

Article

Detecting and Classifying Pests in Crops Using Proximal Images and Machine Learning: A Review

Jayme Garcia Arnal Barbedo 

Embrapa Agricultural Informatics, Campinas, SP 13083-886, Brazil; jayme.barbedo@embrapa.br;
Tel.: +55-19-3211-5880

Received: 30 April 2020; Accepted: 18 June 2020; Published: 24 June 2020



Abstract: Pest management is among the most important activities in a farm. Monitoring all different species visually may not be effective, especially in large properties. Accordingly, considerable research effort has been spent towards the development of effective ways to remotely monitor potential infestations. A growing number of solutions combine proximal digital images with machine learning techniques, but since species and conditions associated to each study vary considerably, it is difficult to draw a realistic picture of the actual state of the art on the subject. In this context, the objectives of this article are (1) to briefly describe some of the most relevant investigations on the subject of automatic pest detection using proximal digital images and machine learning; (2) to provide a unified overview of the research carried out so far, with special emphasis to research gaps that still linger; (3) to propose some possible targets for future research.

Keywords: pest monitoring; machine learning; agricultural crops; insects

1. Introduction

Pests are among the main causes of losses in agriculture [1]. Insects can be particularly damaging, as they can feed from leaves, affecting photosynthesis, and they are also vectors for several serious diseases [2]. There are many chemical and biological methods for pest control [3], but to reach their maximum effectiveness, careful monitoring across the entire property is usually recommended.

In many cases, monitoring is done passively by workers as they carry out their daily activities. The problem with this method is that when the infestation is detected, a lot of damage may have already been done. Early pest detection requires a more systematic approach, especially in large farms. Traps are arguably the most widely adopted tool for systematic pest monitoring [4,5]. If implemented properly, this kind of device can successfully sample insect populations over the entire area of interest. However, without some kind of automation the traps still need to be placed and collected by hand, and infestation evaluation needs to be performed visually, introducing a degree of subjectivity that can lead to a biased assessment of the situation [5,6]. Thus, regardless of the adoption of traps or not, there is a need for methods capable of assessing the status of pests quickly, accurately and autonomously.

Considerable effort has been dedicated to the creation of more effective methods for pest detection and classification. Some techniques try to detect the associated damage instead of the pests themselves [7–9], but direct detection is the predominant approach. Many early studies tried to detect and identify insects by performing an acoustical analysis on the sounds emitted, but interest in this kind of approach seems to have faded in the last decade [10]. Most studies nowadays use digital images for the task [10,11]. While aerial images captured by means of unmanned aerial vehicles (UAV) are being increasingly explored [12,13], in most cases they do not offer enough resolution for the detection of small specimens, and plant canopies may prevent proper detection. Thus, proximal images are still prevalent. Multispectral [14,15], hyperspectral [16,17], thermal [18] and X-ray [10] sensors are

being explored, but conventional RGB (Red-Green-Blue) sensors still dominate due to their low price, portability and flexibility [19].

This article focuses on the combination of digital Red-Green-Blue (RGB) images with machine learning techniques for pest monitoring in the field. There are a few studies dedicated to pest detection in stored products [20–24], but those are not considered here. Also, although the majority of methods for pest monitoring using RGB images employs some kind of machine learning algorithm, there are a few studies that use only image processing techniques such as mathematical morphology [25,26], thresholding [27–29] and template matching [30]. While those methods fall a little outside the scope of this article, they are addressed in the text whenever deemed relevant.

As mentioned before, there are two types of problems tackled by the methods proposed in the literature: detection and classification. Detection refers to the attempt to distinguish a certain target pest from all other elements in an image; it can be viewed as a binary classification (presence/absence of the target). Detection results often are used to estimate the severity of the outbreak. Classification aims to identify different species of pests, which can be viewed as a multiclass detection. Although detection and classification are related problems, they have certain particularities that often guide the choice of techniques used in the algorithms, as emphasized in Section 2.

The creation of an automatic system for pest monitoring can be roughly divided into four stages: data acquisition, model building, encapsulation in a usable tool, and practical use. The vast majority of the studies found in the literature focus on the second stage. In most cases, the first stage is usually only superficially addressed: a description of how the data was collected is usually provided, but a meaningful discussion on the efforts to make the data more representative and on the limitations of the database is rarely present. The third and fourth stages are often beyond the scope of the studies, as the final steps towards practical adoption of the technologies involve aspects that are more related to user experience, marketability, etc.

In this context, although aspects related to all four stages are addressed at some point in the text, this review focuses on the two first stages. It is important to note that the second stage is intimately related to the first. In fact, machine learning tools have evolved to a point in which the data fed to them has a much more prominent role on their success than their intrinsic characteristics, which explains the relatively similar performances yielded by different models when the data used to train and test them is the same [31,32]. Conversely, image sets with different data distributions and variability can lead to significantly disparate performances even the exactly same machine learning model is employed, a fact that is clearly reflected in the accuracies shown in Table 1. This indicates that most of the weaknesses and research gaps that still remain are connected to the quality and variability of the data gathered, and the content of this review reflects this fact.

This article has three main contributions:

- It summarizes the progress achieved so far on the use of digital images and machine learning for an effective pest monitoring, thus providing a complete picture on the subject in a single source (Section 2).
- It provides a detailed discussion on the main weaknesses and research gaps that still remain, with emphasis on technical aspects that discourage practical adoption (Section 3).
- It proposes some possible directions for future research (Section 4).

2. State of the Art of Pest Monitoring Using Digital Images and Machine Learning

An extensive search on the databases Google Scholar, Scopus, ScienDirect and IEEE Xplore was performed to identify as many relevant studies on pest monitoring using digital images as possible. The search was performed using the keywords “pest”, “insect”, “image” and “detection”. These are relatively general terms that yielded a large number of irrelevant results, but it was important to keep the search broad in order to avoid missing relevant matches. Only investigations published in peer-reviewed journals were selected, with the exception of [33], which is a conference article reporting some relevant findings. A second search considering the reference lists of the articles selected in the

first step was also conducted. A total of 34 articles fitted all the searching criteria, while 7 articles that did not employ machine learning techniques but used some interesting strategies were also included. Also, two of the selected articles are literature surveys— Liu et al. [10] reviewed different sensing technologies (acoustic, visible light, multispectral, hyperspectral, X-rays) for detection of invertebrates on crops, while Martineau et al. [11] focused on describing different pest classification strategies found in the literature. Table 1 summarizes all selected investigations, and Figure 1 shows the most common relationships between the types of data extracted from the images and the detectors/classifiers that are adopted.

This section is structured according to the task (detection or classification) addressed by the proposed methods. It is worth noting that a more logical structure in which the methods would be presented as they improved upon the weaknesses of their predecessors was considered. However, because most studies have as main goal to overcome the limitations associated to manual and visual monitoring, instead of targeting current research gaps, following a logical progressive sequence is largely unfeasible. This fact highlights one of the main issues detected in this study: many of the methods proposed in the literature employ similar classification strategies. In most cases, the only major distinction is the species of interest, resulting in studies with significant redundancy and limited novelty. This review identifies many of the research gaps that need to be addressed, which hopefully will serve as inspiration for future efforts.

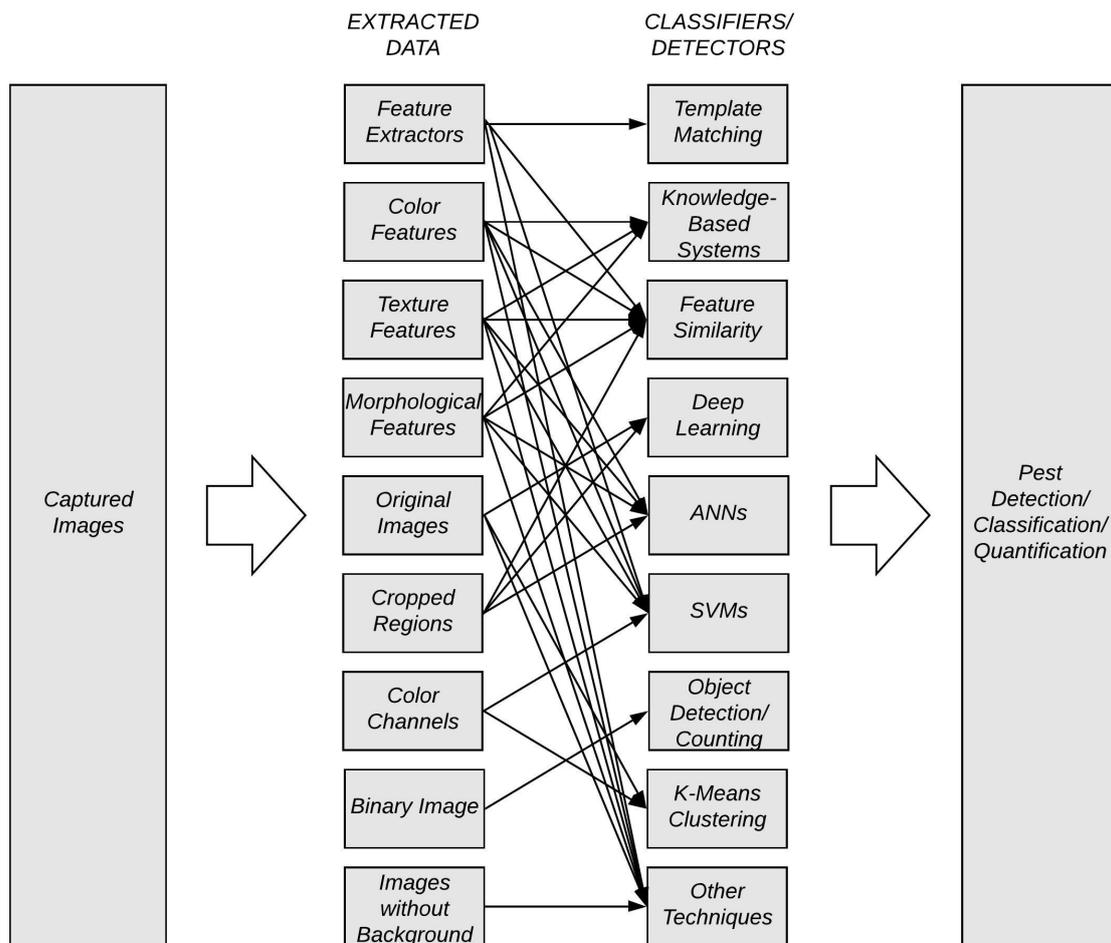


Figure 1. Diagram summarizing the types of data extracted from the original images and which of those are used as inputs to the classifiers and detectors.

Table 1. Basic information about the selected studies.

Ref.	Problem	Image	Pest	Input	Classifier/Detector	Acc.
[30]	Detection	Trap	Red Palm Weevil	Zernike Moments and Region Properties	Template matching	0.88–0.97
[26]	Detection	Controlled	Whiteflies	Binary image	Connected objects counting	0.68–0.95
[5]	Detection	Trap	Psyllids	Image patches	CNN	0.69–0.92
[34]	Detection	Controlled	Whitefly	Spatial, color and texture features	Knowledge-based system	0.70–0.97
[35]	Classification	Field	10 species	Original images	Deep residual learning	0.98
[27]	Classification	Trap	Whiteflies, aphids and thrips	Size, shape and color features	Feature similarity	0.59–1.00
[36]	Classification	Field	10 species	Resized images	AlexNet (CNN)	0.94
[37]	Classification	Field	10 species	SIFT, NNSC, LCP features	SVM	0.85
[38]	Classification	Field	8 species	Segmented images	ANN	0.93
[39]	Detection	Trap	Moth	Patches from original image	CNN	0.93
[17]	Detection	Field	Thrips	HSI Color Channels	SVM	0.98
[40]	Detection	Trap	Western flower thrips, whitefly	Patches from original image	Multi-layer ANN	0.92–0.96
[41]	Classification	Field	10 species	L*a*b* images	K-Means Clustering	N/A
[42]	Classification	Trap	6 species	Morphological and color features	ANN	0.82
[43]	Detection	Trap	24 species	Original Images	CNN + AFRPN	0.55–0.99
[44]	Detection	Field	Whiteflies	Image with background removed	Multifractal analysis	0.88
[45]	Detection	Field	Aphids	HOG Features	SVM	0.87
[46]	Classification	Field	12 species	Cropped regions	CNN	0.83–0.95
[10]	Survey	N/A	N/A	N/A	N/A	N/A
[47]	Classification	Trap	16 butterfly species	Original images	CNN + RPN + PSSM	0.75
[29]	Detection	Controlled	Aphids	Binary image	Object counting	0.81–0.96
[11]	Survey	N/A	N/A	N/A	N/A	N/A
[28]	Detection	Trap	Whiteflies	Binary image	Object counting	0.81–0.99
[48]	Detection	Field	Mexican Bean Beetle, Colorado Potato Beetle	Original images	RSC and LRA	0.76–0.89
[33]	Classification	Trap	Flies, gnats, moth flies, thrips and whiteflies	Original images	Tiny YOLO v3 + multistage CNN	0.92–0.94
[49]	Classification	Trap	6 species	LOSS V2 + SIFT	Feature Similarity	0.96–0.99
[50]	Detection	Trap	Red turpentine beetle	Original images	Downsized RetinaNet	0.75
[51]	Detection	Controlled	Beet armyworm	Morphological and texture features	ANN	0.89
[52]	Classification	Varied	20 insect species	SURF + HOG features	SVM	0.90
[25]	Detection	Field	Whiteflies	Thresholding + morphology	Connected objects counting	0.82–0.97
[53]	Detection	Field	Whiteflies	Original images	K-means clustering	0.92–0.98
[54]	Detection	Field	3 species	Original images	CNN + Decision Net	0.62–0.91
[55]	Classification	Trap	5 species	SIFT features	MLSLC, KNNC, PDLC, PCALC, NMC, SVM	0.77–0.89
[56]	Classification	Trap	8 species	Several features	MLSLC, KNNC, NMC, NDLC, DT	0.60–0.87
[57]	Classification	Trap	9 moth species	154 features	IpSDAE	0.94
[58]	Detection	Trap	Whiteflies	Original images	Multifractal analysis	0.96
[59]	Classification	Trap	Whiteflies, aphids, thrips	Segmented insects	Mahalanobis distance	0.70–0.91
[60]	Classification	Field	24 species	Resized images	VGG19 (CNN)	0.89
[61]	Classification	Trap	4 species of Lepidoptera	156 color, shape and texture features	SVM	0.97
[62]	Detection	Trap	Several	Original images	Normalized Cuts, watershed, K-means	0.96
[63]	Detection	Field	Rice Planthoppers	Haar and HOG features	Adaboost and SVM	0.85
[64]	Detection	Field	Atractomorpha, Erthesina, Pieris	Original images	DSRNLP	0.70–0.92
[65]	Classification	Trap	6 species	Original images, variety of features	YOLO, SVM	0.90–0.92

Legend: ANN—Artificial Neural Network; AFRPN—Anchor-Free Region Proposal Network; CNN—Convolutional Neural Network; DSRNLP—Deep Recursive Super Resolution Network with Laplacian Pyramid; DT—Decision Tree; HOG—Histograms of Oriented Gradient; IpSDAE—Improved Pyramidical Stacked De-noising Autoencoder; KNN—k-Nearest Neighbors; LCP—Local Configuration Pattern; LRA—Limited Receptive Area; MLSLC—Minimum Least Square Linear Classifier; NDLC—Normal Densities Based Linear Classifier; NMC—Nearest Mean Classifier; NNSC—Non-negative Sparse Coding; PCALC—Principal Component Analysis Expansion Linear Classifier; PDLC—Parzen Density Based Linear Classifier; PSSM—Position-Sensitive Score Map; RPN—Region Proposal Network; RSC—Random Subspace Classifier; SIFT—Scale-Invariant Feature Transform; SURF—Speeded-Up Robust Features; SVM—Support Vector Machine; YOLO—You Only Look Once.

2.1. Pest Detection Methods

Detection methods are interested in distinguishing a certain target pest from the rest of the scene in an image. This is equivalent to a binary classification in which the classes are “target present” and “target absent”. The number of detected specimens is often tallied in order to provide a measurement for the degree of infestation.

Some early studies explored image processing techniques without using machine learning algorithms. This type of approach has as main advantage its simplicity, both in terms of implementation and computational complexity. On the other hand, they rely heavily on handcrafted parameters and thresholds, which cause this type of algorithm to be highly susceptible to changes in the characteristics of the images. In one of the earliest studies on the subject [34], two knowledge-based systems (KBS) were combined for detecting and counting whiteflies in traps. The first KBS employed a series of image processing operations to select a few candidate location for the specimens. The second KBS used a series of numerical descriptors to fully characterize the specimens of interest; these, combined with a set of rules, determined if the pest was present or not in the candidate locations. In the same year, Qiao et al. [28] used simple thresholding and a boundary tracking algorithm (sequence of rules delineating each object of interest) for the detection of whiteflies in traps. Al-Saqer and Hassan [30] used Zernike moments and regional properties to describe the shapes of the pest of interest (Red Palm Weevil). The features calculated using by two methods were used to build a library of templates for the targeted pest, which in turn was used to determine the presence or absence of the Red Palm Weevil amidst other insect species. Wang et al. [25] applied a sequence of image processing operations for detection of whiteflies in crops. The proposed algorithm is composed of a median filter, followed by Otsu thresholding and an opening operation to separate clustered objects. The final estimate is obtained by counting the connected objects. Barbedo [26] combined color space transformations, thresholding and mathematical morphology operations (hole filling) for the detection and counting of whiteflies in soybean leaves. Special attention was given to the detection of nymphs at the earliest stages of development. Maharlooei et al. [29] employed color transformations and some image processing operations for the detection of aphids in soybean leaves. Leaf segmentation was performed manually, followed by segmentation of the pests using the hue-saturation-intensity (HSI) color space. Objects of certain sizes were selected as aphids and the remaining connected components were counted.

Multifractal analysis aims to determine the degree of irregularity and complexity of an object. This technique can be a good fit for pest detection due to its robustness to scale and rotation variations, while preserving most of the information contained in the image. Two studies employed this technique. In [58], candidate blobs representing potential locations for whiteflies in traps were determined using two different thresholding schemes. Multifractal analysis was then applied to extract the features that characterize the objects of interest. In [44], image background was first removed by means of a Mahalanobis distance, then multifractal analysis was applied to select some candidate blobs. Features extracted for each blob were combined with size and shape rules for the final whitefly detection in the images.

K-means clustering is a vector quantization technique that aims to group a certain number of observations into k clusters, or classes. This technique was employed by Wang et al. [53] for detecting whiteflies in crops. The algorithm begins by dividing the image into 100×100 blocks. Both the RGB and $L^*a^*b^*$ color spaces are then used as basis for an algorithm that preselects potential cluster center, and then k-means clustering is applied to classify each pixel. Spurious objects are eliminated using ellipse eccentricity rules. Yao et al. [62] investigated the performance of three techniques, normalized cuts (NCuts), watershed and K-means clustering, applied to the separation of pest specimens in traps. NCuts with the optical flow angle as weight function achieved the most accurate results. K-means clustering was one of the classifiers tested in [45], and it has also been used to enhance the results produced by deep learning models [50].

SVMs are among the most widely used machine learning classifiers. They are particularly suitable for binary classifications, as they try to find the hyperplane that best separates two classes. Yao et al. [63] employed a three-layer detection strategy for the detection of rice planthoppers in crops. The first layer was an AdaBoost classifier based on Haar features; the second layer was a SVM classifier based on HOG features; the third layer used color and shape features to remove spurious objects detected in the first two steps. Liu et al. [45] employed SVMs fed with HOG features for detection of aphids in images captured in wheat crops. The HOG features were extracted from regions selected using a maximally stable extremal region (MSER) descriptor. Four other machine learning classifiers were also tested: k-means clustering, Adaboost with Haar feature, and SVM with two different sets of features. Ebrahimi et al. [17] used the HSI (Hue, Saturation, Intensity) color channels as inputs to a SVM for detection of thrips in strawberry flowers. Prior to the SVM application, background was removed by means of a thresholding, and morphological operations were applied to close holes in the binary image.

Artificial neural networks (ANN) are models containing numerous nodes and connections, being loosely inspired by the way neurons are organized and interconnected in a brain. ANNs have been frequently used for the task of image classification. Before 2010, virtually all neural networks had a shallow architecture (few hidden layers). As computational power grew and Graphics Processing Unit (GPUs) became the main piece of hardware for training the models, deep architectures began to rise. Since 2015, research on image classification has strongly veered towards using deep learning [66]. However, a few studies still employ ANNs with shallow architectures. Vakilian and Massah [51] combined image processing techniques with ANNs for the detection beet armyworms in images captured under controlled conditions. Potential pests were first segmented by means of a Canny edge detector, then seven morphological and texture features were extracted and used as input for the three-layer neural networks. Espinoza et al. [40] used a three-step approach for detecting thrips and whiteflies in traps. Potential locations for the objects of interest were selected by means of a threshold, from which 14 color and morphological features were extracted. Those features were used as input to a multilayer feed-forward neural network, which was responsible for the final detection. Roldán-Serrato et al. [48] applied ANNs for the detection of beetles in bean and potato crops. Two types of neural network architectures were used, RSC and LRA, with the former yielding better results both in terms of speed and accuracy.

Deep architectures have been preferred due to their ability to infer spatial relationships without explicit feature extraction [66]. Convolutional neural networks (CNNs) have been particularly popular due to their intimate relationship between layers and spatial information [67]. CNNs have been used in [39] for detecting and counting moths in traps. The system employed a sliding window and, for each image patch, the CNN output a probability for the presence of a moth. Non-maximum suppression (NMS) was used to retain only the patches for which the probability was locally maximal and above a certain threshold, revealing the locations of the specimens. Barbedo and Castro [5] explored the use of CNNs for the detection of psyllids in images of traps. Using the Squeezenet architecture, the authors investigated the influence of sensor quality, image resolution and training/test dataset distribution on the ability of the model to properly detect the psyllids amidst other objects (insects and debris) captured by the trap. CNNs have also been combined with decision networks in order to explore contextual information [54], and with an anchor-free region proposal network for pinpointing pest positions [43].

Although CNNs have been the most popular deep learning architectures for image classification, other types of architectures have also been explored. Sun et al. [50] used a downsized RetinaNet for beetle detection in traps. The RetinaNet, which is a one-stage deep learning detector, was downsized by a depthwise separable convolution and feature pyramid tailoring. The detector was further enhanced by a k-means anchor optimization and residual classification subnet. The system was capable of distinguishing red turpentine beetles from five other beetle species. Yue et al. [64] focused their study on the development of an effective algorithm for increasing the resolution of images captured in the

field. The proposed strategy employed a deep recursive super resolution network with Laplacian Pyramid for high resolution reconstruction of the images. The algorithm was compared with two other image upscaling methods, Bicubic Interpolation and Super-Resolution Convolutional Neural Network (SRCNN).

2.2. Pest Classification Methods

The challenge involved in classifying pests is high because a system like this not only has to properly discriminate between the targeted species, but also deal with non-targeted species, which can be numerous. For this reason, the adoption of sophisticated machine learning techniques is higher than in the case of detection. From the classification studies considered in this work, only two did not employ machine learning tools. Cho et al. [27] used a series of image processing operation for classification of whiteflies, aphids and thrips in traps. Two separate strategies were used: the first, dedicated to the identification of aphids, was based on some size, shape and color features; the second used features extracted from the YUV color space to discriminate whiteflies and thrips. A related approach was adopted in [49], which combined two feature-extraction techniques for recognition of six different insect species. The first technique, called LOSS, extracts a number of shape features to characterize the objects of interest. The second technique, SIFT, is a well-known algorithm that identify salient features in the image. Specimens were then grouped by feature value similarity.

Before the rise of deep learning techniques, SVM was the preferred technique for pest classification in images. However, there have been some studies employing other types of machine learning techniques. Xia et al. [59] used the watershed algorithm for segmentation of the insects, followed by the extraction of color features from the YCrCb color space using the Mahalanobis distance. The classification of each object into whitefly, aphid or thrips was given by the nearest distance between the extracted feature vector and the reference vectors associated to each class. K-means clustering was also applied in [41] for recognition of 10 pest species in images captured in the field.

The objective of some studies was to compare and evaluate several different machine learning models. Wen et al. [55] used the SIFT descriptor to extract features for the characterization of 5 pest species in images captured under controlled conditions. Six classifiers were tested: MLSLC, kNN, PDLC, PCALC, NMC, and SVM. PCALC and SVM yielded the highest accuracies. The same authors tested several geometric, contour, texture and color features as inputs for five different classifiers [56]: MLSLC, KNNC, NMC, NDLC and DT. NDLC yielded the highest accuracies when classifying eight different pest species.

As mentioned above, SVMs have been among the most widely used machine learning techniques. Although interest has dropped with the inception of deep learning techniques, SVMs are still being employed in a wide variety of classification problems. Yao et al. [61] extracted 156 color, shape and texture features and used them as input for a SVM classifier with a radial basis function as kernel, with the objective of classifying four species of Lepidoptera. Venugoban and Ramanan [52] employed SURF and HOG descriptors to extract the features used to feed a SVM dedicated to classify 20 pest species. SVMs have also been used in the context of bio-inspired methods for classification of 10 species in crops [37]. In this work, the Saliency Using Natural statistics model (SUN) was used to generate saliency maps and detect the region of interest (ROI). Features were extracted using the Hierarchical Model and X (HMAX) model combined with SIFT, NNSC and LCP algorithms. All extracted features were then fed to a SVM using a Radial Basis Function (RBF) as kernel function for pest recognition.

As in the case of pest detection, shallow neural network architectures are still being explored for pest classification. Han et al. [42] focused on the proposal of a hardware platform for pest classification in images of traps. The system extracts several morphological and color features that are used as input for the three-layer, 29-neuron ANN dedicated to the classification. Dimililer and Zarrouk [38] used a shallow ANN architecture for classification of eight pest species. The algorithm preprocesses images using grayscale conversion and median filtering, applies thresholding to remove background and

canny edge detection to segment the pest, and the image is rescaled by pattern averaging before being fed to the two neural networks responsible for pest classification.

Deep learning techniques are the current state of the art when it comes to classification based on digital images. As is the case for pest detection, CNNs have been prevalent. A deep CNN model for classification of six pest species in images of traps was tested in [46]. The authors used a global contrast region-based approach to compute a saliency map for localizing pest insect objects. Bounding boxes containing targets were then extracted and used to train and test a CNN with an architecture inspired by AlexNet. Different architectures were explored by shrinking depth and width. Cheng et al. [35] optimized two CNN architectures, ResNet50 and ResNet101, using the deep residual learning method. When used to classify ten different pest species in field images with complex backgrounds, the proposed method outperformed SVM, shallow ANN and plain AlexNet CNN classifiers. The VGG19 deep learning architecture was employed in [60] for classifying 24 pest species in images captured under natural conditions. Region Proposal Network was adopted rather than a traditional selective search technique to generate a smaller number of proposal windows, improving accuracy and speed. The method was compared with a few other deep learning architectures. Dawei et al. [36] applied transfer learning to a pre-trained AlexNet CNN for classification of 10 species in images captured in the field. The model outperformed a CNN trained from scratch, likely due to the relatively small dataset size, which did not contain enough information for proper full model training. A four-part deep learning approach was used in [47] for classification of 16 butterfly species in images of traps. First, channel-spatial attention (CSA) was fused into the CNN backbone for feature extraction and enhancement. A RPN was applied to indicate potential pest positions. A PSSM was used to replace fully connected (FC) layers in the deep CNN architecture for pest classification and bounding box regression. Finally, contextual regions of interest were used to improve detection accuracy. A two-step deep learning approach for classifying and counting five pest types in images of traps was adopted in [33]. The detection step, in which locations of the insects on the sticky paper are determined, was performed by the Tiny YOLO v3 object detector. The classification step was carried out in a three-level hierarchical structure, in which CNN models trained specifically to deal with the classification problems associated to each level were applied successively.

Two other types of deep learning strategies have been investigated for pest classification. Besides being used in [33] (see above), YOLO has been investigated in [65]. This work employed a two-step machine learning approach for the classification of six pest species in images of traps. Objects were first localized by means of the YOLO detector, followed by classification and counting using a SVM fed with color, shape, texture and HOG features. Four kernel functions were tested. In [57], a deep learning model was applied for classification of nine moth species in traps. The algorithm applied a two-step morphological segmentation of the objects of interest (moths). Then, 154 texture, color, shape and local features were extracted and fed to an improved pyramidal stacked de-noising autoencoder (IpSDAE) architecture, the deep learning model dedicated to the classification step. Results were compared to SVM, random forests (RF), BayesNet, logistic regression classifier (LRC), and radial basis function (RBF).

3. Discussion

3.1. The Data Gap Problem

As mentioned before, machine learning algorithms are now capable of solving most classification problems, as long as the data used to fit the models are representative enough. Due to intrinsic characteristics of the agricultural environment, the data collected tends to have a high degree of variability. Many studies try to counteract this by controlling the conditions under which images are acquired. This certainly increases the accuracies obtained in the experiments, but those same models tend to fail with data obtained under more realistic conditions. There are two possible solutions for this problem. The first solution is the development of models and training strategies capable of

learning data distributions from a relatively small number of samples. As suggested in Lake et al. [68], the learning process of models dealing with small datasets should evolve towards the way humans learn, which involves incorporating domain-specific knowledge. Using synthetic data and augmentation to artificially increase the dataset size can also be employed [69]. The use of Generative Adversarial Networks (GANs) may be particularly interesting in this context [70], but more studies are needed to determine if this kind of approach can indeed properly capture the whole variability introduced by the agricultural environment. As promising as these new algorithms are, their development is still in its infancy and will probably not be applicable to agricultural classification problems in the near future. The second solution is to improve the process of image acquisition to better cover all variability found in practice. The remainder of this section discusses the main hurdles towards such a goal, also proposing possible solutions whenever relevant. All issues discussed in this section are summarized in Table 2.

3.2. Difficulties Related to the Strategy Adopted—Images of Traps

A little less than half of the studies considered in this work employ traps to aid the task of pest monitoring. This type of device often offers the highest count accuracy and robustness, and can be applied to multiple types of pests [45,62]. Images of traps tend to have more homogeneous characteristics than those captured in the field. As a result of this reduced variability, it is easier to build a dataset capable of representing the whole range of conditions found in practice. However, images of traps also have some specific challenges associated. Shadows caused by the trap's frame may cause detection problems [40]. Specimens trapped near or on the edges of the frames are hard to detect [28]. Also, depending on the time the trap is left on the field, insects may decay and become unrecognizable [40]. Noise [44] and grid marks [27,57,59] have also been pointed out as a potential sources of error. Fortunately, most of those problems can be prevented.

A common approach for the pest monitoring using traps is to collect images regularly and remotely without removing the trap [39], which enables an analysis of the infestation over time. This type of approach also brings some challenges. As time elapses, new objects appear in the trap (increasing occlusion and overlapping problems), wing poses change, decay effects increase, illumination conditions vary, and background texture changes [39]. These factors may cause inconsistent count estimates overtime. Ding and Taylor [39] argued that if temporal image sequences are provided with a reasonably high frequency, changes may be tracked more easily and errors may be avoided.

The position in which insects get glued can change considerably. Pose variations can cause detection and classification problems [57]. Some studies suggested that the best solution would be to capture images as soon as insects land [56], but this may be impractical in most cases. Including different poses in the dataset used to tune the algorithms and models seem to be a more feasible option, especially taking into consideration that traps usually collect a large number of samples.

Depending on the population density of insects, trapped specimens may touch or even overlap [50,61]. In some cases, traps may contain thousands of objects [40]. Situations like this can be very challenging, often leading pest numbers to be considerably underestimated [34,39,40]. Some image processing techniques are capable of partially separating clusters [62], and some deep learning models can successfully deal with crowded images [47]. However, depending on the degree of overlapping, the only solution may be applying statistical correction techniques. In this case, estimates will be only an approximation, but ultimately not even highly trained humans would be able to provide an accurate estimate under overcrowded conditions [26].

Traps with low population density also pose some problems. In particular, the relative impact of the presence of objects such as dust, debris and other insects becomes more pronounced [28]. If the number of specimens of interest is very low, even a couple of false positives will cause high error rates, a problem that is also observed when field images are used [63]. This type of issue is not easy to address, but should be taken into consideration when evaluating the performance of any trained model.

Table 2. List of challenges linked to pest detection and recognition and potential solutions.

Issue	Source of the Issue	Potential Solutions
Image distortions caused by the trap structure	Trap	Enforce image capture protocols
Object configuration in the trap changes over time	Trap	Increase frequency of image capture
The poses in which insects are glued vary	Trap	Capture images as soon as insects land
Overlapping insects in the traps	Trap	Image processing techniques for separating clusters; statistical corrections
Presence of dust, debris and other insects	Trap	Trap and lure should be designed to minimize the presence of spurious objects
Delay between the beginning of infestation and the moment the trap has enough samples	Trap	Capture images directly in the field environment
Non-airborne pests are missed	Trap	Capture images directly in the field environment
Traps have too many problems associated	Trap	Capture images directly in the field environment
Images captured in the field have a high degree of variability	Condition variety in the field	Employ techniques for illumination correction; build more comprehensive datasets
Contrast between specimens of interest and their surroundings change	Condition variety in the field	Build more comprehensive datasets; enforce image capture protocols
Angle of image capture cause specimens to appear distorted	Condition variety in the field	Build more comprehensive datasets; enforce image capture protocols
Pest occlusion by leaves and other structures	Condition variety in the field	Use of autonomous robots for peering into occluded areas; statistical corrections
Failure to detect specimens at early stages of development	Insect's intrinsic characteristics	Development of more sensitive algorithms; use of more sophisticated sensors
Other objects in the scene cause confusion	Uncontrolled environment	Use more sophisticated detection algorithms; build more comprehensive datasets
Visual similarities between the pest of interest and other species	Insect's intrinsic characteristics	Use more sophisticated detection algorithms; build more comprehensive datasets
Different cameras produce images with different characteristics	Imaging equipment	Build more comprehensive datasets; enforce image capture protocols
More spatial resolution does not necessarily translate into better results	Imaging equipment	Specific experiments need to be designed to investigate the ideal resolution
Data overfitting	Model learning	Apply regularization techniques and other methods for reducing overfitting
Class imbalance	Model learning	Oversample smaller classes and/or undersample the larger ones
Covariate shift issues	Model learning	Apply domain adaptation techniques; build more comprehensive datasets
Inconsistencies in the reference annotations	Model learning	Image annotation performed redundantly by several people
Most datasets are not comprehensive enough	Model learning	Data sharing; citizen science; use of social networks concepts

While traps allow for a more controlled pest monitoring, they also have some major disadvantages. According to Liu et al. [45], traps may not be the best tool for making decisions related to treatment need or timing, as there may be a delay between the beginning of infestation and the moment the trap has enough samples for meaningful conclusions. In addition, traps usually capture only airborne adult pests, missing immature specimens capable of causing severe damage to the crops. For this reason, many studies try to detect and classify pests directly on the leaves, with images being captured either under controlled conditions (detached leaves) or directly in the field.

3.3. Difficulties Related to the Strategy Adopted—Images in the Field

When images are captured in the field, considerable illumination differences may exist [26,45], which can pose a significant challenge [29]. Those differences can be caused by weather conditions (sunny or overcast), angle of insolation, camera settings, presence of shadows, among others [71]. Many datasets do not reflect these variations, as the protocols used to generate them usually prevent images to be captured under conditions far from ideal. In practice, it may be unfeasible to force a potential user to follow the same protocols, especially considering how hostile the agricultural environment can be. Consequently, the variability found in practice is usually much larger than that found in most studies, which is one explanation for the tendency of most proposed methods to fail under more realistic conditions. This is particularly true for methods based on deep learning, as these require a comprehensive training dataset to work properly. In most cases, practical adoption of software for field pest monitoring will require models trained with more realistic and comprehensive datasets. It is worth noting that there are some techniques capable of partially compensating for illumination differences [45], which may be useful depending on the classifier being applied.

The background of images captured in the field can also vary considerably [37], which means that the contrast between specimens of interest and their surroundings will change. This is true even when detection is performed on plant leaves, as these can have colors ranging from light green to dark brown. In addition, the background usually contains artifacts (stems, leaves, spots) that may cause confusion [45]. As mentioned recurrently throughout this article, the training dataset has to represent the entire range of conditions expected to be found in practice, but if the degree of background variation is too high, this may be unfeasible. In cases like this, it may be necessary to enforce some image capture protocols.

Another challenge associated to images captured in the field is that perspective distortions may be common [45], causing specimens to appear with very different visual characteristics depending on the angle of capture. This problem is less prevalent in the case of small pests, but in any case the dataset used for training should contain different viewpoints of the targeted pest, or some image capture protocols need to be created (but these can be difficult to enforce in practice).

Full automation of pest monitoring in the field without using traps is currently unfeasible in most cases. Although it is possible to install a network of image sensors strategically located throughout the crop, many pests of interest are usually found underneath the leaves. Even when pests are located on top of the leaves, chances are that most specimens will be occluded by the crop canopies [45]. In the future, it may be possible to deploy swarming robots that will actively search for the infestations. At present, image-based field monitoring needs to either be paired with other monitoring strategies, or estimates need be corrected by carefully designed statistical models [63].

3.4. Difficulties Related to the Insects Themselves

Most methods are designed to detect adult insects. However, a complete picture about the infestation and how it is evolving may require the identification and counting of specimens at earlier stages of development [17,26,45]. This may pose a significant challenge, because younger specimens may not only be smaller [63], but also have quite different visual characteristics [45]. In some extreme cases, young nymphs may be semi-transparent, which makes them very difficult to be detected [26].

When the targeted pest is small and detection is to be performed directly on the leaves, symptoms and signs produced by diseases, nutritional deficiencies, dead skins and other insects may also become sources of confusion [26,29]. If this is expected to be a common occurrence, the training dataset should include samples of those problematic situations so the model can learn how to properly detect the pest of interest under such conditions.

Another major challenge faced by recognition methods is that the targeted pest may share many visual similarities with other species [59]. This fact is particularly relevant when traps are used, as many different species may be present in a single image [50,55], including many that are not targeted by the classification system [61]. While most studies addressed this problem, the reported results will be intrinsically connected to the specific geographical region where the data was collected [5]. Since other regions may harbor different pest species, the degree of detection difficulty may vary. Thus, while most published results may be viewed as proofs of concept, new experiments need to be carried out and new models should be generated whenever a new area is considered.

3.5. Difficulties Related to the Imaging Equipment

The cameras used for capturing the images may also have considerable impact on the ability of the model to detect the objects of interest [26]. Optical quality plays an important role and, under low illumination conditions, camera settings may also be a relevant factor [29]. Building a dataset including samples captured with all kinds of sensors is impractical; instead, cameras expected to be prevalent in practice should be given priority.

In most cases, the higher the spatial resolution of the images used for recognition, the more information is available and, consequently, better results can be achieved (at the price of increased computational requirements). However, depending on the technique being used, too much detail may lead to oversegmentation, increasing error rates [59]. Thus, it is important to consider that working with the highest possible resolution is not necessarily the best approach.

3.6. Difficulties Related to Model Learning

A problem that is often overlooked is that many machine learning algorithms tend to overfit the data [5], especially with limited data. It is common to find studies reporting accuracies very close to 100%, which may not be realistic depending of the problem being addressed (this may also be caused by too homogeneous datasets). There are several ways to avoid overfitting, including the application of regularization techniques [46], image augmentation [39], unit dropout in deep architectures [46], etc., so more care should be given to the experimental design.

Class imbalance is also a common problem. In any image classification problem, certain classes will usually be much more common than others. Because there is a tendency to build datasets as large as possible, the proportion of samples ends up reflecting such imbalance. The problem with this is that the classifier will tend to become highly biased and, depending on the metrics used to evaluate the model, such bias may remain undetected. The number of samples should be roughly the same for all classes, which can be achieved by either oversampling the smaller classes or undersampling the larger ones [5].

Covariate shift issues are also widespread and rarely addressed. Covariate shift is the phenomenon in which differences between the distributions of the training data used to train the model and the data on which the model is to be applied result in low accuracies [66,71,72]. This problem becomes evident when the same database is used for training and assessing the models, which is common practice in the machine learning community for practical reasons [66,73]. Some studies on recognition of plant diseases have reported steep accuracy drops when the model trained with a given dataset is applied to other data sources [74]. Domain adaptation techniques can mitigate this problem [75], but a more definite solution would involve capturing a much wider variety of training data, which may be impractical [66].

The evaluation of the proposed methods is also not trivial. Samples in the dataset need to be properly annotated to serve as the references (or ground-truth) to which the outputs yielded by the model are compared. The problem is that the annotation process is subjective, labor intensive and, as a result, susceptible to considerable inconsistencies [71,76]. Indeed, some authors have observed significant deviations when different people performed the task [34]. Ideally, image annotation should be performed redundantly by several people, but in most cases this is impractical. Alternatively, authors should always make clear that the references used for evaluation of their algorithms might have some inconsistencies.

Many of the problems mentioned in this section could be minimized by the use of more comprehensive datasets. In a natural environment, conditions cannot be controlled, which means that data variability will be very high. Covering all possible situations found in practice is nearly impossible, especially considering that some conditions are rare and require a certain amount of luck to be properly captured. As observed in [37], pests will occur infrequently and in different locations, and they will not always be ideally positioned for image capture. In the case of plant diseases, there have been some initiatives to build comprehensive datasets using social network and citizen science concepts [66], which may also be a suitable solution for the problem of pest recognition. On the other hand, the labeling process tends to become less rigorous under such a scheme [66], exacerbating the evaluation problem discussed in the previous paragraph. Although there is no definitive solution for the problem of dataset insufficiency, evidence suggests that it is better to have comprehensive datasets that are relatively poorly labelled, than small datasets rigorously annotated [77].

4. Conclusions and Future Directions

Machine learning techniques, and deep learning in special, have been showing a remarkable ability to properly detect and classify pests, either in traps or natural images. Arguably, the main factor preventing a more widespread adoption of automatic pest monitoring systems is their lack of robustness to the vast variety of situations that can be found in practice. This, in turn, is the result of limitations on the datasets used to train the classification models. Building more comprehensive pest image databases is essential to close this gap. However, given the degree of variability associated to practical use, it is very unlikely that substantial progress in this regard will be achieved using conventional approaches. Future efforts could focus on creating mechanisms to facilitate and encourage the involvement of farmers and entomologists in the process of image collection and labelling. This could be achieved, for example, by using the concepts of citizen science [78] to involve more people and scale up the efforts towards a representative database. In the specific case of pest recognition, farmers and field workers could collect images in the field and, after uploaded to a server, those images would be properly labelled by an expert. Initiatives like this are already being carried out in the context of plant disease identification (Barbedo, 2018).

Joint efforts and data sharing can also greatly contribute toward the creation of more comprehensive datasets. If datasets generated by different research groups were made available and properly integrated, the resulting set of images would be much more representative and research results would be more meaningful and applicable to real world conditions. Datasets adhering to the FAIR (Findable, Accessible, Interoperable, and Reusable) principles [79] are particularly useful, as this enables and maximizes the fitness for reuse of research data.

As mentioned before, monitoring pests under natural conditions usually allows for a more timely response to possible infestations. However, monitoring entire farms in real time is a considerable challenge. Having a high density network of imaging sensors covering the entire area would be ideal, but this option is probably not cost effective, and it is technically challenging to properly implement it in an inhospitable environment. An alternative would be to embed sensors in the machinery used in the farm's daily life. This would not allow for real time monitoring, but the whole property would be covered in relatively short time frames. In the more distant future, employing swarming micro-robots

may also become feasible [80]. Other suitable solutions may exist, but more research effort needs to be dedicated to the subject.

Traps have some disadvantages, but provide an easier way for monitoring large areas. Better solutions in edge computing need to be created so the communications network needed for full monitoring automation is not overwhelmed by excessive amounts of data [81]. Models and algorithms also need to be light in order to small, low-cost processing units to be able to process the images and produce useful information in a timely manner. With raw data being properly processed at the edge of the network, it should be possible to transmit only the information needed in the decision making process.

Automating pest monitoring is a challenging task. With the evolution of machine learning algorithms, the tools needed to build accurate systems with real practical applications are already available. Gathering data representative enough of the huge variability found in practice is very challenging, but as devices with imaging capabilities become ubiquitous and mechanisms to enable citizen science are perfected, this might not be a significant problem in the relatively near future. However, as discussed throughout this article, there are still many research gaps that need to be addressed, which means that pest monitoring automation will continue to be a compelling research subject for many years to come.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

References

- Oerke, E.C. Crop losses to pests. *J. Agric. Sci.* **2006**, *144*, 31–43. [[CrossRef](#)]
- Nalam, V.; Louis, J.; Shah, J. Plant defense against aphids, the pest extraordinaire. *Plant Sci.* **2019**, *279*, 96–107. [[CrossRef](#)] [[PubMed](#)]
- Van Lenteren, J.; Bolckmans, K.; Köhl, J.; Ravensberg, W.; Urbaneja, A. Biological control using invertebrates and microorganisms: plenty of new opportunities. *BioControl* **2017**, *63*, 39–59. [[CrossRef](#)]
- Yen, A.L.; Madge, D.G.; Berry, N.A.; Yen, J.D.L. Evaluating the effectiveness of five sampling methods for detection of the tomato potato psyllid, *Bactericera cockerelli* (Sulc) (Hemiptera: Psylloidea: Trioziidae). *Aust. J. Entomol.* **2013**, *52*, 168–174. [[CrossRef](#)]
- Barbedo, J.G.A.; Castro, G.B. Influence of image quality on the identification of psyllids using convolutional neural networks. *Biosyst. Eng.* **2019**, *182*, 151–158. [[CrossRef](#)]
- Sun, Y.; Cheng, H.; Cheng, Q.; Zhou, H.; Li, M.; Fan, Y.; Shan, G.; Damerow, L.; Lammers, P.S.; Jones, S.B. A smart-vision algorithm for counting whiteflies and thrips on sticky traps using two-dimensional Fourier transform spectrum. *Biosyst. Eng.* **2017**, *153*, 82–88. [[CrossRef](#)]
- Huang, M.; Wan, X.; Zhang, M.; Zhu, Q. Detection of insect-damaged vegetable soybeans using hyperspectral transmittance image. *J. Food Eng.* **2013**, *116*, 45–49. [[CrossRef](#)]
- Ma, Y.; Huang, M.; Yang, B.; Zhu, Q. Automatic threshold method and optimal wavelength selection for insect-damaged vegetable soybean detection using hyperspectral images. *Comput. Electron. Agric.* **2014**, *106*, 102–110. [[CrossRef](#)]
- Clément, A.; Verfaillie, T.; Lormel, C.; Jaloux, B. A new colour vision system to quantify automatically foliar discoloration caused by insect pests feeding on leaf cells. *Biosyst. Eng.* **2015**, *133*, 128–140. [[CrossRef](#)]
- Liu, H.; Lee, S.; Chahl, J. A review of recent sensing technologies to detect invertebrates on crops. *Precis. Agric.* **2017**, *18*, 635–666. [[CrossRef](#)]
- Martineau, M.; Conte, D.; Raveaux, R.; Arnault, I.; Munier, D.; Venturini, G. A survey on image-based insect classification. *Pattern Recognit.* **2017**, *65*, 273–284. [[CrossRef](#)]
- Lehmann, J.R.K.; Nieberding, F.; Prinz, T.; Knoth, C. Analysis of Unmanned Aerial System-Based CIR Images in Forestry—A New Perspective to Monitor Pest Infestation Levels. *Forests* **2015**, *6*, 594–612. [[CrossRef](#)]
- Vanegas, F.; Bratanov, D.; Powell, K.; Weiss, J.; Gonzalez, F. A Novel Methodology for Improving Plant Pest Surveillance in Vineyards and Crops Using UAV-Based Hyperspectral and Spatial Data. *Sensors* **2018**, *18*, 260. [[CrossRef](#)] [[PubMed](#)]

14. Liu, H.; Lee, S.; Chahl, J.S. A Multispectral 3-D Vision System for Invertebrate Detection on Crops. *IEEE Sensors J.* **2017**, *17*, 7502–7515. [[CrossRef](#)]
15. Liu, H.; Chahl, J.S. A multispectral machine vision system for invertebrate detection on green leaves. *Comput. Electron. Agric.* **2018**, *150*, 279–288. [[CrossRef](#)]
16. Fan, Y.; Wang, T.; Qiu, Z.; Peng, J.; Zhang, C.; He, Y. Fast Detection of Striped Stem-Borer (*Chilo suppressalis* Walker) Infested Rice Seedling Based on Visible/Near-Infrared Hyperspectral Imaging System. *Sensors* **2017**, *17*, 2470. [[CrossRef](#)] [[PubMed](#)]
17. Ebrahimi, M.; Khoshtaghaza, M.; Minaei, S.; Jamshidi, B. Vision-based pest detection based on SVM classification method. *Comput. Electron. Agric.* **2017**, *137*, 52–58. [[CrossRef](#)]
18. Al-doski, J.; Mansor, S.; Mohd Shafri, H.Z. Thermal imaging for pests detecting—A review. *Int. J. Agric. For. Plant.* **2016**, *2*, 10–30.
19. Barbedo, J.G.A. Detection of nutrition deficiencies in plants using proximal images and machine learning: A review. *Comput. Electron. Agric.* **2019**, *162*, 482–492. [[CrossRef](#)]
20. Neethirajan, S.; Karunakaran, C.; Jayas, D.; White, N. Detection techniques for stored-product insects in grain. *Food Control* **2007**, *18*, 157–162. [[CrossRef](#)]
21. Haff, R.P.; Saranwong, S.; Thanapase, W.; Janhira, A.; Kasemsumran, S.; Kawano, S. Automatic image analysis and spot classification for detection of fruit fly infestation in hyperspectral images of mangoes. *Postharvest Biol. Technol.* **2013**, *86*, 23–28. [[CrossRef](#)]
22. Lu, R.; Ariana, D.P. Detection of fruit fly infestation in pickling cucumbers using a hyperspectral reflectance/transmittance imaging system. *Postharvest Biol. Technol.* **2013**, *81*, 44–50. [[CrossRef](#)]
23. Shah, M.; Khan, A. Imaging techniques for the detection of stored product pests. *Appl. Entomol. Zool.* **2014**, *49*, 201–212. [[CrossRef](#)]
24. Shen, Y.; Zhou, H.; Li, J.; Jian, F.; Jayas, D.S. Detection of stored-grain insects using deep learning. *Comput. Electron. Agric.* **2018**, *145*, 319–325. [[CrossRef](#)]
25. Wang, K.; Zhang, S.; Wang, Z.; Liu, Z.; Yang, F. Mobile smart device-based vegetable disease and insect pest recognition method. *Intell. Autom. Soft Comput.* **2013**, *19*, 263–273. [[CrossRef](#)]
26. Barbedo, J.G.A. Using digital image processing for counting whiteflies on soybean leaves. *J. Asia Pac. Entomol.* **2014**, *17*, 685–694. [[CrossRef](#)]
27. Cho, J.; Choi, J.; Qiao, M.; Ji, C.; Kim, H.; Uhm, K.B.; Chon, T.S. Automatic identification of whiteflies, aphids and thrips in greenhouse based on image analysis. *Int. J. Math. Comput. Simul.* **2007**, *1*, 46–53.
28. Qiao, M.; Lim, J.; Ji, C.W.; Chung, B.K.; Kim, H.Y.; Uhm, K.B.; Myung, C.S.; Cho, J.; Chon, T.S. Density estimation of *Bemisia tabaci* (Hemiptera: Aleyrodidae) in a greenhouse using sticky traps in conjunction with an image processing system. *J. Asia Pac. Entomol.* **2008**, *11*, 25–29. [[CrossRef](#)]
29. Maharlooei, M.; Sivarajan, S.; Bajwa, S.G.; Harmon, J.P.; Nowatzki, J. Detection of soybean aphids in a greenhouse using an image processing technique. *Comput. Electron. Agric.* **2017**, *132*, 63–70. [[CrossRef](#)]
30. Al-Saqer, S.M.; Hassan, G.M. Red Palm Weevil (*Rynchophorus Ferrugineous*, Olivier) Recognition by Image Processing Techniques. *Am. J. Agric. Biol. Sci.* **2011**, *6*, 365–376. [[CrossRef](#)]
31. Kamilaris, A.; Prenafeta-Boldu, F.X. Deep learning in agriculture: A survey. *Comput. Electron. Agric.* **2018**, *147*, 70–90. [[CrossRef](#)]
32. Liakos, K.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D. Machine Learning in Agriculture: A Review. *Sensors* **2018**, *18*, 2674. [[CrossRef](#)]
33. Rustia, D.J.A.; Chao, J.J.; Chung, J.Y.; Lin, T.T. Crop losses to pests. In Proceedings of the 2019 ASABE Annual International Meeting, Boston, MA, USA, 7–10 July 2019. [[CrossRef](#)]
34. Boissard, P.; Martin, V.; Moisan, S. A cognitive vision approach to early pest detection in greenhouse crops. *Comput. Electron. Agric.* **2008**, *62*, 81–93. [[CrossRef](#)]
35. Cheng, X.; Zhang, Y.; Chen, Y.; Wu, Y.; Yue, Y. Pest identification via deep residual learning in complex background. *Comput. Electron. Agric.* **2017**, *141*, 351–356. [[CrossRef](#)]
36. Dawei, W.; Limiao, D.; Jiangong, N.; Jiyue, G.; Hongfei, Z.; Zhongzhi, H. Recognition pest by image-based transfer learning. *J. Sci. Food Agric.* **2019**, *99*, 4524–4531. [[CrossRef](#)] [[PubMed](#)]
37. Deng, L.; Wang, Y.; Han, Z.; Yu, R. Research on insect pest image detection and recognition based on bio-inspired methods. *Biosyst. Eng.* **2018**, *169*, 139–148. [[CrossRef](#)]
38. Dimililer, K.; Zarrouk, S. ICSPI: Intelligent Classification System of Pest Insects Based on Image Processing and Neural Arbitration. *Appl. Eng. Agric.* **2017**, *33*, 453–460. [[CrossRef](#)]

39. Ding, W.; Taylor, G. Automatic moth detection from trap images for pest management. *Comput. Electron. Agric.* **2016**, *123*, 17–28. [[CrossRef](#)]
40. Espinoza, K.; Valera, D.L.; Torres, J.A.; López, A.; Molina-Aiz, F.D. Combination of image processing and artificial neural networks as a novel approach for the identification of Bemisia tabaci and Frankliniella occidentalis on sticky traps in greenhouse agriculture. *Comput. Electron. Agric.* **2016**, *127*, 495–505. [[CrossRef](#)]
41. Faithpraise, F.O.; Birch, P.; Young, R.; Obu, J.; Faithpraise, B.; Chatwin, C. Automatic plant pest detection and recognition using k-means clustering algorithm and correspondence filters. *Int. J. Adv. Biotechnol. Res.* **2013**, *4*, 189–199.
42. Han, R.; He, Y.; Liu, F. Feasibility Study on a Portable Field Pest Classification System Design Based on DSP and 3G Wireless Communication Technology. *Sensors* **2012**, *12*, 3118–3130. [[CrossRef](#)] [[PubMed](#)]
43. Jiao, L.; Dong, S.; Zhang, S.; Xie, C.; Wang, H. AF-RCNN: An anchor-free convolutional neural network for multi-categories agricultural pest detection. *Comput. Electron. Agric.* **2020**, *174*, 105522. [[CrossRef](#)]
44. Li, Y.; Xia, C.; Lee, J. Detection of small-sized insect pest in greenhouses based on multifractal analysis. *Optik Int. J. Light Electron Opt.* **2015**, *126*, 2138–2143. [[CrossRef](#)]
45. Liu, T.; Chen, W.; Wu, W.; Sun, C.; Guo, W.; Zhu, X. Detection of aphids in wheat fields using a computer vision technique. *Biosyst. Eng.* **2016**, *141*, 82–93. [[CrossRef](#)]
46. Liu, Z.; Gao, J.; Yang, G.; Zhang, H.; He, Y. Localization and Classification of Paddy Field Pests using a Saliency Map and Deep Convolutional Neural Network. *Sci. Rep.* **2016**, *6*. [[CrossRef](#)]
47. Liu, L.; Wang, R.; Xie, C.; Yang, P.; Wang, F.; Sudirman, S.; Liu, W. PestNet: An End-to-End Deep Learning Approach for Large-Scale Multi-Class Pest Detection and Classification. *IEEE Access* **2019**, *7*, 45301–45312. [[CrossRef](#)]
48. Roldán-Serrato, K.L.; Escalante-Estrada, J.; Rodríguez-González, M. Automatic pest detection on bean and potato crops by applying neural classifiers. *Eng. Agric. Environ. Food* **2018**, *11*, 245–255. [[CrossRef](#)]
49. Solis-Sánchez, L.O.; Castañeda-Miranda, R.; García-Escalante, J.J.; Torres-Pacheco, I.; Guevara-González, R.G.; Castañeda-Miranda, C.L.; Alaniz-Lumbreras, P.D. Scale invariant feature approach for insect monitoring. *Comput. Electron. Agric.* **2011**, *75*, 92–99. [[CrossRef](#)]
50. Sun, Y.; Liu, X.; Yuan, M.; Ren, L.; Wang, J.; Chen, Z. Automatic in-trap pest detection using deep learning for pheromone-based Dendroctonus valens monitoring. *Biosyst. Eng.* **2018**, *176*, 140–150. [[CrossRef](#)]
51. Vakilian, K.A.; Massah, J. Performance evaluation of a machine vision system for insect pests identification of field crops using artificial neural networks. *Arch. Phytopathol. Plant Prot.* **2013**, *46*, 1262–1269. [[CrossRef](#)]
52. Venugoban, K.; Ramanan, A. Image Classification of Paddy Field Insect Pests Using Gradient-Based Features. *Int. J. Mach. Learn. Comput.* **2014**, *4*, 1–5. [[CrossRef](#)]
53. Wang, Z.; Wang, K.; Liu, Z.; Wang, X.; Pan, S. A Cognitive Vision Method for Insect Pest Image Segmentation. *IFAC Pap. Online* **2018**, *51*, 85–89. [[CrossRef](#)]
54. Wang, F.; Wang, R.; Xie, C.; Yang, P.; Liu, L. Fusing multi-scale context-aware information representation for automatic in-field pest detection and recognition. *Comput. Electron. Agric.* **2020**, *169*, 105222. [[CrossRef](#)]
55. Wen, C.; Guyer, D.E.; Li, W. Local feature-based identification and classification for orchard insects. *Biosyst. Eng.* **2009**, *104*, 299–307. [[CrossRef](#)]
56. Wen, C.; Guyer, D. Image-based orchard insect automated identification and classification method. *Comput. Electron. Agric.* **2012**, *89*, 110–115. [[CrossRef](#)]
57. Wen, C.; Wu, D.; Hu, H.; Pan, W. Pose estimation-dependent identification method for field moth images using deep learning architecture. *Biosyst. Eng.* **2015**, *136*, 117–128. [[CrossRef](#)]
58. Xia, C.; Lee, J.M.; Li, Y.; Chung, B.K.; Chon, T.S. In situ detection of small-size insect pests sampled on traps using multifractal analysis. *Opt. Eng.* **2012**, *51*. [[CrossRef](#)]
59. Xia, C.; Chon, T.S.; Ren, Z.; Lee, J.M. Automatic identification and counting of small size pests in greenhouse conditions with low computational cost. *Ecol. Inform.* **2015**, *29*, 139–146. [[CrossRef](#)]
60. Xia, D.; Chen, P.; Wang, B.; Zhang, J.; Xie, C. Insect Detection and Classification Based on an Improved Convolutional Neural Network. *Sensors* **2018**, *18*, 4169. [[CrossRef](#)]
61. Yao, Q.; Lv, J.; Liu, Q.J.; Diao, G.Q.; Yang, B.J.; Chen, H.M.; Tang, J. An Insect Imaging System to Automate Rice Light-Trap Pest Identification. *J. Integr. Agric.* **2012**, *11*, 978–985. [[CrossRef](#)]
62. Yao, Q.; Liu, Q.; Dietterich, T.G.; Todorovic, S.; Lin, J.; Diao, G.; Yang, B.; Tang, J. Segmentation of touching insects based on optical flow and NCuts. *Biosyst. Eng.* **2013**, *114*, 67–77. [[CrossRef](#)]

63. Yao, Q.; Xian, D.X.; Liu, Q.J.; Yang, B.J.; Diao, G.Q.; Tang, J. Automated Counting of Rice Planthoppers in Paddy Fields Based on Image Processing. *J. Integr. Agric.* **2014**, *13*, 1736–1745. [[CrossRef](#)]
64. Yue, Y.; Cheng, X.; Zhang, D.; Wu, Y.; Zhao, Y.; Chen, Y.; Fan, G.; Zhang, Y. Deep recursive super resolution network with Laplacian Pyramid for better agricultural pest surveillance and detection. *Comput. Electron. Agric.* **2018**, *150*, 26–32. [[CrossRef](#)]
65. Zhong, Y.; Gao, J.; Lei, Q.; Zhou, Y. A Vision-Based Counting and Recognition System for Flying Insects in Intelligent Agriculture. *Sensors* **2018**, *18*, 1489. [[CrossRef](#)] [[PubMed](#)]
66. Barbedo, J.G.A. Factors influencing the use of deep learning for plant disease recognition. *Biosyst. Eng.* **2018**, *172*, 84–91. [[CrossRef](#)]
67. Arel, I.; Rose, D.C.; Karnowski, T.P. Deep Machine Learning—A New Frontier in Artificial Intelligence Research [Research Frontier]. *IEEE Comput. Intell. Mag.* **2010**, *5*, 13–18. [[CrossRef](#)]
68. Lake, B.M.; Salakhutdinov, R.; Tenenbaum, J.B. Human-level concept learning through probabilistic program induction. *Science* **2015**, *350*, 1332–1338. [[CrossRef](#)]
69. Hu, G.; Peng, X.; Yang, Y.; Hospedales, T.M.; Verbeek, J. Frankenstein: Learning Deep Face Representations Using Small Data. *IEEE Trans. Image Process.* **2018**, *27*, 293–303. [[CrossRef](#)]
70. Barth, R.; Hemming, J.; Henten, E.V. Optimising realism of synthetic images using cycle generative adversarial networks for improved part segmentation. *Comput. Electron. Agric.* **2020**, *173*. [[CrossRef](#)]
71. Barbedo, J.G.A. A review on the main challenges in automatic plant disease identification based on visible range images. *Biosyst. Eng.* **2016**, *144*, 52–60. [[CrossRef](#)]
72. Sugiyama, M.; Nakajima, S.; Kashima, H.; Büna, P.v.; Kawanabe, M. Direct Importance Estimation with Model Selection and Its Application to Covariate Shift Adaptation. In Proceedings of the 20th International Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 3–6 December 2007; pp. 1433–1440.
73. Ferentinos, K.P. Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agric.* **2018**, *145*, 311–318. [[CrossRef](#)]
74. Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using Deep Learning for Image-Based Plant Disease Detection. *Front. Plant Sci.* **2016**, *7*, 1419. [[CrossRef](#)]
75. Ben-David, S.; Blitzer, J.; Crammer, K.; Kulesza, A.; Pereira, F.; Vaughan, J.W. A theory of learning from different domains. *Mach. Learn.* **2010**, *79*, 151–175. [[CrossRef](#)]
76. Bock, C.H.; Poole, G.H.; Parker, P.E.; Gottwald, T.R. Plant Disease Severity Estimated Visually, by Digital Photography and Image Analysis, and by Hyperspectral Imaging. *Crit. Rev. Plant Sci.* **2010**, *29*, 59–107. [[CrossRef](#)]
77. Bekker, A.J.; Goldberger, J. Training deep neural-networks based on unreliable labels. In Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, China, 20–25 March 2016; pp. 2682–2686. [[CrossRef](#)]
78. Irwin, A. *Citizen Science: A Study of People, Expertise and Sustainable Development*, 1st ed.; Routledge: Abingdon-on-Thames, UK, 2002.
79. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.W.; da Silva Santos, L.B.; Bourne, P.E.; et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **2016**, *3*. [[CrossRef](#)] [[PubMed](#)]
80. Albani, D.; Ijsselmuiden, J.; Haken, R.; Trianni, V. Monitoring and mapping with robot swarms for agricultural applications. In Proceedings of the 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Lecce, Italy, 29 August–1 September 2017; pp. 1–6. [[CrossRef](#)]
81. Shi, W.; Cao, J.; Zhang, Q.; Li, Y.; Xu, L. Edge Computing: Vision and Challenges. *IEEE Internet Things J.* **2016**, *3*, 637–646. [[CrossRef](#)]

