

## Research Article

# Image Processing Method for Automatic Discrimination of Hoverfly Species

**Vladimir Crnojević,<sup>1</sup> Marko Panić,<sup>1</sup> Branko Brkljač,<sup>1</sup> Dubravko Čulibrk,<sup>2</sup>  
Jelena Ačanski,<sup>3</sup> and Ante Vujić<sup>3</sup>**

<sup>1</sup>Department of Power, Electronic and Telecommunication Engineering, University of Novi Sad, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia

<sup>2</sup>Department of Information Engineering and Computer Science, University of Trento, Via Sommarive 5, Povo, 38123 Trentino, Italy

<sup>3</sup>Department of Biology and Ecology, University of Novi Sad, Trg Dositeja Obradovića 2, 21000 Novi Sad, Serbia

Correspondence should be addressed to Vladimir Crnojević; [crnojevic@uns.ac.rs](mailto:crnojevic@uns.ac.rs)

Received 27 June 2014; Accepted 17 December 2014; Published 30 December 2014

Academic Editor: Andrzej Swierniak

Copyright © 2014 Vladimir Crnojević et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

An approach to automatic hoverfly species discrimination based on detection and extraction of vein junctions in wing venation patterns of insects is presented in the paper. The dataset used in our experiments consists of high resolution microscopic wing images of several hoverfly species collected over a relatively long period of time at different geographic locations. Junctions are detected using the combination of the well known HOG (histograms of oriented gradients) and the robust version of recently proposed CLBP (complete local binary pattern). These features are used to train an SVM classifier to detect junctions in wing images. Once the junctions are identified they are used to extract statistics characterizing the constellations of these points. Such simple features can be used to automatically discriminate four selected hoverfly species with polynomial kernel SVM and achieve high classification accuracy.

## 1. Introduction

Classification, measurement, and monitoring of insects form an important part of many biodiversity and evolutionary scientific studies [1–3]. Their aim is usually to identify presence and variation of some characteristic insect or its properties that could be used as a starting point for further analyses. The technical problem that researchers are facing is a very large number of species, their variety, and a shortage of available experts that are able to categorize and examine specimens in the field. Due to these circumstances, there is a constant need for automation and speed up of this time consuming process. Application of computer vision and its methods provides accurate and relatively inexpensive solutions when applicable, as in the case of different flying insects [1, 2, 4, 5]. Wings of flying insects are one of the most frequently considered discriminating characteristics [4] and can be used standalone as a key characteristic for

their classification [2]. Unlike some other body parts, wings are also particularly suitable for automatic processing [6]. The processing can be aimed at species identification and classification or form the basis for further morphometric analyses once the classification to specific taxonomy is done.

Discriminative information that allows flying insects classification may be contained in wing shape [7], but in the most cases it is contained in the relative positions of vein junctions inside the wing that primarily define unique wing venation patterns [1, 2, 4–6]. Wing venation patterns are the result of specific evolutionary adaptations over a long period of time and are influenced by many different factors [8]. As such, they are relatively stable and can successfully describe and represent small differences between very similar species and taxons, which is not always possible using only shape of the insect's wing. Another useful property of venation patterns is that they are not significantly affected by the current living conditions, present in some specific natural

environment, in comparison to some other wing properties such as colour or pigmentation. This makes them a good choice for reliable and robust species discrimination and measurement. The advantage of using venation patterns is also that patterns of previously collected wing specimens do not change with the passing of time, as some other wing features, so they are suitable for later, off-field analyses.

Discrimination of species in the past was based on descriptive methods that proved to be insufficient and were replaced by morphometric methods [6]. These methods rely on geometric measures like angles and distances in the case of standard morphometry or coordinates of key points called landmarks, which can be also used for computing angles and distances, in the case of more recent geometric morphometrics. In the wing-based discrimination each landmark-point represents a unique vein junction, whose expected position on the wing is predefined and which needs to be located in the wing before discrimination. Manually determined landmarks require skilled operator and are prone to errors, so automatic detection of landmark-points is always preferred.

Some systems for automatic classification of insects are designed to perform recognition tasks in uncontrolled environments with variability in position and orientation of objects [3], while others are designed to operate under controlled working conditions [2, 6].

Methods for automatic detection of vein junctions in wing venation of insects usually consist of several preprocessing steps which include image registration, wing segmentation, noise removal, and contrast enhancement. In order to extract lines that define wing venation pattern, in the next stage are often applied edge detection, adaptive thresholding, morphological filtering, skeleton extraction, pruning, and interpolation, usually in the precisely given order. In this way the locations of landmark-points corresponding to vein junctions are found [1, 4] or moreover a polynomial model of the whole venation pattern is made on the basis of line junctions and intersections [1, 2, 5]. This may be easier to achieve if the light source is precisely aligned during the image acquisition phase, so that it produces uniform background [4], or when it is allowed to use additional colour information, as in the case of leaf venation patterns [9], but it is not always possible. Some of the reasons can be noisy and damaged images due to the dust, pigmentation, different wing sizes, image acquisition, or bad specimen handling.

The Syrphidae family of hoverflies are of special interest due to a number of important roles they have in pollination, indication of biodiversity level, and evolution research. The paper presents an approach to their automatic classification based on the method for automatic detection of landmark-points in wing venation of flying insects, which is utilizing supervised learning on a dataset of vein junction images extracted by the human experts from the real-world images of specimens' wings.

Section 2 provides an overview of the dataset and the proposed landmark-points detection method. The proposed classification methodology based on automatically detected landmark-points is presented in Section 3, while results are given in Section 4. Finally, conclusions are drawn in Section 5.

TABLE 1: Number of wing images per each class (species) in created dataset.

<i>Chrysotoxum</i>			<i>Melanostoma</i>		
<i>festivum</i>	<i>vernale</i>	other	<i>mellinum</i>	<i>scalare</i>	other
248	154	22	267	105	72

## 2. Landmark-Points Detection

The proposed method for landmark-points (vein junctions) detection consists of computing specific, window based features [10–13], which describe presence of textures and edges in window and subsequent classification of these windows as junctions (i.e., positives) or not-junctions (i.e., negatives) using a junctions detector previously obtained by some supervised learning technique.

**2.1. Wing Images Dataset.** The set of wing images used in the presented study consists of high-resolution microscopic wing images of several hoverfly species. There is a total of 868 wing images of eleven selected hoverfly species from two different genera, *Chrysotoxum* and *Melanostoma*, Table 1.

The wings have been collected from many different geographic locations during a relatively long period of time of more than two decades. Wing images are obtained from the wing specimens mounted in the glass microscopic slides by a microscopic device equipped with a digital camera with image resolution of  $2880 \times 1550$  pixels and are stored in TIFF format. Each image is uniquely numbered and associated with the taxonomy group it belongs to. Association of each wing with a particular species is based on the classification of the insect at the time when it was collected and before the wings were detached. This classification was done after examination by a skilled expert. The images were acquired later by biologists under relatively uncontrolled conditions of nonuniform background illumination and variable scene configuration without previous camera calibration. In that sense, obtained images are not particularly suitable for exact measurements.

Other shortcomings of the samples in the dataset are result of variable wing specimens quality, damaged or badly mounted wings, existence of artifacts, variable wing positions during acquisition, and dust. In order to overcome these limitations and make images amenable to automatic hoverfly species discrimination, they were first preprocessed. The preprocessing consisted of image rotation to a unified horizontal position, wing cropping, and scaling. Cropping eliminates unnecessary background containing artifacts, while aspect ratio-preserving image scaling enables overcoming the problem of variable size among the wings of the same species. After computing average width and average height of all cropped images, they were interpolated to the same width of 1680 pixels using bicubic interpolation. Wing images obtained in this way form the final wing images dataset used for sliding-window detector training, its performance evaluation, and subsequent hoverfly species discrimination using the trained detector. Number of images per species is not uniform, Table 1, so only four species with significant

number of images are selected for later experimental evaluation of the proposed method for species discrimination (classification) based on detected landmark-points. These four species include 774 images from the two different genera of the Syrphidae family and are illustrated in Figure 1.

**2.2. Training/Test Set.** In order to analyze the applicability and the efficiency of the proposed methodology, when it comes to the problem of landmark-points detection, special vein junctions training/test set was created from the collected images in the wing images dataset described in Section 2.1. It consists of the characteristic wing regions (patches) that correspond to vein junctions in the wing venation pattern, that is, positives, and randomly selected patches without vein junctions, that is, negatives, which still can contain parts of wing venation. From each wing image 18 uniquely numbered positive patches, whose centers are shown as red dots in Figure 1, were manually extracted and saved using specially created user application. In the case of severely damaged wings which were damaged or missing landmarks, the corresponding patches were not selected. As a result, the training/test set with 15590 positives and 22466 manually selected negatives was created by using all available hoverfly wing images, where the total number of positives was slightly smaller than expected due to the mentioned reasons. The created set was then used for the detailed study of the effects of various implementation choices on the detector's performance, as described in Section 2.3.

**2.3. Landmark-Points Detector.** Discriminative descriptors of vein junctions that are used in the proposed landmark-points sliding-window detector are HOG (histogram of oriented gradients) [12] and the robust version of CLBP (complete local binary pattern) [14], proposed in [11]. HOG and LBP operators were first presented in [10, 12]. In order to determine and compare performance of different detectors based on these image descriptors and evaluate the impact of different sets of pertinent parameters, descriptors were considered separately and combined, in the same way as described in [13, 15].

Since the wing color is varying characteristic in the given classification problem due to many factors and consequently cannot be reliably used for discrimination among different species, it was decided that all descriptors should be based only on features derived from grayscale images. Therefore, the first step in all computations is the conversion of input color images into their grayscale versions by standard conversion of RGB color space into HSI and selection of computed intensity channel as the final grayscale image.

The CLBP is one of the numerous improvements of LBP descriptor which have been proposed in recent years. The robust version of CLBP, that is, RCLBP, is just utilization of the idea implemented in the conventional LBP [10] that was suggested recently by one of its original authors [11].

When it comes to HOG, a feature vector consisting of a number of discrete histograms of image gradient orientation [12] is employed. Discrete histograms are computed over small rectangular spatial regions in image, called cells, which

are obtained by subdivision of the main feature extraction window. The first step in the histogram computation is gradient discretization, done for each pixel by projection of pixel gradient onto the two closest allowed bins, that is, the two closest of several predefined uniformly spaced discrete directions. Before finally computing discrete histograms of gradient orientation for each cell, 2-D CTI (convoluted trilinear interpolation) filtering described in [13] is additionally applied. The CTI filtering is used to smooth results of gradient discretization from the previous step, and it is achieved through interpolation of computed discretized gradient values between spatially adjacent pixels. The filtering is performed by convolution with Gaussian-like kernel of each of gradient orientation planes, that is, each of the created gradient images corresponding to different predefined possible orientations of image gradient. Thus, instead of only two nonzero values representing the discretized image gradient at some pixel (spatial position), after filtering the image gradient at each pixel is represented as the sum of several components with different magnitude corresponding to all predefined discrete directions. As suggested in [12], before the construction of the final feature vector, values of discrete histogram are locally normalized by the normalization procedure which includes accumulating histograms over somewhat larger overlapping spatial regions, called blocks, and using the  $L_2$  vector norm. These values, representing normalized values of several spatially adjacent discrete histograms which belong to the cells inside the same block, are then concatenated block by block to form the final HOG feature vector per window.

The HOG vector length and the dimensionality of the corresponding feature space depend on the choice of parameters that define window, cell and block size, extent of block overlapping, and a number of allowed discrete histogram values (orientation bins). We used nine bins evenly spaced over  $0^\circ$ – $180^\circ$ ,  $64 \times 64$  pixels for detection window, blocks containing  $2 \times 2$  cells, and the one cell wide blocks' overlapping width. In order to measure detector's performance different cell sizes (8, 16, and 32 pixels) were used. As a result, depending on the cell size, possible dimensions of used HOG feature vectors are: 1764 (hog8), 324 (hog16), and 36 (hog32). Extraction of HOG features for different cell sizes on the example of one of the vein junction images from the training/test set is illustrated in Figure 2, along the main phases in the computation of HOG feature that are also shown.

The CLBP descriptor integrates information about the sign and magnitude of difference  $d$  computed between central pixel  $p_c$  and pixel  $p_i$  in its predefined neighborhood in some grayscale image. On the other hand, the conventional LBP utilizes only information about the sign of  $d$ , as can be seen in

$$\text{LBP}_{n,r} = \sum_{i=0}^{n-1} s(p_c - p_i) 2^i, \quad s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0, \end{cases} \quad (1)$$

where  $n$  denotes the number of neighbouring pixels at the radius  $r$  that are compared with the central pixel  $p_c$ .

As the main parameters of CLBP descriptor were used the circular neighbourhood geometry with eight surrounding

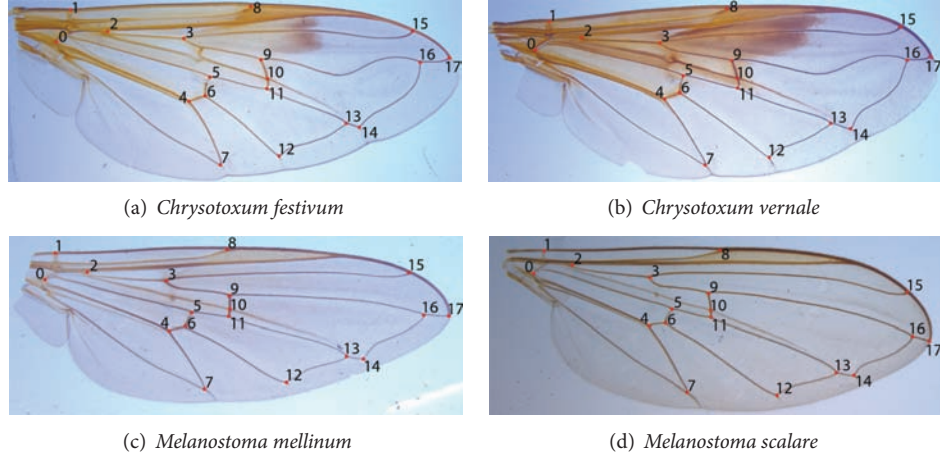


FIGURE 1: Images of the wing of four selected hoverfly species belonging to two different genera. Numbered red dots represent positions of manually marked predefined landmark-points in wing venation which can be used for species discrimination.

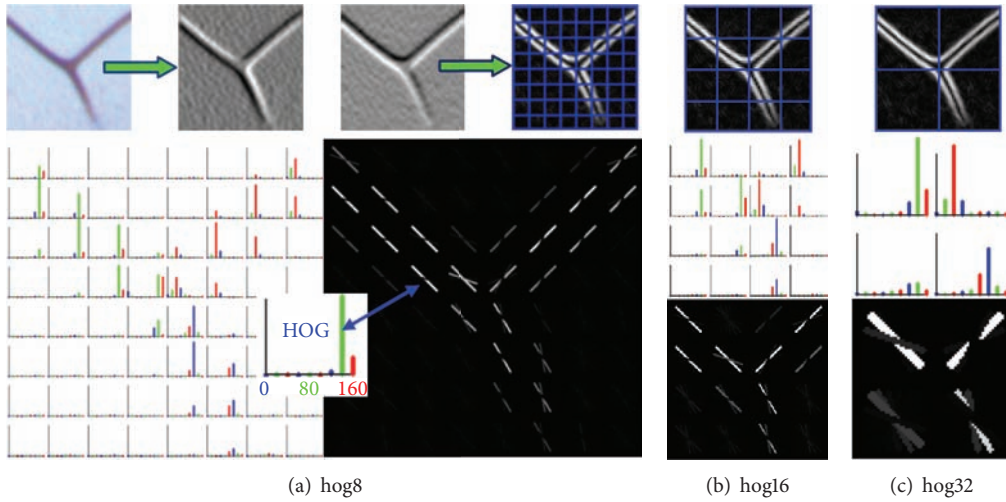


FIGURE 2: Illustration of HOG features on the example of one of the images from the training/test set, Section 2.2. In (a)–(c) are visualized HOG features corresponding to different values of cell size:  $8 \times 8$  pixels (a),  $16 \times 16$  pixels (b), and  $32 \times 32$  pixels (c). In each case (a)–(c), for visualization are used two images: one depicting discrete histograms of gradient orientation on the level of single cells in the image (images consisting of small histograms with blue, green, and red bars) and the second image which represents its grayscale counterpart (on the level of each cell, i.e., histogram, are drawn lines which correspond to particular orientation of image gradient, while intensity of lines, i.e., their normalized grayscale value, corresponds to the magnitude of the image gradient in the given direction after gradient discretization). Blue arrow in (a) indicates equivalence between magnified histogram and its grayscale counterpart, with lines associated with the bars in the histogram. Different color of bars in the histograms is used to ease distinction between allowed discrete orientations, 9 values in the range  $0-180^\circ$ . Green arrows in the upper part of (a) describe the process of image gradient computation using Sobel filter, while the blue grid depicts the cell size, also overlayed over gradient images in (b) and (c).

pixels at unit distance, and the same window size of  $64 \times 64$  pixels, as in the case of HOG. The value of each difference  $d_i = p_c - p_i$  can be decomposed into two components  $s_i$  and  $m_i$ , which represent sign and magnitude of difference, respectively, as given in (2). These components are used for the construction of two types of CLBP codes which describe local pixel intensity variations. The information about the sign of difference is used for the construction of the CLBP\_S code, in a similar way like in the case of conventional LBP code, while the information about the difference's magnitude is used for the construction of CLBP\_M code, which is

introduced in order to provide additional discriminative power. Consider

$$d_i = s_i * m_i, \quad s_i = \text{sign}(d_i), \quad m_i = |d_i|. \quad (2)$$

Definition of CLBP\_S code is the same as the one of LBP in (1), while in the definition of CLBP\_M code in (3) an additional threshold  $c$ , which is determined adaptively, is introduced:

$$\text{CLBP\_M}_{n,r} = \sum_{i=0}^{n-1} t(m_i, c) 2^i, \quad t(x, c) = \begin{cases} 1, & x \geq c \\ 0, & x < c. \end{cases} \quad (3)$$



The adaptive threshold  $c$  used for obtaining CLBP\_M codes, (3), is computed as the mean value of all  $m_i$  on the level of whole image. However, in our situation this computation is restricted to the regions of  $32 \times 32$  pixels inside the patches (windows) of  $64 \times 64$  pixels.

There also exists the third type of CLBP code under the name CLBP\_C which refers to the intensity of the central pixel  $p_c$ , but due to the issue we consider here, in order to recognize junctions in the vein structure and distinguish them from the rest of the wing, we are more interested in the local variation of pixel intensity values represented by CLBP\_S and CLBP\_M codes. Therefore, CLBP\_C was not used, but for the sake of completeness we still give its definition in the following equation, where  $c_l$  is the mean value of pixel intensities over the observed region:

$$\text{CLBP}_{C_{nr}} = t(p_c, c_l). \quad (4)$$

Before computation of histograms of CLBP\_S and CLBP\_M codes in the next step, a small but very important additional code variation is made in both cases. Since we use the circular pixel surrounding of  $n = 8$  pixels at radius  $r = 1$ , the following two substrings “010” and “101” in the binary representation of codes are substituted with “000” and “111,” respectively. The same code variation was performed in [11], but it was applied to conventional LBP instead of CLBP\_S and CLBP\_M. Under the assumption that these two substrings are most likely caused by noise, this variation removes the noise from features. Also, it substantially reduces the number of bins in code histograms from 255 to 46 values. All 46 binary codes obtained in this way are termed “uniform,” which means that 42 among them are characterized by two 0-1 or 1-0 transitions, that is, their uniformity measure is 2, while the remaining four are with four transitions; that is, their uniformity measure is 4. This uniformity measure was defined in [10] and expresses the fact that some local binary patterns describe fundamental properties of texture which makes them more important than some other. Hence, one of the reasons for choosing RCLBP instead of uniform LBP is that it is based on higher values of uniformity measure for some codes. The common property of proposed RCLBP codes is the uniform circular structure with very few transitions, which makes them suitable to faithfully describe expected form of edges in the local region determined by the parameters  $n$  and  $r$ . They will be denoted by RCLBP\_S and RCLBP\_M in the rest of the paper.

The described RCLBP\_S and RCLBP\_M codes are graphically illustrated in Figure 3, where differences between these two types of CLBP codes utilized in the vein junctions detection are shown. Figure 3 contains examples of images (patches) of the vein junctions from the training/test set described in Section 2.2, which correspond to the rectangular image regions around different landmark-points in Figure 1.

The values of the formed histograms of CLBP\_S and CLBP\_M codes are at the end normalized with the so called min-max norm to the range between 0 and 1 and after that concatenated to form RCLBP (robust complete local binary pattern) feature vector per each region inside the window. As a result the final RCLBP feature vector per each  $64 \times 64$

window is composed of 492 features, obtained from four nonoverlapping regions (blocks) with the size of  $32 \times 32$  pixels.

The combined feature vectors are formed by appending described RCLBP feature vector at the end of the corresponding HOG feature vector. Both HOG and RCLBP feature vectors were used separately and in all combinations in order to measure their window based performance on the training/test set using the same classifier. Performance comparison was made using support vector machine (SVM) classifier that has good generalization properties and ability to cope with small number of samples in the case of high feature space dimensionality [16].

Feature extraction was implemented in C++ using OpenCV library [17], on computer with Intel i5 CPU 3.20 GHz and 8 GB of RAM, without any parallelization, special adaptations, or GPU acceleration. Computation time of HOG feature is determined by the chosen cell size, so as expected hog8 has the highest computation time, which per window is approximately 10% higher than in the case of hog32 on same configuration, while CLBP feature has approximately 5% higher computation time than hog8.

Detector's performance testing was done in the machine-learning package Weka [18] using LibSVM library [19], which contains an implementation of SVM classifier. It consisted of analyzing accuracy of the same classifier with different types of window based features, whereas the classifier used SVM with polynomial kernel defined in (5) and the following set of parameters:  $C = 1$ ;  $\gamma = 1$ ;  $c_0 = 0$ ; and  $d = 3$ . In all cases, classifier's performance was measured using 10-fold cross-validation on the training/test set. The cross-validated window level results in terms of the true positives and the false positives rates are shown in Figure 4. Consider

$$K(u, v) = \gamma (u \cdot v + c_0)^d. \quad (5)$$

The usage of HOG and RCLBP features as descriptors of vein junctions shows acceptable results with miss rate smaller than 1% in most cases, Figure 4. When used separately, RCLBP features give better result than HOG features. The HOG features with the cell size of 32 pixels are too coarse to properly describe vein junction in the middle of the window, because in this case window contains only 4 cells, as shown in Figure 2(c). On the other hand the smallest cell size of 8 pixels, illustrated in Figure 2(a), gave the best result among all HOG features. As can be seen from Figure 4, combined HOG and RCLBP features have the best performance but are more memory and time demanding during the training phase due to larger dimensionality of their feature space. Nevertheless, presented results were motivation for the construction of vein junctions sliding-window detector.

As a result, combined HOG-RCLBP features with the cell size of 16 pixels were selected as the best choice for the automatic hoverfly species discrimination based on the sliding-window landmark-points detection. Computation time per image for the chosen set of features and sliding-window step size was approximately 57 s on the given computer configuration.

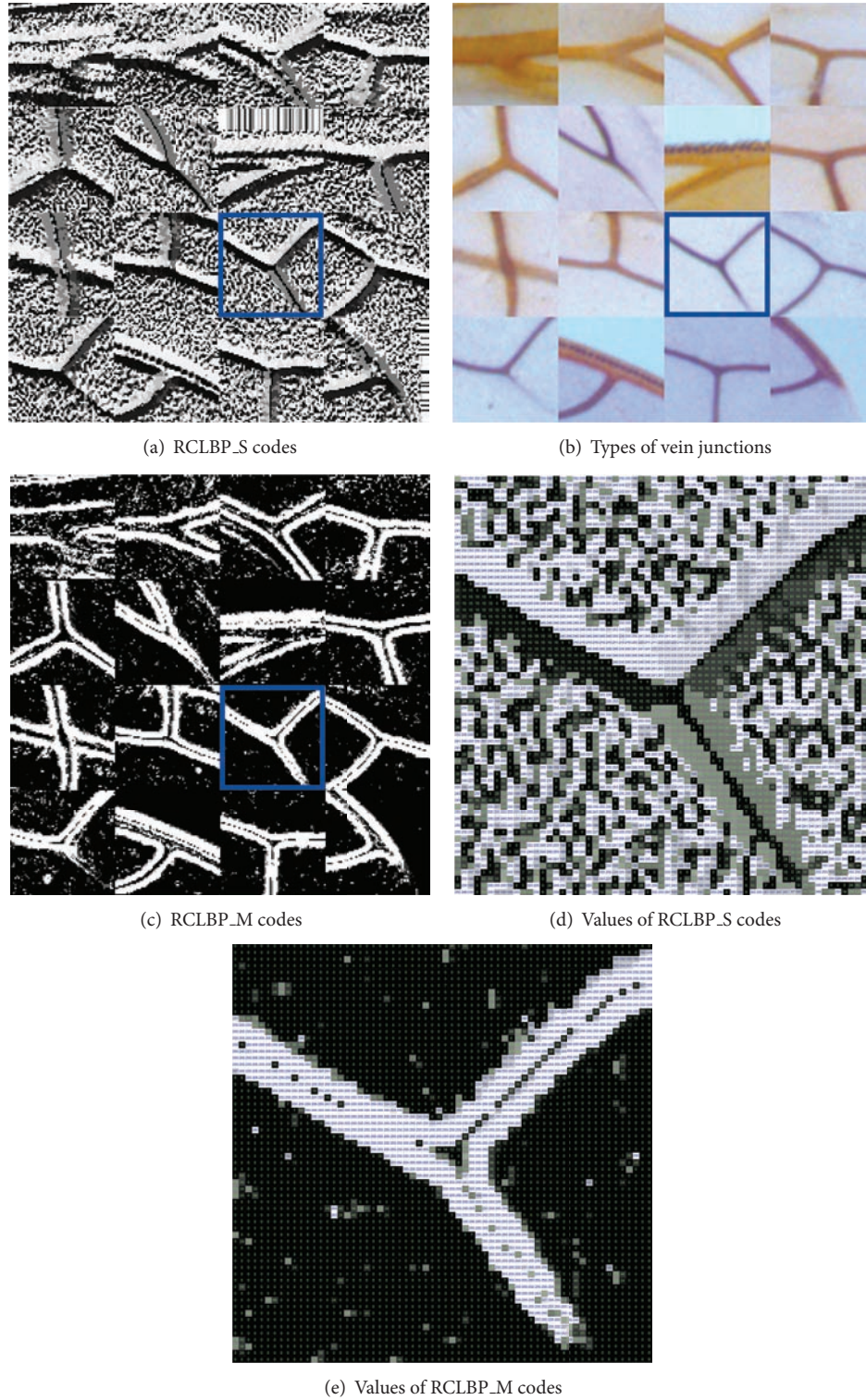


FIGURE 3: Visualization of RCLBP\_S and RCLBP\_M codes. In the middle, (b), is an artificial mosaic formed by images which represent different types of vein junctions in wings of hoverflies. Positive class of training/test set used for supervised learning of vein junctions detector, Section 2.2, consists of images like those shown in (b). Mosaic images on the left (a) and on the right (c) sides of (b) represent generated grayscale visualizations of the computed RCLBP\_S and RCLBP\_M codes, respectively. Values of computed 8-bit binary codes are in the  $[0, 255]$  grayscale range, which makes them suitable for direct visualization. In (d) and (e) are magnified details from (a) and (b). These details represent values of the corresponding codes, computed for each pixel of vein junction outlined with blue frame in (b). Additionally, in order to emphasize difference between values of binary codes corresponding to junctions and those corresponding to the surrounding background, in (d)-(e) with green and blue color are also written their integer values, zoom figure.

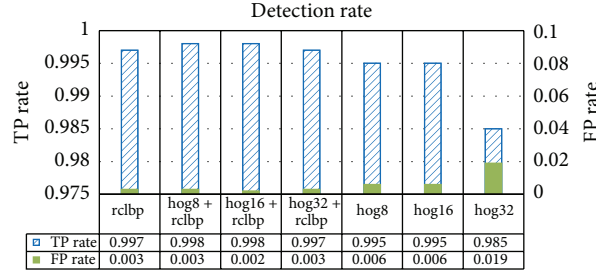


FIGURE 4: Performance comparison of detectors with different input features using SVM with polynomial kernel and true positives and false positives rates as complementary performance measures.

### 3. Species Discrimination

Automatic hoverfly species discrimination was limited to four selected hoverfly species from the wing images dataset which have significant number of instances, Table 1. The discrimination is based on the output of the functional block performing automatic detection of vein junctions in the wing image. The vein junctions detection is done using a sliding window that is densely searching through the image using the proposed sliding-window detector described in Section 2. For better performance, an optimally trained SVM classifier with the polynomial kernel implemented in [17] is used. Its optimal parameters were determined through exponential parameter grid search using 10-fold cross-validation across the whole training/test set described in Section 2.2, as searching criteria used the minimum of false positives rate. Once the optimal values were determined, that is,  $C = 12.5$  and  $\gamma = 0.03375$ , where other kernel parameters  $c_0 = 0$  and  $d = 3$  were set in advance, the whole training/test set was used once again in order to train the final detector.

The constructed detector scans the wing image and returns discrete responses indicating whether a vein junction is present or not in the current window. The same size of the sliding window step is used for both image dimensions. In the case of detection, center coordinates of the current window which correspond to possible vein junction are saved together with classifier's soft response value. This value describes how far from the separating hyperplane defined by support vectors is the current feature vector or how trustworthy the detector's decision is. This soft information is later used to improve the precision of final landmark-points detections.

Due to the multiple detections of the same vein joint and possible false detections, additional postprocessing of obtained detections is needed at the end of the sliding-window search, that is, once the detector finishes scanning through the image. In Figure 5 with the red dots are shown four examples of detection results, which include multiple detections of the same vein joint, as well as the false detections.

The postprocessing consists of clustering of detected points, where the clusters which have less than 3 points (detections) associated with them at the end of clustering are discarded from further consideration in order to eliminate possible false detections, since it is expected that in the region where the true vein junction exists there will be

more than 3 detections due to the dense scanning, that is, small window step size. The centroids are computed by using previously obtained detector's soft response values which are normalized at the level of each cluster using  $L_1$  norm so that they can correspond to probability of correct vein junction detection and consequently can be used as appropriate weight coefficients of multiple detections inside the cluster.

The clustering procedure is based on an iterative algorithm that in each iteration searches through the detections that have not yet been associated with any existing cluster until all detections are assigned to some cluster. It uses a distance criterion based on the sliding window step size and initializes clusters with the existing unassociated detections. Once the clustering is completed, as it has been mentioned, clusters with 3 or less points are discarded and the centroids of the remaining clusters are determined as weighted average of all detections inside the cluster. Obtained centroids represent possible vein junctions that have been found in an image by the sliding window detector.

Even though we tried to remove false detection by eliminating clusters with less than 3 detections, there is no guarantee that the remaining clusters contain all expected landmark-points or that there are no false detections. There are multiple reasons for this, damaged wings, presence of artifacts, dust, and wing specimens with missing parts. Consequently, fixed length feature vectors, which would be based on obtained automatic detections, are not an appropriate choice for image classification. Therefore, we propose generalized approach that is not sensitive to the number of detected landmark-points.

For the purpose of the characterization of points constellation in the wing, we computed convex hull of obtained centroids. Let us denote the set of cluster centroids with  $S$ . Convex hull of set  $S$  denoted by  $\text{conv}S$  is the set of all convex combinations of points in  $S$ :

$$\text{conv}S = \{\theta_1 k_1 + \dots + \theta_n k_n \mid k_i \in S, \theta_i \geq 0, i = 1, \dots, n, \theta_1 + \dots + \theta_n = 1\}, \quad (6)$$

where  $k_i$  represents centroid of cluster  $i$ . Because  $\text{conv}S$  is the smallest convex set that contains  $S$ , we are particularly interested in those points from the set  $S$  which belong to the boundary of  $\text{conv}S$ . The boundary point of  $\text{conv}S$ ,  $x \in \text{bdconv}S$ , satisfies the following property: for all  $\epsilon > 0$ , there exist  $k \in \text{conv}S$  and  $z \notin \text{conv}S$  with  $\|k - x\|_2 \leq \epsilon$ ,  $\|z - x\|_2 \leq \epsilon$ ;



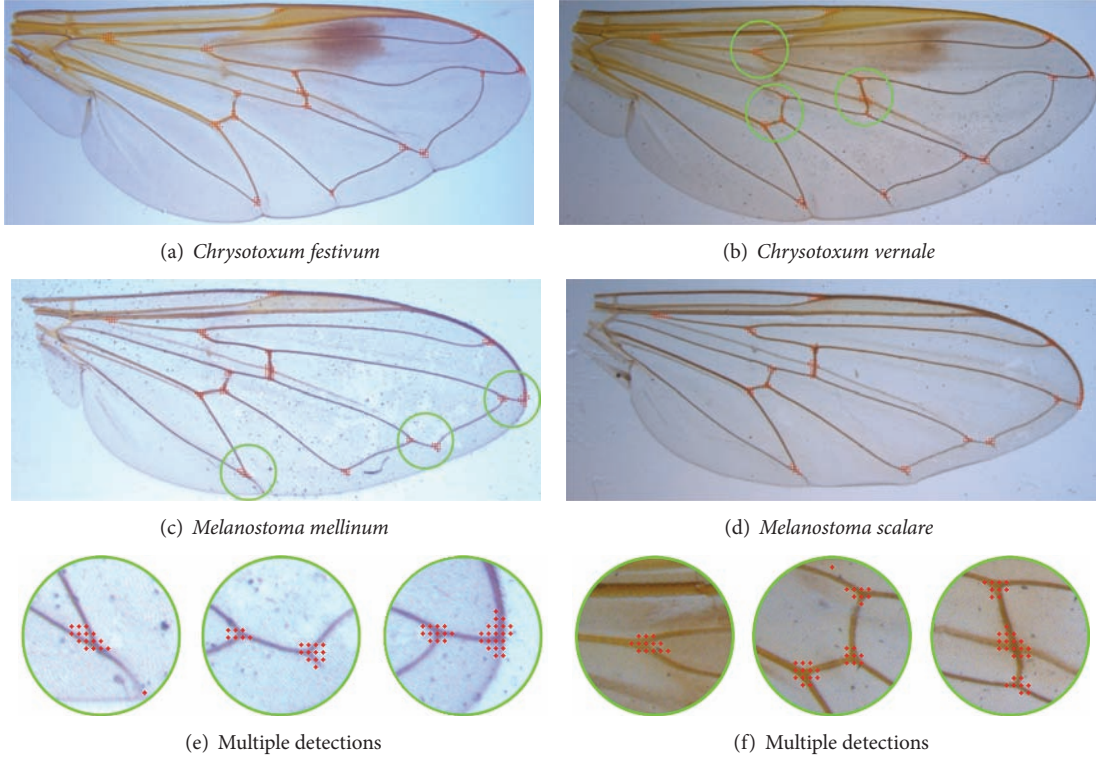


FIGURE 5: An example of automatic detection of vein junctions in wing images of four selected hoverfly species: (a)–(d), detections are shown as red dots. Since detector is based on sliding-window, each vein junction is detected several times in the same image. Based on detector's confidence level, different significance is given to each of the multiple detections corresponding to the same junction. Several magnified details from (b)–(c) are illustrated in (f) and (e), respectively.

that is, there exist arbitrarily close points in **convS** and also arbitrarily close points that are not in **convS**.

Hence, after determining boundary points of **convS**, we compute the following measures which characterize the convex hull: the centroid of points in **bdconv** which belong to  $S$ , median of distances between all boundary points of **convS** which belong to  $S$  and their centroid, root mean square difference of previously described distances from their median, perimeter of contour which envelops **convS**, and area of **convS**. Their summary is given in Table 2.

A common property of these features is that they, as descriptive statistics, do not depend significantly on the number of landmark-points used for their computation and are also rotation-invariant. Under the assumption that they are discriminative enough to distinguish different hoverfly species and do not change significantly inside the same species, they are used as elements of the feature vector that describes particular wing image.

The described procedure for characterization of convex hull of detections' centroids is repeated consecutively 3 times for each image, with elimination of detections' centroids belonging to the boundary of current convex hull after each characterization step. This means that, after computing convex hull in each step, centroids on the boundary of convex hull are removed and the same procedure is repeated again. At the end, the result is that each wing image is characterized by 18 features (values) which describe the properties of 3

TABLE 2: Summary of convex hull based features.

$F1$	$x_c$
$F2$	$y_c$
$F3$	$\text{median} \left( \sqrt{(x_b - x_c)^2 + (y_b - y_c)^2}, \right.$ $\left. (x_b, y_b) \in A = \{\mathbf{bdconvS} \cap S\} \right)$
$F4$	$\sqrt{\frac{1}{ A } \sum_b \left( \sqrt{(x_b - x_c)^2 + (y_b - y_c)^2} - F3 \right)^2}, (x_b, y_b) \in A$
$F5$	$\text{Perimeter}(\mathbf{bdconvS})$
$F6$	$\text{Area}(\mathbf{convS})$

convex hulls constructed per each image, while each of them is characterized by 6 previously described features, Table 2. As an illustration, in Figure 6 are shown examples of detections' centroids and created convex hulls in the case of four selected hoverfly species.

Since during evaluation some of the predefined landmark-points (landmarks numbered 0 and 1 in Figure 1) proved to be not descriptive enough to properly and reliably describe wing images of different species, they were discarded from further analysis, although at first they were marked as landmarks. The reason is their greater variability due to specific position in the wing, which in combination with relatively small dataset makes their detection and even proper manual selection during detector's training phase



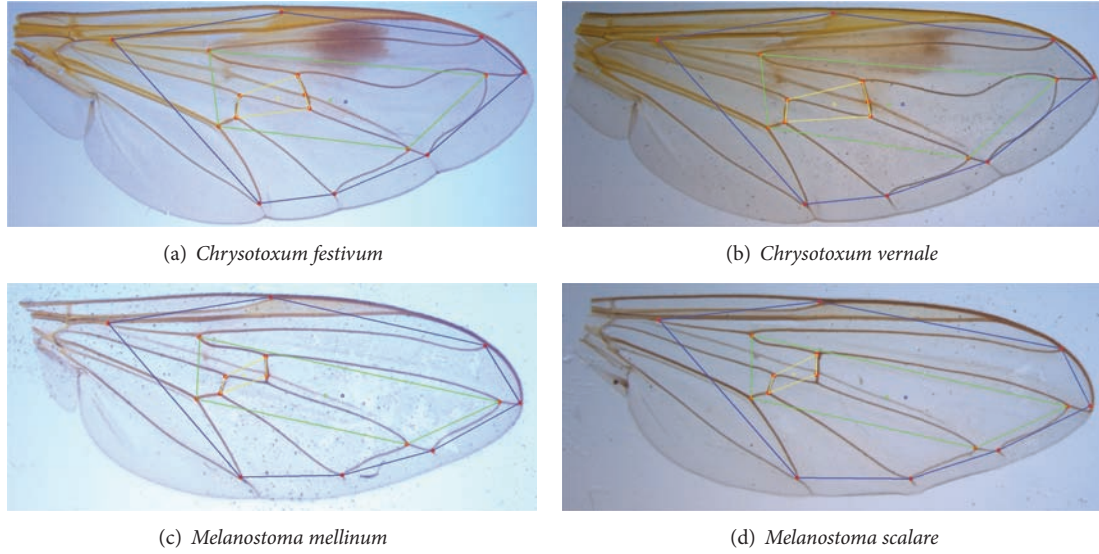


FIGURE 6: Illustration of constructed convex hulls used for characterization of each wing image in automatic discrimination of four selected hoverfly species. Convex hulls are created using detected landmark-points, which are drawn as red dots. Position of each landmark-point is computed as a centroid of multiple detections of the same junction. Multiple detections are illustrated in Figure 5.

much harder. Therefore, before characterization of each wing image by its convex hulls, detections which are expected to correspond to these landmark-points are removed.

Automatic discrimination of four selected hoverfly species is then made using SVM classifier with polynomial kernel (5) implemented in [18], with the following set of parameters:  $\gamma = 1$ ,  $C = 1.5$ ,  $d = 3$ , and  $c_0 = -0.5$ . The results of 10-fold cross-validation using 774 wing images are presented in Tables 4 and 3 and discussed in the following section.

#### 4. Results

The performance of automatic landmark-points detection using different sliding window step sizes was analyzed. The step sizes of 8, 16, and 32 pixels were used and different degrees of landmark-points detection per image were achieved. Using the sliding window with the largest step size is significantly faster than the alternatives but with the smallest number of detected landmarks per image and it is the most imprecise due to the absence of multiple detections. The highest detection accuracy was achieved using the smallest step size, so this sliding-window detector was selected to serve as the basis for species discrimination (classification) using polynomial SVM described in Section 3. Classification results, obtained using 10-fold cross-validation, are presented using different performance measures in Table 4 and by accuracy assessment matrix given in Table 3. The accuracy assessment matrix (confusion matrix), Table 3, is a representation of misclassification errors and ideally contains nonzero values only on the main diagonal. The values outside the main diagonal in Table 3 show the number of images of a class indicated on the left (reference data) that have been labeled by the classification algorithm as the class indicated on the top (classified data). The average classification accuracy of

TABLE 3: Classification results, accuracy assessment matrix for four selected hoverfly species, denoted by letters (a)–(d): *Chrysotoxum festivum* (a), *Chrysotoxum vernale* (b), *Melanostoma mellinum* (c), and *Melanostoma scalare* (d).

		Classified				Total samples
		a	b	c	d	
Reference	a	211	35	0	2	248
	b	45	102	4	3	154
	c	3	1	243	20	267
	d	2	3	25	77	105

81.6% was achieved among the four different species, while from the accuracy assessment matrix in Table 3 it can be observed that the classification accuracy between the genera (*Chrysotoxum* and *Melanostoma*), that is, the two different groups of hoverfly species, is much higher and is 97.7%. The reason is that the intergenera differences are much higher than the differences between the species inside the same genus.

These results confirm the applicability of the proposed approach in the sense that the used features, based on obtained automatic detections, enable very high discrimination between the two genera inside the same flying insects family, Table 3.

Finally, in order to better understand and further investigate properties of the given classification problem and descriptive capacity of the proposed convex hull based and simple low-dimensional features, binary classification of selected four species was also analyzed. Instead of a performance of a single multiclass classifier, performance of four binary classifiers was measured. The same type of classifier and the same convex hull based features as in the

TABLE 4: Classification performance of multiclass classifier which utilizes convex hull based features: true positives (TP) rate, false positives (FP) rate, precision, and recall.

	<i>Chrysotoxum</i>		<i>Melanostoma</i>		Weighted average
	<i>festivum</i>	<i>vernale</i>	<i>mellinum</i>	<i>scalare</i>	
TP rate	0.851	0.662	0.910	0.720	0.816
FP rate	0.095	0.063	0.057	0.037	0.067
Precision	0.808	0.723	0.893	0.755	0.813
Recall	0.851	0.662	0.910	0.720	0.816

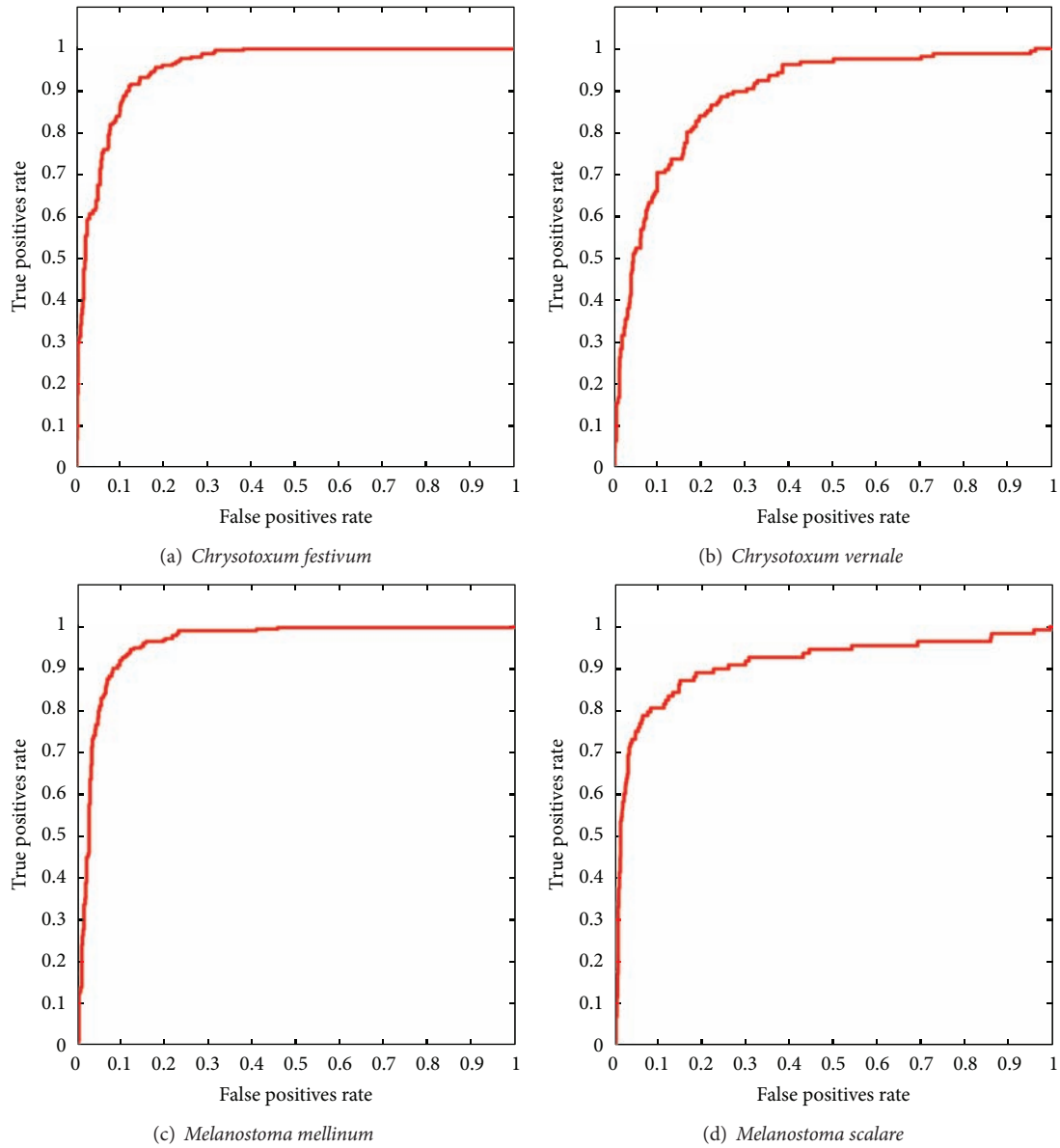


FIGURE 7: ROC curves corresponding to four binary classification problems, (a)–(d). In all cases instances of the class denoted in subcaption were considered as positives, while instances of the remaining three classes were labeled as negatives. The same type of SVM binary classifier with different sets of parameters was used in all experiments.

previously described multiclass scenario were used in all experiments. Receiver operating characteristics (ROC) of the corresponding binary classifiers with different sets of parameters are shown in Figures 7(a)–7(d). Classifiers represented by the curves (a) and (c) exhibit better performance, that is, have steeper curves than the classifiers described by the curves (b) and (d). This is interesting behaviour because positive instances of the first two classifiers, (a) and (c), belong to hoverfly species from different genera, which is also true in the case of the last two classifiers, (b) and (d). It is a consequence of the fact that the positive classes corresponding to curves (a) and (c) have significantly higher number of samples in comparison to the number of instances corresponding to the positive classes in curves (b) and (d), as it can be observed from Table 3 or Table 1. It also suggests that the complexity of the classification task, which is based on the proposed features derived from automatically detected vein joints in wing images, is more influenced by the unbalanced training set than by some wing image characteristics specific for particular genera (*Chrysotoxum* or *Melanostoma*). These results are also in accordance with multiclass case in Table 3, where the most misclassification errors appeared between the species inside the same genus.

All datasets that were used in this paper and the accompanying C++ code, which was used for feature extraction and classification, can be found at <http://www.biosensecenter.com/index.php/hoverflies-classification>.

## 5. Conclusion

Systems for automatic classification of insects are generally intended for field use. Therefore, it is desirable that they are robust and as general as possible. At the present time image based systems are considered as the preferred choice in comparison to some other alternative solutions, like for example, DNA analysis, since they are mobile and more affordable. An image processing approach to hoverfly species discrimination presented in this paper showed promising results on the collected wing images dataset. Its advantage is that it utilizes robust method for detection of landmark-points in wing venation patterns of flying insects which is based on the proposed combination of HOG and RCLBP descriptors and can cope with different image imperfections. Simple rotation-invariant features chosen for later wing classification are one possible solution for the problem of unpredictable number of automatic detections and proved to be discriminative enough to distinguish correctly between two different hoverfly genera with the very high accuracy of 97.7%, and to a lesser extent between the two species that comprise each genus. The proposed classification method may be used for engineering of a complex automated hoverfly species identification system which would achieve high accuracy and significant level of robustness.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgment

This research work has been supported by the Ministry of Education, Science and Technological Development of Republic of Serbia, as part of the research project III43002.

## References

- [1] D. Houle, J. Mezey, P. Galpern, and A. Carter, "Automated measurement of *Drosophila* wings," *BMC Evolutionary Biology*, vol. 3, article 25, 2003.
- [2] T. Arbuckle, S. Schroder, V. Steinhage, and D. Wittmann, "Biodiversity informatics in action: identification and monitoring of bee species using ABIS," in *Proceedings of the 15th International Symposium Informatics for Environmental Protection*, pp. 10–12, 2001.
- [3] N. Larios, H. Deng, W. Zhang et al., "Automated insect identification through concatenated histograms of local appearance features: feature vector generation and region detection for deformable objects," *Machine Vision and Applications*, vol. 19, no. 2, pp. 105–123, 2008.
- [4] N. MacLeod, Ed., *Automated Taxon Identification in Systematics: Theory, Approaches and Applications*, CRC Press, Boca Raton, Fla, USA, 2007.
- [5] Y. Zhou, L. Ling, and F. Rohlf, "Automatic description of the venation of mosquito wings from digitized images," *Systematic Zoology*, vol. 34, no. 3, p. 346, 1985.
- [6] A. Tofilski, "Using geometric morphometrics and standard morphometry to discriminate three honeybee subspecies," *Apidologie*, vol. 39, no. 5, pp. 558–563, 2008.
- [7] F. Rohlf and J. Archie, "A comparison of Fourier methods for the description of wing shape in mosquitoes (Diptera: Culicidae)," *Systematic Zoology*, vol. 33, no. 3, p. 302, 1984.
- [8] W. Thompson, *On Growth and Form*, Cambridge University Press, Cambridge, UK, 1945.
- [9] X. Zheng and X. Wang, "Fast leaf vein extraction using hue and intensity information," in *Proceedings of the International Conference on Information Engineering and Computer Science (ICIECS '09)*, pp. 1–4, IEEE, December 2009.
- [10] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.
- [11] J. Chen, V. Kellokumpu, G. Zhao, and M. Pietikäinen, "RLBP: robust local binary pattern," in *Proceedings of the British Machine Vision Conference (BMVC '13)*, pp. 122.1–122.10, BMVA Press, Bristol, UK, 2013.
- [12] N. Dalal and B. Triggs, "HOG for human detection," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 886–893, IEEE, 2005.
- [13] X. Wang, T. Han, and S. Yan, "An HOG-LBP human detector with partial occlusion handling," in *Proceedings of the IEEE 12th International Conference on Computer Vision (ICCV '09)*, pp. 32–39, Kyoto, Japan, September 2009.
- [14] Z. Guo, L. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1657–1663, 2010.
- [15] B. Brkljač, M. Panić, D. Čulibrk, V. Crnojević, J. Ačanski, and A. Vujić, "Automatic hover species discrimination," in *Proceedings of the 1st International Conference on Pattern Recognition Applications and Methods*, vol. 2, pp. 108–115, SciTePress, 2012.



- [16] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, NY, USA, 2000.
- [17] G. Bradski and A. Kaehler, *Learning OpenCV*, O'Reilly Media, Sebastopol, Calif, USA, 2007.
- [18] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. Witten, "The WEKA data mining software: an update," *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 1, pp. 10–18, 2009.
- [19] C.-C. Chang and C.-J. Lin, "LibSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 1–27, 2011.

