunta de Galicia (regional government) project *Study of dairy selection of polineferas plant in apis mellifera and influence of protein content of pollenkit and texture of polen loads* (PGIDT01PXI38305PR).

References

[1] P. Carrión, E. Cernadas, P. Sá-Otero, and E. Díaz-Losada. Could the Pollen Origin be Determined using Computer Vision? An Experimental Study. In *IASTED International Conference on Visualization, Imaging, and Image Processing*, pages 74–79, 2002.

[2] P. Carrión, E. Cernadas, J.F. Gálvez, and E. Díaz-Losada. Determine the composition of honeybee pollen by texture classification. In *Lecture Notes on Computer Science 2652, Proceedings of the 1st Iberian Conference on Pattern Recognition and Image Analysis (IBPRIA 2003)*, pages 158–167, 2003.

[3] E. Díaz Losada, E. Fernández Gómez, C. Álvarez Carro, and M.P. Saa Otero. Aportación al conocimiento del origen floral y composición química del polen apícola de Galicia (Spain). *Boletín de la Real Sociedad Española de Historia Natural*, 92(1-4):195–202, 1996.

[4] P. Sá-Otero, P. Canal-Camba, and E. Díaz-Losada. Initial data on the specific heterogeneity foundin the bee pollen loads produced in the "Baixa-Limia-Serra do Xurés" Nature Park, 2002.

[5] E. Díaz Losada, A. V. González Porto, and M.P. Saa Otero. Étude de la culeur du pollen apicole recueilli pa Apis mellifer L. en nord-ouest d'Espagne. (Galice). *Acta Botanica Gallica*, 145(1):39–48, 1998.

[6] D. Hodges. The pollen loads of the honeybee. In *Bee Research Association*, page 48. 1984.

[7] M.I. Hidalgo and M. L. Bootello. About some physical characteristics of the pollen loads collected by Apis mellifera L. *Apicoltura*, 6:179–191, 1990.

[8] P. Li and J.R. Flenley. Pollen texture identification using neural networks. *Grana*, pages 59–64, 1999.

[9] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural Features for Image Classification. *IEEE Trans. on Man and Cibernetics*, 3(6):610–621, 1973.

[10] R.M. Haralick and L. Shapiro. *Computer and Robot Vision. Volumen I*. Addison-Wesley, 1992.

[11] M. Sonka, V. Hlavac, and R. Boyle. *Image Processing, Analysis, and Machine Vision*. International Thomsom Publishing (ITP), 1999.

[12] L. H. Siew, R. M. Hodgson, and E. J. Wood. Texture Measures for Carpet Wear Assessment. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 10(1):92–104, 1988.

[13] S. Theodoridis and K. Koutroumbas. *Pattern recognition*. Academic Press, 1999.

[14] S. Mallat. A threary for multiresolution signal descomposition: the wavelet representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(7):674–693, 1989.

[15] A. Laine and J. Fan. Texture classification by Wavelet Packet signatures. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 15(11):1186–1191, 1993.

[16] I. Daubechies. Ortonormal bases of compactly supported wavelets. *Commun. Pure Appl. Math.*, XLI:909–996, 1988.

[17] K. I. Laws. Rapid texture identification: image processing for missile guidance. In *SPIE*, volume 238, pages 376–380, 1980.

[18] P. Pudil, J. Novovicova, and J. Kittler. Floating search methods in feature selection. *Pattern Recognition Letters*, 15:1119–1125, 1994.

[19] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. Jonh Wiley Sons, 2001.

[20] S. Haykin. *Neural Networks*. Prentice-Hall, 1999.

[21] R. Collobert and S. Bengio. SVMTorch: Support Vector Machines for large-scale regression problems. *Journal of Machine Learning Research*, 1:143–160, 2001.

[22] C. Cortes and V. Vapnik. Support Vector Networks. *Machine Learning*, (20):273–297, 1995.

[23] T.M. Cover. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Trans. on Electronics Computers*, 14:326–334, 1965.

[24] V. Vapnik. *The nature of statistical learning theory*. Springer-Verlag, 1995.

[25] N. Cristianini and J. Shawe-Taylor. *An introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge University Press, 2000.

[26] J. Platt. Fast training of support vector machines using sequential minimal optimization. In B. Schölkopf, C.J.C. Burges, and A.J. Smola, editors, *Advances in Kernel Methods - Support Vector Learning*, pages 185–208. MIT Press, 1999.

[27] K.I. Kim, K. Jung, S.H. Park, and H.J. Kim. Support Vector Machines for texture classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(11):1542–1550, 2003.

View publication stats