Support Vector Machine (SVM) Recognition Approach adapted to Individual and Touching Moths Counting in Trap Images

M. C. BAKKAY*, S. CHAMBON*, H. A. RASHWAN*, C. LUBAT† and S. BARSOTTI†
*University of Toulouse, IRIT, France
†SiConsult, Labège, France
Email: schambon@enseeiht.fr

Abstract—This paper aims at developing an automatic algorithm for moth recognition from trap images in real-world conditions. This method uses our previous work for detection [1] and introduces an adapted classification step. More precisely, SVM classifier is trained with a multi-scale descriptor, Histogram Of Curviness Saliency (HCS). This descriptor is robust to illumination changes and is able to detect and to describe the external and the internal contours of the target insect in multi-scale. The proposed classification method can be trained with a small set of images. Quantitative evaluations show that the proposed method is able to classify insects with higher accuracy (rate of 95.8%) than the state-of-the-art approaches.

I. INTRODUCTION

Insecticides are expensive and dangerous for plants and humans. Therefore, farmers attempt to survey insect species and to evaluate their density in the fields in order to adapt and to reduce the use of insecticides. For this purpose, it is needed to catch insects and then, to manually count these insects in order to analyze the evolution of the insect population and to take the decision for using insecticides or not. Unfortunately, manual counting of these insects from trap images is slow, expensive, and sometimes error-prone. Thus, developing a system, which can achieve a completely automated detection, and that can recognize and count insects is very advantageous.

In the field of detection/segmentation of objects, many computer vision techniques have been introduced. These techniques are, to name a few, edge detection [2], snake contour detection [3], clustering with k-means [4] and mean-shift [5]. For recognition/classification, usual and well-known techniques concern: Support Vector Machine [6], and more recently Deep Learning [7], sometimes based on well known descriptor, such as Histogram Of Gradient, HOG [8] or Scale Invariant Feature Transform (SIFT) descriptor [9].

Insect classification is a challenge because it needs to recognize a small object with poor color and shape characteristics in a non-homogeneous background that can contain some difficulties. In particular, trap images may contain noise: very small insects or herbs, the pheromone cap or some lines of glue, see Fig. 1. Moreover, since the trap is installed in outdoor environment, the trap images are also affected by illumination changes [10]. Finally, touching and overlapping insects can also be found in the trap yielding a complex counting task.

For insect recognition/classification, many methods have been applied to butterfly classification, like [11], [12] but few publications are dedicated to agricultural insects. We can consider two different possibilities for doing this task. On one side, methods have considered insect specimens [11], [13] where images contain few difficulties and are at high resolution. In this case, classification can be applied directly without the need of a segmentation step [14]. On the other side, in wild trap images, two steps are needed: a first insect segmentation before applying a classification approach in a second part. These methods encounter many challenges: low
image quality, illumination, movement of the trap, movement of the moth, camera out of focus and presence of noise (such as leaves, broken wings, etc). In this work, we are interested in working with trap images. In the literature, for the segmentation, the authors can use color, shape and texture features [15], [16], active contour segmentation [15] or morphological-based segmentation [17]. All these methods do not consider the presence of touching insects in the trap. Most of descriptors used are not robust to illumination changes and do not detect occluded contours. So, this is why, in our previous work [1], we both propose to use features that use most of the characteristics of the moth and to take into account the problem of touching insects.

For the classification, many techniques have been tested, like k-nearest neighbor classification [15] or Support vector machine, SVM [16]. In [17], the authors even consider a pose estimation-dependent classification using deep learning. Deep learning based methods need large training datasets that are not always available for agricultural insects. And, in particular, in this work, we do not have such a database.

This work aims to study the invasion of a particular moth, which is Lobesia Botrana (Eudemis), a European vine moth, for adapting the pesticide treatment of grape culture. More precisely, wine producers usually capture this particular moth, they count the number of insects and then they analyze the evolution of this counting in order to confirm the use of the pesticides or not. Consequently, in this paper, we propose an automatic algorithm for moth classification in trap images. First, a hybrid segmentation approach, introduced in our previous work [1], is used to eliminate noise and to separate touching insects. Then, a SVM classifier is trained with a proposed multi-scale descriptor, named Histogram Of Curviness Saliency (HCS) [18]. This descriptor is robust to illumination changes and is able to describe the external and the internal contours of the target insect in multi-scale. Moreover, the proposed classification method can be trained with a small set of images.

In the next section, the proposed method for insect recognition is presented. To demonstrate its effectiveness, experimental results with a comparative study are detailed in Section III before the conclusion and perspectives, in Section IV.

II. SVM RECOGNITION BASED ON HCS DESCRIPTION

A. Overview

The method allows to recognize individual moths in images of a trap that contain many difficulties (noise, touching moths and elements that are not insects), as presented in Fig. 1. The method that starts by detecting any kind of insects, steps (1) and (2), that have been previously published and described in [1] and then classifies the insects of interest, Lobesia Botrana, Eudemis, from the other insects, like, for example, flies or spiders, see Fig. 2. This last step (3) corresponds to the main contribution presented in this paper. However, for the interest of the reader, we briefly described the previous published steps (1) and (2).

(1) Robust contour detection – First of all, we apply a robust contour detection [18]. Then, we apply a k-means algorithm to classify the estimated contours into different categories: contours due to noise (class 1), contours related to individual insects (class 2) and contours that contain touching insects (class 3 to k).

(2) Region-based segmentation – After this automated clustering step, the next task attempts to separate the obtained possible touching insects by applying a region-based segmentation. The idea is to use the contours classified into class 3 to k that delimit regions as seeds for the watershed algorithm [20]. This region-based segmentation part contains three ordered steps: the contour dilation, the region merging and the watershed segmentation. In fact, contour dilation is
used to remove discontinuities of the detected contours while the region merging step avoids the over-segmentation of the watershed algorithm. Finally, two results are possible:

(a) The watershed algorithm detects two or more regions inside the contour, thus touching insects will be separated to two or more insects.
(b) The algorithm detects only one big region and the shape of this insect is just refined.

All this previous work provided an encouraging detection. More precisely, the qualitative results obtained are: 87% of noises or objects that are not insects are detected as noises, 70% of insects are detected as insects and 82% of touching insects are detected as touching insects. Moreover, the proposed method classifies 13% of noises as insects, however, no noise is detected as touching insects. Then, 13% of insects are detected as noises, and 17% of insects are detected as touching insects. Finally, 18% of touching insects are detected as insects and no touching insects are detected as noise.

(3) Insect classification – The next and final step, that corresponds to the contribution presented in this paper, concerns the insect classification. For that purpose we introduce an approach based on Support Vector Machine (SVM) that uses the Histogram Of Curviness Saliency (HCS). In the rest of the section, we will present the two aspects: how HCS is adapted to this specific recognition and how this descriptor is introduced in the SVM techniques.

B. Histogram of Curviness Saliency (HCS)

In the aforementioned detection step, we used the detector proposed in [18] and based on curviness saliency that is an estimation of curvature. This detector generates both the magnitude and the orientation of curvature features. Thus, we naturally expand the concept of the famous classical HOG, Histogram Of Gradient [8], widely used in the feature description literature, to work on these curviness saliency features as proposed in [21]. More precisely, the orientation and the magnitude of the curviness saliency are used for building a descriptor called the Histogram Of Curviness Saliency (HCS). That HCS is used in a sliding window fashion in a Region Of Interest (ROI) (i.e., in this work a detected insect) to generate dense features based on binning the curvature orientation over a spatial region. The orientation and the magnitude of the curviness saliency can be computed, as described in [21], with:

\[ CS = \left( (I_{xx} - I_{yy})^2 + 4I_{xy}^2 \right), \]  

(1)

where \( I_{xx} \), \( I_{yy} \) and \( I_{xy} \) are the second derivatives of the image. Using this curviness saliency in multi-scale leads to this equation, as illustrated in [18]:

\[ \text{MCS}^2 = \text{MCS}'^2. \]  

(2)

It means that the multi-scale of \( CS \) (MCS) of a ROI is multiplied by the eigenvector \( e_1 \) corresponding to the curviness saliency of a pixel.

Using the same principle of HCS presented in [21], we propose a descriptor that contains the orientation of the curvature of MCS binned into sparse per-pixel histograms.

C. SVM classification based on HCS

We used SVM [6] to construct a hyperplane that separates the two classes of training data HCS descriptors (moths and other insects). Consequently, to be adapted to the problem, we suppose that the two classes are not linearly separable, and, so, the function of this surface is given by:

\[ f(\vec{x}) = \text{sgn} \left( \sum_{i=1}^{p} \alpha_i^* \vec{y}_i k(\vec{x}_i, \vec{x}) + w_0 \right), \]  

(3)

where \((\vec{x}_i, \vec{y}_i)\) is related to the training data, \(\vec{x}_i\) is the \(p\)-dimensional HCS descriptor, \(\vec{y}_i \in \{-1, 1\}\) is the class label, \(k\) is a kernel function, \(\alpha_i^*\) are optimal Lagrange multiplier and \(w_0\) is the bias. For the kernel function, many kernels are possible, see [6] to have an overview. In this work, the polynomial kernel function of degree 6 defined as:

\[ k(\vec{x}_i, \vec{x}_j) = (\vec{x}_i \cdot \vec{x}_j + 1)^6, \]  

(4)

provided the best classification rate.

III. EXPERIMENTATION

A. Dataset

We have collected annotated image patches (2865 negative patches and 746 positive patches) from moth trap images. In these images, captured under different illumination conditions, there are many insects of varying types and sizes, different elements that can induce false detections (herbs, very small insects, lines of glue, pheromone cap, etc.), see Fig. 3.

We performed data augmentation to increase the number of positive patches from 746 to 3577. This allows also to incorporate invariance to basic geometric transformations and to noise into the classifier. Therefore, the data augmentation that we used consists in applying geometric transformations (rotation and translation), blurring, adding Gaussian noise and aspect ratio transformations to the original patches [14].

To select training and testing sets from collected patches, we used a 5-fold cross-validation that avoids over-learning and under-learning. Patches are randomly partitioned into 5 subsamples. The cross-validation process is then repeated 5 times. In each iteration, one of the 5 subsamples is retained as the validation data for testing the model, and the remaining 4 subsamples are used as training data. This is the standard algorithm for validating any classification approach [22].

B. Parameter study

In the proposed approach, the parameters for computing HSC descriptor are:

1) detection window
2) block size
3) block stride
4) cell size
5) number of bins per cell

These parameters are explained in details in [18] and [21]. The parameter for SVM is the soft margin \(C\). In these experiments, detection window are set to \((64 \times 48)\), and block size to \((8 \times 8)\). In addition, block stride are set to \((4 \times 4)\) and cell
C. Experimental Results

In Fig 3, we present visual results obtained. As shown, most insects (green and purple rectangles) and noise (blue rectangles) are detected by the segmentation algorithm. In addition, touching insects (red rectangles) are separated in most of the cases. The recognition step correctly separates moths (green rectangles) to other insects (black rectangles).

We implemented some approaches that have been presented in section I and we compared them to the proposed approach in section I and we compared them to the proposed approach. In table I, for all the tested methods, the accuracy rates for all tested images have been computed. The best results are obtained with 0.1.

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed method</td>
<td>95.8%</td>
</tr>
<tr>
<td>Ding et al. [14]</td>
<td>93.2%</td>
</tr>
<tr>
<td>Yao et al. [16]</td>
<td>90.1%</td>
</tr>
<tr>
<td>Yalcin et al. [15]</td>
<td>87.6%</td>
</tr>
</tbody>
</table>

IV. CONCLUSION

In this paper, a classification method adapted to moth recognition has been introduced. It first uses a detection that is able to detect individual and touching insects from trap images. Second, SVM classifier is trained with a proposed multi-scale HCS descriptor which is robust to illumination changes in multi-scale way. The proposed classification method can be trained with a small set of images. Compared to state of the art methods, this new method brings the best classification rate. For future works, we plan to generalize all the proposed method to other species of insects.

REFERENCES


ACKNOWLEDGMENT

This work was conducted as part of the project VITI OPTIMUM 2.0. The project has been approved by the French competitiveness clusters: Agri Sud-Ouest Innovation and IAR. The project has been cofinanced by BPIFrance, European Regional Development Fund and the Occitania Region.