Pattern Recognition 65 (2017) 273–284

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

A survey on image-based insect classification

Maxime Martineau[a,*], Donatello Conte[a], Romain Raveaux[a], Ingrid Arnault[b], Damien Munier[b], Gilles Venturini[a]

[a] LI, Université François Rabelais de Tours, 37200 Tours, France
[b] CETU Innophyt, Université François Rabelais de Tours, 37200 Tours, France

ARTICLE INFO

Keywords:
Image-based insect recognition
Classification
Automated entomology
Arthropods
Biology
Recognition
Image

ABSTRACT

Entomology has had many applications in many biological domains (i.e insect counting as a biodiversity index). To meet a growing biological demand and to compensate a decreasing workforce amount, automated entomology has been around for decades. This challenge has been tackled by computer scientists as well as by biologists themselves. This survey investigates fourty-four studies on this topic and tries to give a global picture on what are the scientific locks and how the problem was addressed. Views are adopted on image capture, feature extraction, classification methods and the tested datasets. A general discussion is finally given on the questions that might still remain unsolved such as: the image capture conditions mandatory to good recognition performance, the definition of the problem and whether computer scientist should consider it as a problem in its own or just as an instance of a wider image recognition problem.

1. Introduction

Building accurate knowledge of the identity, the geographic distribution and the evolution of living species is essential for a sustainable development of humanity as well as for biodiversity conservation. Insects are a class of invertebrates within the arthropods that have a exoskeleton, a three-part body (head, thorax and abdomen), three pairs of jointed legs, compound eyes and one pair of antennae. This is a subclass of the Arthropods class which also includes Arachnids, Myriapods and Crustacean and they have been the dominant component of animal species diversity for all of the past 520 million years, since the main burst of the Cambrian radiation. They have been considered to be the earliest colonizers of land [1,2]. Fossil remains demonstrate that Arthropod species are robust over long periods of time, and that, given the opportunity, they migrate with changing conditions rather than evolving new species [3]. Arthropods are of exceptional value in the reconstruction of paleoenvironments because they are provide detailed, precise information on vegetation, soils, water quality, vertebrate species composition, forest composition and degree of stress [3]. Arthropods represent up to 80% of animal phylum [4] and make up for the largest proportion of species richness at any spatial scale [5]. With 1,5 millions of species, they are more representative for wholesale organism biodiversity than any other group. Arthropods have been recognized as efficient indicators of ecosystem function and recommended for use in conservation planning [6–9]. Many researchers have assessed habitat quality and measured habitat differences using Arthropods within different ecosystems as well as agricultural, forest and urban landscapes [10–14]. More specifically, arthropod diversity has been used to indicate the impacts of habitat modification [7,15,16] to measure the effects of human disturbance [17].

Species identification of Arthropods is a fundamental part of recognizing and describing biodiversity. Traditionally, identification has been based on morphological diagnoses provided by taxonomic studies. Only experts such as taxonomists and trained technicians can identify taxa accurately, because it requires special skills acquired through extensive experience. However, the number of taxonomists and other identification experts has drastically decreased. In some cases a trained technician could make routine identifications using morphological ”keys” (step-by-step instructions of what to look for), but in most cases an experienced professional taxonomist is needed. The use of Arthropods as biotic indicator is slown down by the high level of expertise required to identify insects. It is an obstacle to a broad and easy application [18]. Consequently, alternative and accurate identification methods non-experts can possibly use are required.

Finding automatic methods for such identification is an important topic with many expectations. One of the most common data that can be used in this context are images. Images of arthropods can be acquired and further processed by an image classification system. Such images also have some specificity (see the next sections) and might

* Corresponding author.
E-mail addresses: maxime.martineau@univ-tours.fr (M. Martineau), donatello.conte@univ-tours.fr (D. Conte), romain.raveaux@univ-tours.fr (R. Raveaux), ingrid.arnault@univ-tours.fr (I. Arnault), damien.munier@univ-tours.fr (D. Munier), gilles.venturini@univ-tours.fr (G. Venturini).

http://dx.doi.org/10.1016/j.patcog.2016.12.020
Received 4 August 2016; Received in revised form 3 November 2016; Accepted 19 December 2016
Available online 23 December 2016
0031-3203/ © 2016 Elsevier Ltd. All rights reserved.

require adaptation of existing image recognition methods, or new methods.

The aim of this paper is to provide an overview of this domain as thorough as possible. It is structured as follows: Section 2 describes with more precision the scope of this survey and the methodology used to analyze the literature. Section 3 considers the first step of all approaches, i.e., image acquisition. Section 4 is devoted to features extraction, some of them being specific to arthropods. Section 5 describes the classification of the data into known classes. Section 6 details the existing datasets and, finally, Section 7 provides a discussion on this domain, with some conclusions and perspectives.

2. Scope of this survey and methodology

2.1. Scope and previous surveys

The scope of this survey can be stated as follows: the kind of animals we deal with are arthropods. In a vernacular manner, we will abusively use the term insects to denote these animals. The data acquired from these animals are 2D images: we do not consider other kinds of data, such as 3D images [19], sound [20,21], genomic data [22], etc. We are especially interested in methods that can process 2D images with “normal lighting and shooting” conditions, that is when images are acquired by standard cameras (even mobile or embedded devices). Finally, we consider the systems that extract features from images and that perform a classification of the animals into known classes.

Two previously published surveys are relevant to these scope and aims. In [23], the authors take the general point of view of species identification and discuss the current state of the field and possible perspectives. They underline the key difficulties to solve this problem and the reasons why it is yet not so much developed. They discuss examples of successful achievements that were initially made and provide motivations for further studying this problem. In [24], the authors list a few papers from recent years. In this survey, we adopt a more focused and detailed level of description than in [23]. In comparison to [24], we provide more references and a common framework to analyze and compare them. The proposed survey is more centered on computer science point of view than the cited surveys. Consequently, it can be very useful for researchers in Pattern Recognition field. Furthermore, final discussions provide valuable questions that could be the starting points for researcher who intend embarking on automatic arthropods identification.

2.2. Problem definition

Image-based arthropod classification could be seen as an application of image classification. Based on some photograph depicting the specimen, its biological identity is to be determined. The peculiarities of the problem are three-fold : taxonomy, image variations and incrementality.

Arthropod species, as for any living species, are grouped into taxons at several scales. These scales constitute the biological taxonomy which can be seen as a tree grouping species based on genetic similarity. In Fig. 1 are shown the scales relevant to arthropod classification. The biologists might be willing to use a scale or another depending on their scopes and applications. For instance, biodiversity studies need very fine information (down to species precision) and, conversely, pest management applications require less precise information (say to a genus information). The problem can therefore be seen differently along granularity in the taxonomy.

Aside from classical object image variations (such as rotation, scale, parallax, background or lightning), insect images have more particular properties such as pose (because specimen appearance varies with the orientation they are been shown) and deformation (because the specimens are most of the time composed of articulated parts). These

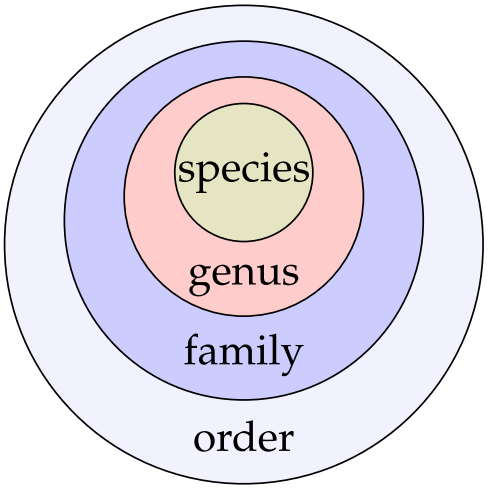

Fig. 1. The taxonomic scales relevant to arthropod classification.

aforementioned variations can be referred to as *capture-specific* variations in the sense they only depend on capture factors. About the objects themselves (*object-specific* variations), age, sex and genetic mutations are the main factors of visual variations. The most instructive example is that of lepidoptera (commonly referred to as butterflies) which can have extremely different visual aspects along time (being successively caterpillars, chrysalises and butterflies).

Finally, given the considerable diversity amongst arthropods, classes are highly numerous and, besides, new classes remain undiscovered. This property can be seen as incrementality.

2.3. Methodology

To describe this state of art on image-based recognition of insects, we used the following methodology. Starting from a set of fourty-four papers, we described each paper along several analytic points of view. These analytic points of view are related to a classical image recognition system. Such a system can be decomposed into three sequential phases: a) Image capture; b) Feature extraction and c) Classification. Image capture consists of the tools and methods used to take photos of the insects (discussed in Section 3). Features extraction consists of extracting useful visual information out of the pictures (discussed in Section 4). Finally, classification is the actual step at which recognition is performed based on the extracted information (discussed in Section 5). The datasets are the images used to train and test the system (discussed in Section 6).

A set of criteria were chosen to analyze the literature from different perspectives and summarize it into useful information for further use. Are images capture with or without constraints? In the wild or in the lab? Are features domain-dependent or conventional from the computer vision community? What are the classifiers used in the literature? How is the performance evaluation carried out? What are the number of classes or the data set size? Are data sets publicly available or not? All these criteria can be seen as different dimensions or angles to categorize, summarize and identify relationships between approaches. Once these pieces of information are obtained, the papers were clustered based on them. Both dimensions and clusters are further exposed along the next sections.

3. Image capture settings and apparatuses

A first categorization which can be made is based on the type of image capture setting and on the context of the acquisition. In fact, the acquisition mode might influence the techniques used for the identification of the insects, but also the context of application of such algorithms. In this context, several analytic dimensions can be con-

sidered. The first is about acquisition conditions, that basically deals with the insects (both in the way they are captured and taken care of) and the way the images are produced (the capture tools and the overall protocol). The second is what we name as pose constraint, that is how much we allow the point of view on the insect to be variating. A constrained pose setting specifies we want to see, for instance, only the front side of the insect we capture the image of. An unconstrained pose setting, in the contrary, induces variability in orientation of the individual and therefore in their apparent shapes, details and colours. Next, closely related to the pose, the part or area of the individual is to be considered. Some insects are better recognized from their wings as others from their entire body.

Therefore we have divided the considered works into two broad categories: lab-based setting and field-based setting. In a lab-based setting there is a fixed protocol for image acquisition. This protocol governs the insect trapping, its placement and the material used for the acquisition (capture sensor, lighting system, etc.). Lab-based setting is mainly used by entomologists bringing the insects to the lab to inspect them and to identify them. Therefore, they have the possibility to manage the image acquisition system. By field-based setting we mean a capture of insects directly in cultivated fields, without any particular constraints to the images capture system. Typically the acquisition is made by a mobile device and the insect is alive when the picture is taken. The application context of this type of setting is that of a farmer who wishes to identify an insect to know, for example, if it belongs to a devastating species or not. In this context no, or little, knowledge of the identification system is required by the user other than the user interface. At these two categories we add a third, smaller, which includes the works whose acquisition system is such that in a picture there are more insects than one. In such setting, a more complex segmentation phase might be required before the identification to isolate each insect.

A taxonomy of the reviewed papers is shown in Table 1. For each paper the taxonomy reports the corresponding bibliographic reference. We will highlight, in the next paragraphs, the most interesting acquisition systems from the considered papers (or a group of papers using similar settings).

In lab-based setting (see examples on Fig. 2a and b), most of the acquisition systems are manually manipulated. In [58,32,33,35,34,36–40,62,46,52] the acquisition is made manually on a microscope slide, with some standard illumination systems (fiber optic light sources, desk lamps, etc.). In [58], to change the pose of the insect, an interesting trick has been used: body rotation was achieved by first placing two short nylon fishing lines on either side of the specimen and gently moving a cover slip placed over them by using the tip of the index finger. In [32,35,34,36,46,52,51] some zoom lenses are used to have more details of the insect or to capture only some parts (wing, genitalia, etc.). In [36] live bees captured in the wild are cooled down using an icebox: this makes them immobile for a time sufficient for the recognition image to be acquired without causing the bee long-term harm. Also in [37] active moths were refrigerated to prevent them escaping while being imaged. In [51] a stereoscope is used in order to avoid a view with obstructed parts. In [61] the insects were frozen for 20 min and then randomly and manually placed on a white balance panel. The images were taken under different orientations and two poses (top and side view).

In [63] the main goal of the paper is not the insect identification but the detection of insects in bulk wheat samples. Nevertheless, it is interesting to note that their acquisition system could be used also for insect identification. Samples, that are a blend of wheat kernels and dockage including some grass seeds, are placed in a crate cell, are captured by a color capture device, with a backlight illumination. To reduce reflection, a dome with flat matte white paint inside was placed over the samples.

There are also semi-automatic systems. In [57,59] the authors propose a semi-automated mechanical manipulation and imaging system that allows the positioning of the specimen and its rotation to obtain the best view for classification. The decision of the view is taken by the human operator. In [19], a similar but more complex system is proposed: a mirror system is providing a split image, with two views of the insect approximately 90° apart. Providing two simultaneous views increases the likelihood of good dorsal views for 2D classification, and it brings greater robustness for 3D reconstruction since the camera viewpoints of the two views are fixed.

In [41,45,49] the acquisition system is similar to the above-described ones, but the insect is captured with a classic color digital camera without microscope, so the insect has a “natural” size in the image. In [49] a transparent plastic stick was fixed to a piece of foam and held at the far end of a needle that was inserted into the specimen, in order to eliminate the shadow due to the fill lights used.

Typically in field-based setting (see Fig. 2c and d) there is no particular acquisition hardware and the insects are captured by standard cameras or mobile devices in open field. In [27], even if the acquisition is performed on open fields, pheromone traps are used and cameras are being integrated to the trap setup in order to improve the precision of insect information captured and facilitate the subsequent segmentation phase.

In multi-individuals (see examples on Fig. 2e and f) setting [30,66,65,29], the positioning of the insect is also manual and is similar to the setups described for manual setting in lab-based category. In [30] the rice pests caught in traps are spread on the glass plate with as little overlap as possible, which will make the following image segmentation and identification easier. The acquisition setting is quite standard, but an interesting particularity is that two digital cameras above and below the glass plate are fixed on a stainless holder in order to have multiple poses of the insect without manipulating either the insect or the acquisition system.

In conclusion we can say that in field of insect identification, the lab-based setting is still the most widely used setup and the positioning of the insects is made mostly manually with a constrained pose (see Table 1). More recently, the study is done considering several individuals in a single image, but the acquisition system is still largely manual. There are still few works that develop insect identification systems in open field, probably because of a more difficult recognition context (cluttered background, live insect, etc.). In fact in some open-field setting some expedients are used to have an image simple to process.

Table 1
Taxonomy of reviewed papers on the basis of image capture setting.

Category Levels			Refs.
Lab-based	Manual Positioning	Constrained Pose	[31–40,19,41–45,26,25,46–55]
		Unconstrained pose	[56–59,30,60–62,52]
	Automatic Positioning		[63,57,59]
Field-based			[28,64,27]
Multi-individuals			[29,65,66,30]

3.1. Advantages and drawbacks

With the large increase of mobile devices, an equivalent growing open-field captures is expected. In fact such a system is more flexible and it can be used by a wider range of users. Open-field captures do not require special hardware and there are no constraints on insect pose. However this kind of systems are more complex in terms of algorithm design and it is hard to obtain good performances.

Anyway, lab-based settings still remain a configuration that will be studied in the future. Indeed, biologists have some specific research

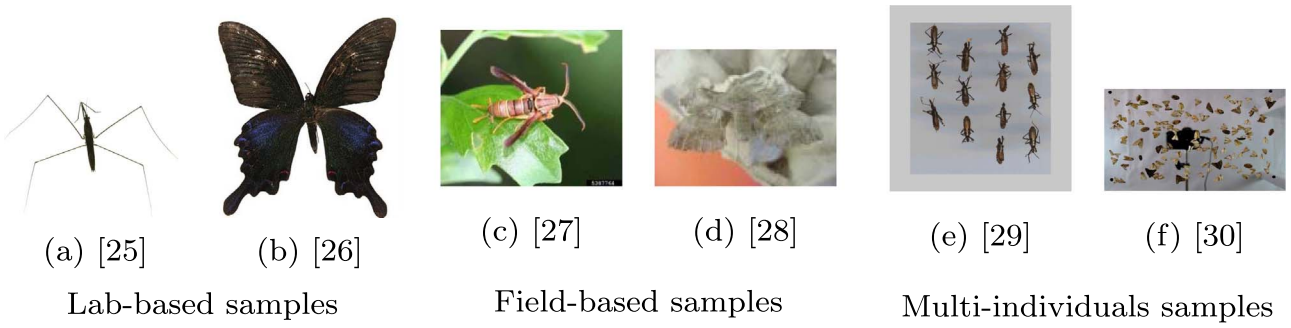

Fig. 2. Image samples.

that need insect to be studied in laboratory. Therefore, while open-field capture will be used by non-expert people, lab-based setting will continue to grow because of specific need of biologists. The advantage of lab-based setting, from the point of view of the classification system, is the possibility of constraining the insect so that the recognition task is less complex thanks to less variability. Nevertheless, the main drawback of lab-based setting is the manual manipulation of the capture system (posing of the insect, illumination fixing, etc.). This manual task is very tedious and time-consuming and researcher ask more and more for automatic manipulation systems.

In order to speed up the process of recognition, multi-individuals setting are conceived. In this case, the time is reduced, because of multiple classification in a single capture. However, again the classification system has to deal with more image processing problems, like detection of insects that is a difficulty not present in original lab-based setting.

4. Feature extraction

Focusing on feature extraction, we listed the two following aspects: the segmentation which is the data extraction phase such that the insect is isolated (foreground/background detection) and the types of features used (whether they are local, global and what information they model).

4.1. Foreground/background detection

Before the features can be extracted, the system needs to know which zone of the image is important. This part is taken care of by segmentation. This segmentation problem was tackled in several ways, as listed in Table 2.

In some cases, segmentation is learned in a supervised manner as in [63,27,46]. [46] feeds the entire image to the classifier, which is supposed to learn to discriminate only on foreground and therefore to recognize it. In [27], the approach relies on learning on both positive and negative example images. Negative example images are images where there is no insect but only foreground. Learning on such images enables the system to ignore the background.

Table 2
Taxonomy of the reviewed papers on the basis of segmentation.

Category Levels		Refs.
No segmentation		[63,27,46]
User segmentation		[37,25,26,48,39,55,32,40,33]
Thresholding	Manual or static	[56,54,51,29,28,68,38,25,43,52,44,49,41,45]
	Clustering	[64,28,68,62,41,45]
	Edge detection	[58,28,68,25]
Other	Image apriori	[36,50,38]
	Background subtraction	[30,31,35,34]
	Snake	[43,66]

When segmentation is irrelevant, it can be asked to the user as in [37,25,26,48,39,55,32,40,33]. An interface is shown to the user so that he or she can select the region of interest by drawing its outlines. Of course, manual segmentation might be a tedious and time-consuming task considering how numerous the images can get. That is why segmentation is worth automated most of the time.

Some segmentation methods rely on thresholding which is basically splitting the image histogram into several groups which correspond respectively to the object and the background. The simplest way of doing thresholding is to set the intensity value that separates the two groups. This value can be set statically in the program or by the user who can select the one which yields the best result as in [53,54].

Another way of performing thresholding is to see it as a clustering problem where two or more clusters (which are the regions) have to be formed [64,28,68,62,41,45]. Otsu's method criterion is about choosing the clusters such that the intra-cluster variation is minimized while the inter-cluster one is maximized [41,45,68]. *k*-means is used onto the color space to search for centroids representing the different regions in the image based on color similarity [62,64]. [28] uses ISODATA, a clustering algorithm that builds clusters with a given standard deviation threshold. [41,45] use meanshift clustering in the color space as a preprocessing step to Otsu.

Some articles build their segmentation over edge detection. Most of them use Sobel filters to get the gradient of the image and use filling operators to get a first mask [28,68,25]. [58] uses order-statistic filters to get the edges of the images.

There are other techniques used to segment the images. [43,66] use active contours (snake) that take a simple thresholding mask as a seed point to get a more accurate segmentation. [30,31,35,34] use background subtraction (which implies the background to be constant).

Finally some studies base their segmentation pipelines on image apriori. [38] assumes the object in the image yields the longest outline while [36] searches for lines on wing images and [50] uses the symmetry of lepidoptera as a detection criterion.

4.2. Three general feature extraction methods

For decades, constructing a pattern-recognition or machine-learning system required careful engineering and considerable domain expertise to design a feature extractor that transformed the raw data (such as the pixel values of an image) into a suitable internal representation or feature vector from which the learning subsystem, often a classifier, could detect or classify patterns in the input. To increase feature performance, learning mid-level features on top of handcrafted ones has been put forward. During the last decade, the amount of human labor has been reduced by representation learning. Representation learning is a set of methods that allows a machine to be fed with raw data (pixels) and to automatically discover the represen-

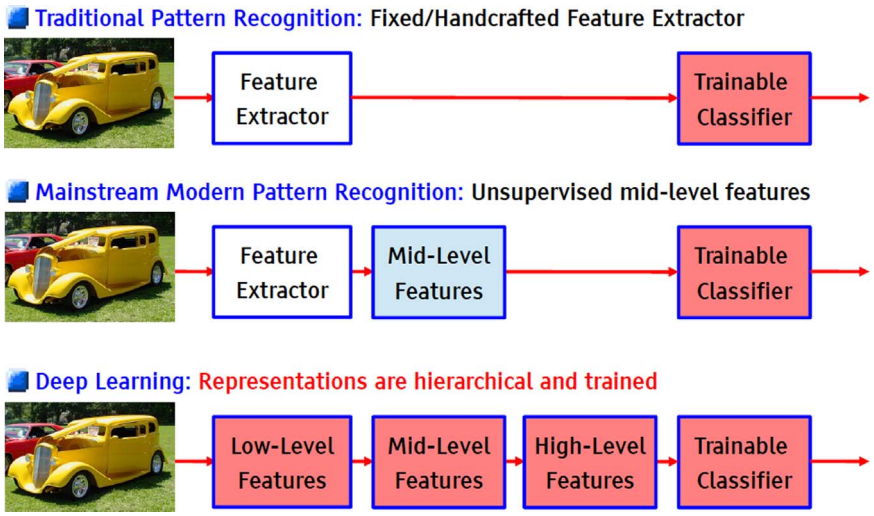

Fig. 3. Feature extraction evolution [67].

Table 3
Feature taxonomy for insect recognition.

Category Levels			Refs.
Handcrafted features	Domain-dependent	Wing' Venations	[36,38,48]
		Geometry	[25,53,54,26,40,49,63]
	Global and generic image features	Shape	[66,29,58,56,43,30,52,44,26]
		Color	[56,41,43,28,30,52,26]
		Texture	[33,41,43,30,39,52,51,26]
		Raw Pixel	[46]
	Local features	SIFT	[19,42,65,47,57,60,61,59]
		Others	[28,50,27,45]
Mid-level features	Unsupervised representations	BoW	[19,57,60,61]
		PCA	[32,55,31,35,34]
	Supervised representations	MLP	[56,33,46]
		Sparse Coding	[27,42,47,64]
Hierarchical representations	Auto-encoder		[52]

tations needed for detection or classification. Fig. 3 illustrates the concepts of handcrafted, mid-level and learned features. In Table 3, the papers from the literature are organized according the following criteria: a) Handcrafted features; b) Mid-level features; c) Learning Hierarchical Representations.

4.3. Fixed/handcrafted feature extractor

4.3.1. Domain-depend features

In [36], the feature is based on the venations present in the bees' wings, and the cells they enclose, to generate numerical feature vectors. Authors in [36] could say the venations in the bees' wings are almost as fingerprints. In the same direction, [38] the coordinates of some characteristic points on the wing are used to compare vein patterns. In [49], owlflies wings are described by Elliptic Fourier coefficients. The feature extraction algorithm incorporates prior expert knowledge about wing veination. More recently, [48] studied the impact of feature selection on top of descriptors presented in [38].

In [53,54,26,25], series of geometrical features including area, perimeter, holes number, eccentricity and roundness have previously been tested. All these features are intuitive because they can be directly measured or simply calculated from images. However, these precisely extracted features are easily affected by factors such as the posture of insects or the shooting angle. Furthermore, it is usually difficult to compute the real size of insects from the images because of the lack of some shooting parameters such as object distance.

4.3.2. Global image features

General image features are generic descriptors that can be used to represent any kind of images not specifically insects. However, in the literature there is distinction between local and global features whether features ares extracted locally on specific area of the image or globally on the whole image.

In [66] Hu moments (Hu), Elliptic Fourier Descriptors (EFD), Radial Distance Functions (RDF) and Local Binary Patterns (LBP) are extracted as shape descriptors. LBP features seem to outperform the rest of the features in recognition rate based on the individual performance of each method. The results from the underlying features are then fused using weighted majority voting to obtain a decision. In [26], a series of shape, color and texture features was developed that draw on CBIR and allow the identification of butterfly images to the taxonomic scale of family. In [64], the objective is to detect pests into an image, and $L^*a^*b^*$ color space is used as features.

In [43,41,30,56] global features are used to describe the entire insect. A set of 54 features including geometric features of insect, contour features, moment features, texture features and color features are extracted. In [29,58] features were extracted from the shape descriptors computed on the binary mask of the Region Of Interest (ROI). The measurements taken were area, convex area, eccentricity, major axis length, minor axis length, perimeter, solidity, roundness, compactness, extent, aspect ratio. In [33,51] wavelet decomposition of the images is performed.

4.3.3. Local features

A crucial step for identification or classification and retrieval is to describe images by extracting a set of local feature descriptors, encoding them into high-dimensional vectors and then fusing them into image-level signatures. In [28], images are preprocessed to detect insect location and then global color features are measured (color frequency histograms). Finally, local color features are extracted from patches. In [50], the core of the system is a probabilistic model that infers Semantically Related Visual (SRV) attributes from low-level visual features of moth image. The main motivation is to obtain a human-readable representation of the image. In [45], the approximate shift invariant property of the dual tree complex wavelet transform (DTCWT) and its good directional selectivity in 2D make it a very appealing choice for insect image recognition. These local and global approaches are combined: a first distance between images is obtained by conventional region-based bag of words techniques and a second

distance is computed through DTCWT. The two distances are then merged together. Finally, the very well-known SIFT (Scale Invariant Feature Transforms) [69] features are extracted in [19,57,61,42,47,65,59]. SIFT features were completed by multiple order gradient histogram (MOGH) in [60]. In the same direction, [27] SIFT features are also completed by with color, texture and shape features. In [59,65], SIFT features are extracted and directly matched between paired-images. The number of positive putative matches is used to estimate the similarity of the two images.

4.4. Mid-level features

In the Digital Automated Identification System (DAISY) [35,34,32,31,55], images are processed using filters to detect edges (top-hat filer) and then noise is removed by applying a Gaussian blur. It adopts a principal component analysis (PCA) method to acquire image features that contain nearly all the information of an image. PCA features are more suitable for species identification because the great amount of detailed information collected with PCA can weaken differences between high-level categories such as orders. The basic algorithm of DAISY is based on the methods for human face detection and recognition via eigen-images. DAISY requires user interaction for image capture and segmentation, because specimens must be aligned in the images. In Russell et al. [39], SPIDA (Species IDentification Automated) was created as a ground spiders identification system. SPIDA's feature vector is built using images of the spider's external genitalia from a subset of components of the wavelet transform using the Daubechies 4 function [70]. This method also has the drawback that the specimen has to be manipulated by hand, and the image capture, preprocessing, and region selection also require direct user interaction.

There is a large body of literature [19,57,61,60] showing the successful application of the bag-of-keywords approach. In this approach, local-feature data is extracted from the available images and characterized by computing a feature vector called a descriptor. A visual dictionary is created by clustering a subset of the training data in the descriptor space, creating clusters that define dictionary words. This dictionary is then used to map the region descriptors of a novel image into words, and to map the bag of features for an image into a bag of words that is used to construct a representation of the image. Histograms of these words have been successfully used as intermediate image representations. Once an intermediate representation is constructed, the remainder of the training data is used to train a classifier to recognize object classes.

In this series of studies regarding insect identification, the progression of recognition methods from simple pattern recognition approaches to unsupervised and later discriminative dictionaries can be seen.

To improve the classification accuracy, [42,47,27] develop an insect recognition system using advanced multiple task sparse representation and multiple-kernel learning (MKL) techniques. As different features of insect images contribute differently to the classification of insect species, the multiple-task sparse representation technique can combine multiple features of insect species to enhance the recognition performance.

Another way to build discriminative mid-level features is to feed a Multi Layered Perceptron (MLP) with handcrafted features as in [33,56] or directly from the raw pixels as in [46]. However the number of feature abstraction layers is limited to two because of the vanishing gradient problem.

4.5. Learning hierarchical representations

Over the last few years, a large amount of research on visual recognition has focused on learning low-level and mid-level features using unsupervised learning, supervised learning, or a combination of both. The ability to learn multiple levels of good feature representations in a hierarchical structure helps to construct sophisticated recognition systems.

In [52], series of geometric, shape, invariant moment, texture and color features are given as an input to a Stacked Denoising Autoencoder. This deep learning architecture consists of multiple levels of non-linear operations and is an effective way to represent high-level abstractions. Automatically learning features at multiple levels of abstraction allows a deep learning system to learn complex functions mapping the input to the output directly from data [71]. The goal in [52] is to solve the pose variety problem. Compared with general auto-encoder (AE) architecture, DAE architecture adds the noise to the training data, which means that the DAE learns to detach noise from real data. It makes the encoder learning more robust to the expression of the input signal.

4.6. Advantages and drawbacks

The first attempts of image-based recognition [53,54] used morpho-metrics as features. Morpho-metrics are distances or ratios between venations present on the wings. These morpho-metrics were extracted and used to separate species using Linear Discriminant Analysis. Morpho-metrics being specific to winged species, more generic approaches were to be found. [35,34,32,44,36] established the DAISY system to classify wasp insect images using PCA. To improve classification accuracy, [55] applied DAISY to recognize insect images by analyzing their wing patterns and shapes. [36] proposed an improved ABIS (Automatic Bee Identification System) system using support vector machine (SVM) and kernel discriminate analysis based on geometric features of wings (such as length, angle, and area). Moreover, many literature works have focused on constructing object appearance models, a key part of object classification. Generally, based on their appearance models, most object feature descriptors can be categorized as either global features or local features. [39] adopted global features (including color, texture, and geometry) to classify insect images and obtained good results using high-quality images. However, because the features are very sensitive to rotation, scale, translation, and viewpoint changes, this classification method did not work well on objects with large intra-species variation or high inter-species similarity. To address these issues, [61] developed a local feature-based insect identification scheme to account for variations in insect appearance. Furthermore, [62] devised an image-based automated insect identification and classification method using three models: an invariant local feature model, a global feature model, and a hierarchical combination model. [65] extended the LOSS algorithm [68] for analyzing the geometrical characteristics of insects to improve insect classification. A drawback is the limited expressiveness of these models because insect appearances can only be represented by handcrafted descriptors which makes it difficult to handle significant insect view and pose changes simultaneously. For the same reason, some local features based on SIFT such as CFH [57], bag of words [61,19,61,60] or ScSPM [42] that are currently used in insect species recognition are not quite suitable in high-level (order level) insect identification. It is well known that natural images can be sparsely represented using a sparse linear combination of a few elements from a trained dictionary. In contrast to most existing insect-classification methods that directly operate on low-level features or cues, sparse coding can learn insect appearances from raw features to quantify insect appearances by means of sparsity [42,47,27].

Over the last few years, in [52] the ability to learn multiple levels of good feature representations in a hierarchical structure helps to construct sophisticated recognition systems. The benefit of the sophistication of recognition systems is the release of the imaging constraints. The perspective is therefore about handling more and more noise and variations so that both the capture step gets less and less tedious and the classification gets improved in handling more complex class

configurations.

5. Classification methods

From the perspective of classification, we used the following pieces of information: the granularity of the classification (the scale in the taxonomy tree whether it's to a species, families, order or another level precision), the number of classes, classifier(s) used, type of validation and the accuracy of the results.

The original pattern recognition framework assumes there is a clear boundary between extracting information from raw data and classification. The algorithms fulfilling these two tasks are thought as independent processes and seen as blackboxes from one another point of view. However, the boundary is becoming blurrier with some recent work (see representation learning [72] for example). In these approaches, the features and the classifier function are learnt at the same time during the learning stage. However, most of the papers studied in this article keep a sharp boundary between feature extraction and classification. In this section, the classification stage is investigated as well as the way it has been applied to insect image recognition until now.

As for the previous chapters, we consider the papers under a certain categorization, which is depicted in Table 4: a) Monolithic classifiers; b) Combinations; c) Instance-based.

5.1. Monolithic classifiers

The basic way to do classification in pattern recognition is to train a single system on the actual data.

Classification is the process of actually recognizing an individual as belonging to a certain population among others. There are two ways to see classification conceptually: discriminatively and generatively speaking. This will be discussed about along the next paragraphs.

5.1.1. Discriminative methods

As a discriminative process, classification is setting boundaries on a given feature space between the different populations in order to distinguish them. Stochastically speaking, it is said that discriminative approaches try to model $p(C = c|x)$, i.e. the probability of occurence of class c given x (features vector).

A first approach to discriminative classification is to find a linear boundary between classes. [43,61] uses the least squares approximation method to find this boundary. In [54,53,38,40,63], Fisher criterion (Linear Discriminant Analysis) is used to the same end on morphometrics. However, data may not be linearly separable and therefore linear discrimination is likely to fail. In [30,49,47,51,42,27,61,28,48,25] SVMs are in use. While [51,47,42,28] use linear SVM, [61,30,49,48] use kernelized SVMs. [48,61] use SVMs with the polynomial kernel. [30,49,25] use the gaussian Radial Basis Function of standard deviation σ as a kernel. σ is considered as a free parameter in the frame of such a use. The usual way to tune σ is trial and error. However, [73] proposes a genetic algorithm to find this parameter which is used in [49]. [27] uses Multiple Kernel Learning, which is a SVM using an optimally chosen combination of different kernel types in order to efficiently learn on different feature modalities.

As used in [52,33,58,46,44,48,25], neural networks are another discriminative approach where non-linearity is induced through activation functions. In [46], a neural network is trained with scaled conjugate gradient-based backpropagation. Several parameters and network sizes (from no to two hidden layers) are tested on the following recognition task: discriminate a single pest from many non-damaging insects. While a one-hidden-layer network is proven to be a universal approximator, the authors couldn't get satisfying result with one single hidden layer. A possible explanation is, because the inputs are raw pixel, the MLP must learn a feature level. Finally, [44] uses Quality Threshold ARTMAP, which is an incremental learning network technique. An ARTMAP network is a combination of unsupervised modules (which are self-organizing maps): one operating on features, one on the classes and a last one in between.

Decision trees are used in [28,48,57]. [28,48] use classic decision tree generated with C4.5 algorithm. [59] uses Logistic Model Trees which differ from classic trees in that its nodes are logistic regression classifiers.

5.1.2. Generative methods

From a generative perspective, classification is trying to estimate the likelihood for an individual to belong to a given population. To keep using the probabilistic point of view, generative approaches are modeling $p(x|C = c)$, i.e. the probability of occurence of feature vector x if considering the class c.

[59,43,28,48] use parametric probability density functions to estimate $p(x|C = c)$ for each class c. [59,48,28,43] use Naïve Bayes classifiers which assume the features are statistically independent. Logistic classifiers (used in [48]), however, don't postulate on statistical independance. Rather than trying to find a distribution by tuning probability density functions, non-parametric methods rely on the instances themselves. [43,61] use Parzen-Rosenblatt window classifiers to estimate the probabilities.

Other generative methods are used besides probability density estimation. In its first version, DAISY [31,32,35,34] uses a non-parametric statistical test to evaluate the quality of the PCA reconstructed image with respect to the original image. The test used is Kendall rank correlation coefficient (more commonly referred to as the τ-test). [64] uses a correspondence filter to detect a given shape and localize it on image. It is a frequency domain filter which is computed through an optimisation process such as a linear optimisation program. The idea is to create such a filter so that the peak of correlation and the localisation of the object we search in the image match up.

Table 4
Taxonomy of the reviewed papers on the basis of classification.

Category Levels			Refs.
Monolithic	Discriminative	Linear	[54,53,38,40,63]
		SVM	[30,49,47,51,42,27,61,28,48,25]
		Neural nets	[52,33,58,46,44,48,25]
		Decision trees	[28,48,57]
	Generative	Parametric	[57,43,28,48]
		Density estimation	[43,61]
		Other	[31,32,35,34,64]
Combinations	Committees		[66]
	Boosting and bagging		[19,57,28]
	Other		[33,39,29,62]
Instance-based	k-NN		[66,61,43,28,26,45,41,65]
	Other		[50]

5.2. Combinations

But when it comes to issues such as overfitting, a monolithic classifier is somehow not enough and that is here where ensemble learning is relevant. The idea of ensemble learning is to use a combination of said weak classifiers to get a strong classifier [74]. Classifier combinations also aim at improving simple classifiers performances.

In [66] four k-Nearest Neighbour classifiers are trained on different feature descriptors. Their decisions is combined into a weighted majority vote: it is a committee of k-NN classifiers. [19,57] combine decision trees in more complex ways. In [57], logistic model trees are boosted with AdaBoost. The classifiers are incrementally trained, each

classifier being trained on classification error weighted according to the previous classifier error so that it can learn more where it was weaker. [29,60] use several classifiers to perform dichotomies on the dataset. In [60], the decision is made by making SVMs compete against each other in a tennis-tournament-like process.

SPIDA [33] not only experiments flat neural network classifiers but also tests multi-granularity discrimination: a first neural net discriminates individuals at the genus level before a second one performs likewise at the species-level in the actual genus. The system is found to have no difficulty to discriminate between genuses but some between species inside different genera. In [39] a different structural approach is adopted: A set of neural nets were trained (one per taxon). Each neural network has two outputs: one that is trained to be stimulated when the individuals belongs to the taxon (known as pro-species output) and the other when it's not (the anti-species output). The final decision is given as follows: the top-3 scores from all the pro neurons are considered and those who are above a given threshold are considered as answers and shown to the user. In [29], a stacked ensemble of neural networks is used to get satisfying classification results (compared to those of a single neural net classifier). It is in fact a decision-tree where every node decision is given by a neural network. [62] uses a different framework which is based on recognition relevance. If a first classification correlation level is not high enough, a second classification is operated. It stacks up a global feature classifier (based on Bayes classifiers) on top of a local feature classifier (with Nearest Mean classifier) in order to become invariant to pose variabilities. The global classifier is first called and, if the matching probability is too low (under a given threshold), the local feature classifier is called to give the final decision.

5.3. Instance-based

The previously mentioned methods rely on characterizing classes in the feature space either generatively or discriminatively. Another way of tackling the classification problem is to compare the unknown image not to classes themselves but to known occurences (which turn out to be the learning images).

[66,26,45,41,65] use nearest neighbour classification to get a final decision. [26] uses a nearest neighbour classifier with a weighted L_2 norm: the authors want to test the classification within modifying the weights of the different features. [41,45,65] base their approaches on refining a candidate list at each level. A first classifier gives a match list based on some similarity measure while a second one refines that list. [41] experimented a two-level classification workflow. The first level is a coarse matching step where a histogram similarity measure is done and candidates are filtered based on a given similarity-threshold value. The second level classification is done by a k-NN classifier which performs on some texture features. [45] first uses an integrated region matching process on clustered image to find k image candidates. The final answer is given by a Nearest Neighbour classifier.

[50] relies on a CBIR approach which is not about getting a final answer but just providing a match list to the user. The decision task falls therefore to them. In the scope of this study, Relevance Feedback is used as a performance booster. The user sends a query and is asked about the answer relevance. If trust has been granted to the user (because he passed a recognition test successfully), the system distorts the feature spaces so that the non-relevant items are further away from each others.

5.4. Methods comparison studies

In some studies, features are extracted from the images and the choice of a classifier is not really guided. They usually take a set of classifiers to apply to extracted features and keep the one that maximizes the recognition rate on sample data. Results of some of the tested classifiers are somehow investigated. Table 5 details the different methods used in each study and gives the recognition rates per method.

Table 5
Recognition rates in comparison studies (bold figure for best rate).

	[61]	[43]	[28]	[48]	[25]
Least Square	84.2	83.3			
LDA				64.0	
SVM	88.4		**85.0**	60.0	**92.0**
Decision tree			58.3	36.0	
MLP				58.5	76.0
Bayes		**89.9**	65.9	**65.2**	
Logistic				63.2	
Parzen density	86.3	84.7			
Random forest			83.2		
k-NN	77.4	83.8	71.6		
Nearest Mean	**89.5**	63.6			

5.5. Advantages and drawbacks

Although discriminative methods are powerful tools to accurately split the feature spaces, their lack of incrementality can be a restriction in the scope of broad insect recognition where new species can be encountered. Generative methods are more fitted to this end since the classes probabilities are considered independently. Adding a class is then not a problem.

When it comes to evolving classes (unknown individuals encountered that belong to already existing classes), instance-based are handier since the new individuals can be integrated to the instance. Going further, the nature of insect recognition as a pattern recognition problem can be called into question. Given the potentially high cardinality of the classes set and the high intra-class variability, it might be more relevant to try to match the unknown individual to a known one than to an entire class. This feature is interesting from the user perspective because they get the closest known individuals instead of the class information. The problem could then be seen more as a metric learning one. In such a frame, the similarity function is to be learnt so that the individuals in a class are close to each others and as far as possible to other classes in the similarity space. The learning process would therefore be class-agnostic in the sense the learning phase only needs to know which individuals pairs are similar and which ones aren't. The class is induced by the individuals and their labels.

6. Datasets

To characterize the datasets, we noted the number of instances (images) per class and overall as well as per training, validation and testing subset. The granularity is a good indicator as well. Granularity is the scope in the taxonomy where a study is located. It can be as specific as to a few species in a single family or as wide as over several orders (see taxonomical scales in Fig. 1). Every dataset used in the studied literature are listed in Table 6.

The first comment one can make is that the granularity is often very thin and the variability very subtle. We most of the time talk about species in a few specific orders or families. In fact, studied sets are tightly bound to the projects biological and environment-related problems such as recognizing a very specific group of pests for a given orchard. Some studies are in the frame of Integrated Pest Management [75] Pests for different orchards were investigated such as rice [30], palm [46], apples [56] cotton [29] or miscellaneous vegetables [64,65,68] pests. [52,61,43,37,66] focus on moths. Some others deal with biodiversity assessments like the BugID project [19,57,59,60,76,77] which focus on aquatic myriapods. These insects, when found in rivers, are used to perform bioassessments regarding water quality. [58] does likewise to assess marine life by recognizing copepods populations. ABIS [36] aims at quantifying bees populations

Table 6
Datasets used in the selected studies.

Insects	Granularity	Number of taxons	Avg. Nb. of images per taxon	Capture type	Constrained pose?	Part of insect	Publicly available	Used in	Best reco rate (* means avg reco rate)
Lepidopterans	species	100	4	lab-based	Yes	Wings	No	[41,45]	85.00%
Ichneumonidae	species	5	50	lab-based	Yes	Wings	No	[31,34,32]	94%
Midges	species	49	20	lab-based	Yes	Wings	No	[32,35]	86%
Macromoths	species	237	3	lab-based	Yes	Whole	No	[28,37]	85.00%
Moths, but-terflies and caterpillar	species	4500	8	lab-based	Yes	Whole	Yes(http://janzen.sas.upenn.edu/)	[50] (par-tial use)	53.00%
Leafhoppers	species	3	N/A	lab-based	Yes	Whole	No	[54]	73.00%
Copepods	species	8	30	lab-based	No	Whole	No	[58]	98.00%
Miscellaneous orchard insects (multiple poses)	species	9	162	lab-based	Yes	Whole	No	[61,62,52] (par-tial use)[43]	96.60%
Miscellaneous orchardinsects (trap-based)	species	5	128	lab-based	No	Whole	No	[52,62](partial use)	94.10%
Tephritidae	species	20	7	lab-based	Yes	Whole or	No	[42,47]	88.50%
Many orders	orders	9	123	lab-based	Yes	Whole	No	[26]	32%
Apple pests	species	6	200	lab-based	No	Whole	No	[56]	99.50%
Butterflies	species	9	25	lab-based	Yes	Whole	No	[25,27]	90.30%
Miscellaneous insects	species	20	12	lab-based	Yes	Whole	No	[44]	99%
Benthic macro-invertebrates	species	8	191	lab-based	No	Whole	No	[60]	88%
Palm tree in-sects	species	21	20	lab-based	Yes	Whole	No	[46]	95%*
Cotton insects	species	12	N/A	multi-individuals	No	Whole	No	[29]	75%*
Rice insects	species	4	195	multi-individuals	No	Whole	No	[30]	97.5%*
Stoneflies	species	9	425	lab-based	No	Whole	Yes (http://web.engr.edu/~tgd/bugid/stonefly9/)	[59,19,57]	96.50%
Ephemeroptera, Plecoptera and Tri-choptera	species	29					Yes (http://web.engr.edu/~tgd/bugid/ept29/)		
Spiders	species	13	974	lab-based	Yes	Whole	No	[39]	100%
Lycosidae	species	6	18	lab-based	Yes	Whole	No	[33]	100%
Several pests	species	6	160	multi-individuals	No	Whole	No	[65]	N/A
Field crop in-sects	species	24	26	field-based	No	Whole	No	[27]	89.50%
Moths	species	N/A	6	field-based	No	Whole	Yes (http://www2.ahu.edu.cn/pchen/web/insectRecognition.htm)	[66]	88%
Several pests	species	N/A	N/A	field-based	No	Whole	No	[64]	N/A
Wasps and dragonflies	species	3	8 40	lab-based	Yes	Wings	No	[38]	N/A
Honeybees	subspecies	26	70	lab-based	Yes	Wings	No	[48]	65%
Bombus bees	species	8	65	lab-based	Yes	Wings	No	[36]	95%
Various bees	subspecies	5	448	lab-based	Yes	Wings	No	[40]	97.50%
Owlflies	species	7	17	lab-based	Yes	Wings	No	[49]	100%
Fruitflies	species	72	25	lab-based	Yes	Wings	No	[51]	86.20%
Mosquitoes	species	79	100	lab-based	Yes	Wings	No	[51]	80.30%

regarding their different species. Bees are the main pollinators and are therefore essential for crops to grow.

Some studies deal with coarser problems [25,26,56] such as families of butterflies [26] or different orders [25] while some studies test the performances of their methods on different granularity levels [33].

There are almost a dataset per article, in most cases because every study has its focus on both the capture of images and the insect taxons involved. The result here is we miss a reference set.

7. General discussion

Automated insect identification has been intensively studied over the last three decades, including computer vision-based systems for the classification of insect species.

From all the considered works emerge the following questions:

1. What is the least imaging effort to make in order to get satisfying results?
2. How to define the problem with respect to the taxonomy tree?
3. Is image-based insect recognition more than an application of image classification?

7.1. What is the least imaging effort to make?

The first approaches of insect-image classification used very constrained and low-tolerance imaging protocols such as with flatbed wing images. Then came methods using region-of-interest segmentation on plain or weakly cluttered backgrounds. Finally, the plunge towards dealing with complex-background images has yet to be taken although some local methods have been used to this end.

Considering generic image classification, Convolutional Neural Networks have been acknowledged as state of the art in a quite short span of time regarding their high and surpassing performance in terms of both accuracy and invariances. In the set of studies being scrutinized here, none has experimented on-pixel classification further than through classical shallow MLP [46] on one hand. On the other hand, deep hierarchical approaches aren't represented further than Stacked Denoising Autoencoders in [52].

However, important characteristics regarding deep learning in general should be taken into account regarding the insect-image classification problem: learning deep hierarchical representations requires a notably larger set of example images than for shallow learning. Assuming the insect images harvest is a more tedious task than for generic images (which is tedious enough for getting millions of images), this issue is the main limitation to the use of CNNs. Finally, no article applied the transfer learning paradigm in order to take the benefit of a pre-trained deep hierarchical representation. Using knowledge gained on larger but generic dataset such as ImageNet as a learning basis to recognize insects could be a way to tackle the image volumetrics problem.

7.2. How to define the problem with respect to the taxonomy tree?

As for any other classification task, the problem has to be modeled considering both the classes and the relations between them. Are the classes close to each other? Do they cover variability? Does that variability make the classes cross each other? To sum up, what are the variabilities both between the categories (*inter-class*) and inside (*intra-class*) them?

The first intuition is these properties vary along the taxonomy tree (see Fig. 1). The smaller the scale is, the finer the intra-class variations. Small scale insect recognition problems tend to be fine grained object recognition problems. At this scale, the main issue is the intra-class variations can be as strong as and even stronger than the inter-class ones [61,62,52,57,39].

Does it go to less fine-grained when we climb up the tree towards a coarser level? Even though macro-taxons are not formed on morphological criteria, the genetical proximity induces likelihood of visual similarity. [33] genus-level results are way superior than the species-level ones. Even further, their hierarchical experiment seems to yield good results: a first set of neural networks identifies the genus and a second set deals with contained species. This assertion has to be taken carefully due to the low number of classes that are tested here.

On another point of view, natural phenomena (biodiversity resulting in inter-class similarities and intra-class differences) bring more difficulties for insect identification at the order level than that at the species level [25]. However, it is most often found that the similarities in species within a single family and the differences among species of separate families are prominent enough to [allow] identification [25]. This assumption calls another one: To make upscaled insect recognition efficient, every sub-class (or sub-taxon) must be collected. Same observation is made in [26] at the family level: higher results are observed for higher taxon coverage.

The macro-groups cannot be considered as uniform populations where the individuals are equivalent. These groups must be learnt by considering their respective subgroups equally.

7.3. Is image-based insect recognition more than an application of image classification?

This question could be rephrased as “Are there intrinsic properties to insect images so that insect recognition is a specific computer vision problem in itself? ” To provide the beginning of an answer, the characteristics regarding insect recognition must be considered. Aside from classical object image variations, pose is a fairly critical variation type that is tackled in [52]. Deformation is one potential characteristics of these objects due to articulated parts [57,60]. However, these variations can occur under equivalent forms in other objects categories. To take the example of face recognition, age is an object-specific variation factor which can be compared to the insect age. Regarding the visual statistical and structural properties of such images, apparently no study investigated these aspects.

Observations on the classification perspective emerge from the conclusions in Section 7.2: taxonomy does not reflect visual similarity at any location. This heterogeneity creates complex classes configurations at a coarse level as previously shown but also brings about the dynamics of a potential evolving classification tool. In a context where specimens are being encountered along time, i.e. the construction of the classes is incremental in some sense, new groups of specimens might appear at any moment. As a consequence on a fine level, class are likely to add up when unencountered taxons appear. A dual assertion can be made at the coarse level: when sub taxons add up and since taxonomy is not heterogeneous, the macro-taxon is incrementally defined by its sub-taxons. Depending on the granularity, the problems can reveal two kinds of incrementality : class add-ups and incremental class definitions. These systemic concerns have not been addressed yet in the literature since the datasets are fixed.

Even though this problem has interesting properties, its singularity regarding pattern recognition and image analysis remains unsure. Besides, generic methods are the most used methods in the recent works and it will not be possible to compare methods as long as the experiments are not performed on a common dataset.

Conflict of interest

None.

Acknowledgement

This work was supported by a research grant from the Région Centre-Val de Loire, France. Thanks go to Prof. Hubert Cardot,, Dr.

Nicolas Ragot and the rest of the RFAI team for their feedback.

References

[1] D. Pisani, L.L. Poling, M. Lyons-Weiler, S.B. Hedges, The colonization of land by animals molecular phylogeny and divergence times among arthropods, BMC Biol. 2 (2004) 1.

[2] J.-P. Lin, S.M.G. Iii, J.G. Gehling, L.E. Babcock, Y.-L. Zhao, X.-L. Zhang, S.-X. Hu, J.-L. Yuan, M.-Y. Yu, J. Peng, A parvancorina-like arthropod from the cambrian of south china, Hist. Biol. 18 (1) (2006) 33–45.

[3] S.A. Elias, Quaternary Insects and their Environments, Smithsonian Institution Press, Washington, USA, 1994 https://www.cabdirect.org/cabdirect/abstract/19941106757.

[4] F. Ødegaard, How many species of arthropods? Erwin's estimate revised, Biol. J. Linn. Soc. 71 (4) (2000) 583–597.

[5] P.M. Hammond, Species inventory, in: G. B (Ed.), Global biodiversity, Chapman and Hall, London, 1992, pp. 17–39.

[6] D.M. Rosenberg, H.V. Danks, D.M. Lehmkuhl, Importance of insects in environmental impact assessment, Environ. Manag. 10 (6) (1986) 773–783.

[7] C. Kremen, Assessing the indicator properties of species assemblages for natural areas monitoring, Ecol. Appl. 2 (2) (1992) 203–217.

[8] A. Finnamore, The Advantages of Using Arthropods in Ecosystem Management, Biological Survey of Canada (Terrestrial Arthropods) for Canadian Museum of Nature and Entomological Society of Canada, Ottawa.

[9] N.M. Van Straalen, Evaluation of bioindicator systems derived from soil arthropod communities, Appl. Soil Ecol. 9 (1) (1998) 429–437.

[10] J. Niemelä, D. Langor, J.R. Spence, Effects of clear-cut harvesting on boreal ground-beetle assemblages (Coleoptera: carabidae) in western Canada, Conserv. Biol. 7 (3) (1993) 551–561.

[11] M. Pollet, P. Grootaert, An estimation of the natural value of dune habitats using Empidoidea (Diptera), Biodivers. Conserv. 5 (7) (1996) 859–880.

[12] J.J. Rykken, D.E. Capen, S.P. Mahabir, Ground beetles as indicators of land type diversity in the green mountains of vermont, Conserv. Biol. 11 (2) (1997) 522–530.

[13] R. Kitching, A.G. Orr, L. Thalib, H. Mitchell, M.S. Hopkins, A.W. Graham, Moth assemblages as indicators of environmental quality in remnants of upland Australian rain forest, J. Appl. Ecol. 37 (2) (2000) 284–297.

[14] H. Gibb, D.F. Hochuli, Habitat fragmentation in an urban environment large and small fragments support different arthropod assemblages, Biol. Conserv. 106 (1) (2002) 91–100.

[15] A.N. Andersen, The use of ant communities to evaluate change in Australian terrestrial ecosystems: a review and a recipe, in: Proceedings of the Ecological Society of Australia, vol. 16, 1990, pp. 347–357.

[16] J.H. Lawton, D. Bignell, B. Bolton, G. Bloemers, P. Eggleton, P. Hammond, M. Hodda, R. Holt, T. Larsen, N. Mawdsley, et al., Biodiversity inventories, indicator taxa and effects of habitat modification in tropical forest, Nature 391 (6662) (1998) 72–76.

[17] D.N. Kimberling, J.R. Karr, L.S. Fore, Measuring human disturbance using terrestrial invertebrates in the shrub-steppe of eastern Washington (USA), Ecol. Indic. 1 (2) (2001) 63–81.

[18] W. Büchs, Biodiversity and agri-environmental indicators—general scopes and skills with special reference to the habitat level, Agric. Ecosyst. Environ. 98 (1) (2003) 35–78.

[19] M.J. Sarpola, R.K. Paasch, E.N. Mortensen, T.G. Dietterich, D.A. Lytle, A.R. Moldenke, L.G. Shapiro, An aquatic insect imaging system to automate insect classification, Trans. ASABE 51 (6) (2008) 2217–2225.

[20] Y. Chen, A. Why, G. Batista, A. Mafra-Neto, E. Keogh, Flying insect classification with inexpensive sensors, J. Insect Behav. 27 (5) (2014) 657–677.

[21] Z. Le-Qing, Insect sound recognition based on mfcc and pnn, in: Proceedings of the Multimedia and Signal Processing (CMSP), 2011 International Conference on, vol. 2, IEEE, 2011, pp. 42–46

[22] A.K. Banerjee, K. Kiran, U. Murty, C. Venkateswarlu, Classification and identification of mosquito species using artificial neural networks, Comput. Biol. Chem. 32 (6) (2008) 442–447 (ISSN 1476-9271).

[23] K.J. Gaston, M.A. O'Neill, Automated species identification: why not?, Philos. Trans. R. Soc. B: Biol. Sci. 359 (1444) (2004) 655–667 (ISSN 0962-8436, 1471-2970).

[24] S.N. Asiah Hassan, N.N.S. Abdul Rahman, Z. Zaw Htike, S. Lei Win, Vision based entomology, Surv. Int. J. Comput. Sci. Eng. Surv. 5 (1) (2014) 19–32 (ISSN 09763252, 09762760).

[25] J. Wang, C. Lin, L. Ji, A. Liang, A new automatic identification system of insect images at the order level, Knowl.-Based Syst. 33 (2012) 102–110. http://dx.doi.org/10.1016/j.knosys.2012.03.014 (ISSN 09507051).

[26] J. Wang, L. Ji, A. Liang, D. Yuan, The identification of butterfly families using content-based image retrieval, Biosyst. Eng. 111 (1) (2012) 24–32. http://dx.doi.org/10.1016/j.biosystemseng.2011.10.003 (ISSN 15375110).

[27] C. Xie, J. Zhang, R. Li, J. Li, P. Hong, J. Xia, P. Chen, Automatic classification for field crop insects via multiple-task sparse representation and multiple-kernel learning, Comput. Electron. Agric. 119 (2015) 123–132. http://dx.doi.org/10.1016/j.compag.2015.10.015 (ISSN 01681699).

[28] M. Mayo, A.T. Watson, Automatic species identification of live moths, Knowl.-Based Syst. 20 (2) (2007) 195–202.

[29] H. Gassoumi, N.R. Prasad, J.J. Ellington, Neural network-based approach for insect classification in cotton ecosystems, in: Proceedings of the International Conference on Intelligent Technologies (InTech 2000), Bangkok, Thailand, vol. 7, 2000.

[30] Y. Qing, J. LV, Q.-J. LIU, G.-Q. DIAO, B.-J. YANG, H.-M. CHEN, J. TANG, An insect imaging system to automate rice light-trap pest identification, J. Integr. Agric. 11 (6) (2012) 978–985.

[31] P.J.D. Weeks, I.D. Gauld, K.J. Gaston, M.A. O'Neill, Automating the identification of insects a new solution to an old problem, Bull. Èntomol. Res. 87 (02) (1997) 203–211.

[32] M.A.O'Neill, I. D. Gauld, K. J. Gaston, P. Weeks, Daisy: an automated invertebrate identification system using holistic vision techniques, in: Proceedings of the Inaugural Meeting BioNET- International Group for Computer- Aided Taxonomy (BIGCAT), 1997, pp. 13–22.

[33] M. Do, J. Harp, K. Norris, A test of a pattern recognition system for identification of spiders, Bull. Èntomol. Res. 89 (03) (1999) 217–224.

[34] P.J.D. Weeks, M.A. O'Neill, K.J. Gaston, I.D. Gauld, Species-identification of wasps using principal component associative memories, Image Vis. Comput. 17 (12) (1999) 861–866.

[35] P.J.D. Weeks, M.A. O'Neill, K.J. Gaston, I.D. Gauld, Automating insect identification exploring the limitations of a prototype system, J. Appl. Entomol. 123 (1) (1999) 1–8.

[36] T. Arbuckle, S. Schröder, V. Steinhage, D. Wittmann, Biodiversity informatics in action: identification and monitoring of bee species using ABIS, in: Proceedings of the 15th International Symp. Informatics for Environmental Protection, vol. 1, 2001, pp. 425–430

[37] A.T. Watson, M.A. O'Neill, I.J. Kitching, Automated identification of live moths (Macrolepidoptera) using digital automated identification System (DAISY), Syst. Biodivers. 1 (3) (2004) 287–300.

[38] A. Tofilski, DrawWing, a program for numerical description of insect wings, J. Insect Sci. 4 (1) (2004) 17.

[39] K.N. Russell, M.T. Do, J.C. Huff, N.I. Platnick, Introducing SPIDA-Web: Wavelets, Neural Networks and Internet Accessibility in an Image-Based Automated Identification System, in: N. MacLeod (Ed.), Automated Taxon Identification in Systematics: Theory, Approaches and Applications, chap. 9, Boca Raton, FL: CRC Press, 2008, pp. 131–152.

[40] T.M. Francoy, D. Wittmann, M. Drauschke, S. Müller, V. Steinhage, M.A.F. Bezerra-Laure, D. De Jong, L.S. Gonçalves, Identification of Africanized honey bees through wing morphometrics two fast and efficient procedures, Apidologie 39 (5) (2008) 488–494 (ISSN 0044-8435, 1297-9678).

[41] L.-Q. Zhu, Z. Zhang, Auto-classification of insect images based on color histogram and GLCM, in: Proceedings of the Fuzzy Systems and Knowledge Discovery (FSKD), Seventh International Conference on, vol. 6, IEEE, 2010, pp. 2589–2593.

[42] A. Lu, X. Hou, C. Lin, C.-L. Liu, Insect Species Recognition using Sparse Representation, in: Proceedings of the British Machine Vision Conference, British Machine Vision Association, 2010, pp. 108.1–108.10.

[43] C. Wen, Q. Zhu, Dimension reduction analysis in image-based species classification, in: Intelligent Computing and Intelligent Systems (ICIS), vol. 3, 2010, pp. 731–734.

[44] S.N. Yaakob, L. Jain, An insect classification analysis based on shape features using quality threshold ARTMAP and moment invariant, Appl. Intell. 37 (1) (2012) 12–30 (ISSN 0924-669X, 1573-7497).

[45] L.-q. Zhu, Z. Zhang, Insect recognition based on integrated region matching and dual tree complex wavelet transform, J. Zhejiang Univ. Sci. C. 12 (1) (2011) 44–53 (ISSN 1869-1951, 1869-196X).

[46] S.M. Al-Saqer, G.M. Hassan, Artificial neural networks based red palm weevil (Rynchophorus Ferrugineous, Olivier) recognition system, Am. J. Agric. Biol. Sci. 6 (2011) 356–364.

[47] A. Lu, X. Hou, C.-L. Liu, X. Chen, Insect species recognition using discriminative local soft coding, in: Proceedings of the Pattern Recognition (ICPR), 21st International Conference on, IEEE, 2012, pp. 1221–1224.

[48] F.L. da Silva, M.L. Grassi Sella, T.M. Francoy, A.H.R. Costa, Evaluating classification and feature selection techniques for honeybee subspecies identification using wing images, Comput. Electron. Agric. 114 (2015) 68–77. http://dx.doi.org/10.1016/j.compag.2015.03.012 (ISSN 01681699).

[49] H.-P. Yang, C.-S. Ma, H. Wen, Q.-B. Zhan, X.-L. Wang, A tool for developing an automatic insect identification system based on wing outlines, Sci. Rep. 5 (2015) 12786 (ISSN 2045-2322).

[50] L. Feng, B. Bhanu, J. Heraty, A software system for automated identification and retrieval of moth images based on wing attributes, Pattern Recognit. 51 (2016) 225–241 (ISSN 00313203).

[51] C. Favret, J.M. Sieracki, Machine vision automated species identification scaled towards production levels machine vision ID, Syst. Entomol. 41 (1) (2016) 133–143. http://dx.doi.org/10.1111/syen.12146 (ISSN 03076970).

[52] C. Wen, D. Wu, H. Hu, W. Pan, Pose estimation-dependent identification method for field moth images using deep learning architecture, Biosyst. Eng. 136 (2015) 117–128.

[53] C.H. Dietrich, T.H. Emigh, L.L. Deitz, Morphometric discrimination among females of sibling species of Aconophorini (Homoptera: membracidae), Syst. Entomol. 16 (3) (1991) 311–318.

[54] C.H. Dietrich, C.D. Pooley, Automated identification of leafhoppers (Homoptera: cicadellidae: Draeculacephala Ball), Ann. Èntomol. Soc. Am. 87 (4) (1994) 412–423.

[55] M.A. O'Neill, DAISY: a practical tool for semi-automated species identification, in: N. MacLeod (Ed.), Automated Taxon Identification in Systematics: Theory, Approaches and Applications, chap. 9, Boca Raton, FL: CRC Press, 2008, pp. 131–152.

[56] P. Boniecki, K. Koszela, H. Piekarska-Boniecka, J. Weres, M. Zaborowicz, S. Kujawa, A. Majewski, B. Raba, Neural identification of selected apple pests, Comput. Electron. Agric. 110 (2015) 9–16.

[57] N. Larios, H. Deng, W. Zhang, M. Sarpola, J. Yuen, R. Paasch, A. Moldenke,

D.A. Lytle, S.R. Correa, E.N. Mortensen, L.G. Shapiro, T.G. Dietterich, Automated insect identification through concatenated histograms of local appearance features feature vector generation and region detection for deformable objects, Mach. Vis. Appl. 19 (2) (2008) 105–123. http://dx.doi.org/10.1007/s00138-007-0086-y (ISSN 0932-8092, 1432-1769).

[58] L.K. Leow, L.-L. Chew, V.C. Chong, S.K. Dhillon, Automated identification of copepods using digital image processing and artificial neural network, BMC Bioinforma. 16 (Suppl 18) (2015) S4.

[59] D.A. Lytle, G. Martínez-Muñoz, W. Zhang, N. Larios, L. Shapiro, R. Paasch, A. Moldenke, E.N. Mortensen, S. Todorovic, T.G. Dietterich, Automated processing and identification of benthic invertebrate samples, J. North Am. Benthol. Soc. 29 (3) (2010) 867–874. http://dx.doi.org/10.1899/09-080.1 (ISSN 0887-3593, 1937-237X).

[60] V. Tirronen, A. Caponio, T. Haanpää, K. Meissner, Multiple order gradient feature for macro-invertebrate identification using support vector machines, in: Adaptive and Natural Computing Algorithms, Springer, 2009, pp. 489–497.

[61] C. Wen, D.E. Guyer, W. Li, Local feature-based identification and classification for orchard insects, Biosyst. Eng. 104 (3) (2009) 299–307.

[62] C. Wen, D. Guyer, Image-based orchard insect automated identification and classification method, Comput. Electron. Agric. 89 (2012) 110–115. http://dx.doi.org/10.1016/j.compag.2012.08.008 (ISSN 01681699).

[63] I.Y. Zayas, P.W. Flinn, Detection of insects in bulk wheat samples with machine vision, Trans. ASAE 41 (3) (1998) 883–888.

[64] F. Faithpraise, P. Birch, R. Young, J. Obu, B. Faithpraise, C. Chatwin, Automatic plant pest detection and recognition using k-means clustering algorithm and correspondence filters, Int. J. Adv. Biotechnol. Res. 4 (2) (2013) 189–199.

[65] L.O. Solis-Sánchez, R. Castañeda-Miranda, J.J. García-Escalante, I. Torres-Pacheco, R.G. Guevara-González, C.L. Castañeda-Miranda, P.D. Alaniz-Lumbreras, Scale invariant feature approach for insect monitoring, Comput. Electron. Agric. 75 (1) (2011) 92–99.

[66] H. Yalcin, Vision based automatic inspection of insects in pheromone traps, in: Proceedings of the Agro-Geoinformatics (Agro-geoinformatics), Fourth International Conference on, IEEE, 2015, pp. 333–338.

[67] Y. LeCun, What's Wrong With Deep Learning?, slide #3 — Available at: ⟨http://www.pamitc.org/cvpr15/files/lecun-20150610-cvpr-keynote.pdf⟩, 2015.

[68] L. Solis-Sánchez, J. García-Escalante, R. Castañeda-Miranda, I. Torres-Pacheco, R. Guevara-González, Machine vision algorithm for whiteflies (Bemisia tabaci Genn.) scouting under greenhouse environment, J. Appl. Entomol. 133 (7) (2009) 546–552.

[69] D.G. Lowe, Object recognition from local scale-invariant features, in: Proceedings of the Computer vision, of the seventh IEEE international conference on, vol. 2, IEEE, 1999, pp. 1150–1157.

[70] I. Daubechies, The wavelet transform, time-frequency localization and signal analysis, IEEE Trans. Inf. Theory 36 (5) (1990) 961–1005.

[71] Y. Bengio, Learning deep architectures for AI, Found. Trends Mach. Learn. 2 (1) (2009) 1–127.

[72] Y. Bengio, A. Courville, P. Vincent, Representation learning a review and new perspectives, IEEE Trans. Pattern Anal. Mach. Intell. 35 (8) (2013) 1798–1828.

[73] L. Guo, D.-S. Huang, W. Zhao, The optimization of radial basis probabilistic neural networks based on genetic algorithms, in: Proceedings of the International Joint Conference on Neural Networks, vol. 4, IEEE, 2003, pp. 3213–3217.

[74] L. Rokach, Ensemble-based classifiers, Artif. Intell. Rev. 33 (1–2) (2010) 1–39.

[75] P. Geier, L. Clark, An ecological approach to pest control, in: Proceedings of the Eighth Technical Meeting. International Union of Conservation of Nature and Natural Resources, Warsaw, 1960, pp. 10–18.

[76] E.N. Mortensen, E.L. Delgado, H. Deng, D. Lytle, A. Moldenke, R. Paasch, L. Shapiro, P. Wu, W. Zhang, T.G. Dietterich, Pattern recognition for ecological science and environmental monitoring an initial report, Autom. Taxon. Identif. Syst. (2007) 189–206.

[77] E.N. Mortensen, H. Deng, L. Shapiro, A SIFT descriptor with global context, in: Proceedings of the CVPR 2005, vol. 1, 2005, pp. 184–190.

Maxime Martineau graduated in computer engineering in 2015. He is now a Ph.D. student in pattern recognition and image processing at the LI laboratory, Université François Rabelais of Tours, France.

Donatello Conte received a Ph.D. degree in information engineering from the University of Salerno, Fisciano, Italy in 2006. He is currently Associate Professor of Computer Science with the University of Tours in France and at the LI Laboratory. His current research interests include structural pattern recognition, real-time video analysis, and document images processing.

Romain Raveaux obtained two M.Sc. in 2006, one in networking and telecommunication and another one in computer sciences from the University of Rouen (France). From 2006 to 2010, he had worked on his Ph.D. Thesis at the L3I laboratory of the University of La Rochelle. Since 2011, he is an Assistant Professor at the LI lab of the University of Tours.

Ingrid Arnault obtained a M.Sc. in analytical chemistry and a M.Sc. in the control of insects's populations. She is a researcher ingeneer and the supervisor of CETU Innophyt of University François Rabelais since ten years. She develops applicated projects in biocontrol (macroorganisms, botanicals, biodiversity of Natural ennemies)

Damien Munier obtained a M.Sc. in Insect Science in 2011 from the University François Rabelais of Tours (France). From 2011 to 2016, he worked at CETU Innophyt of University of Tours on biological and integrated control and an innovative method of arthropods classification and biodiversity measures.

Gilles Venturini is a Professor in Computer Science at the Université François Rabelais of Tours, France. He is working on user interfaces between data or data mining tools and users or domain experts. He is also studying 3D data processing. His work include 3D interfaces, virtual reality, data mining and complex data visualization.