Ecological Informatics 67 (2022) 101516

Contents lists available at ScienceDirect

Ecological Informatics

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

InsectCV: A system for insect detection in the lab from trap images

Telmo De Cesaro Júnior [a,*], Rafael Rieder [b], Jéssica Regina Di Domênico [a], Douglas Lau [c]

[a] Sul-rio-grandense Federal Institute of Education, Science and Technology (IFSul), Passo Fundo, RS, Brazil
[b] University of Passo Fundo (UPF), Passo Fundo, RS, Brazil
[c] Brazilian Agricultural Research Corporation (Embrapa Trigo), Passo Fundo, RS, Brazil

ARTICLE INFO

Keywords:
Convolutional neural network
Mask r-cnn
Object detection
Pest detection
Aphids
Warning systems

ABSTRACT

Advances in artificial intelligence, computer vision, and high-performance computing have enabled the creation of efficient solutions to monitor pests and identify plant diseases. In this context, we present InsectCV, a system for automatic insect detection in the lab from scanned trap images. This study considered the use of Moericke-type traps to capture insects in outdoor environments. Each sample can contain hundreds of insects of interest, such as aphids, parasitoids, thrips, and flies. The presence of debris, superimposed objects, and insects in varied poses is also common. To develop this solution, we used a set of 209 grayscale images containing 17,908 labeled insects. We applied the Mask R-CNN method to generate the model and created three web services for the image inference. The model training contemplated transfer learning and data augmentation techniques. This approach defined two new parameters to adjust the ratio of false positive by class, and change the lengths of the anchor side of the Region Proposal Network, improving the accuracy in the detection of small objects. The model validation used a total of 580 images obtained from field exposed traps located at Coxilha, and Passo Fundo, north of Rio Grande do Sul State, during wheat crop season in 2019 and 2020. Compared to manual counting, the coefficients of determination ($R^2 = 0.81$ for aphids and $R^2 = 0.78$ for parasitoids) show a good-fitting model to identify the fluctuation of population levels for these insects, presenting tiny deviations of the growth curve in the initial phases, and in the maintenance of the curve shape. In samples with hundreds of insects and debris that generate more connections or overlaps, model performance was affected due to the increase in false negatives. Comparative tests between InsectCV and manual counting performed by a specialist suggest that the system is sufficiently accurate to guide warning systems for integrated pest management of aphids. We also discussed the implications of adopting this tool and the gaps that require further development.

1. Introduction

The increase in the aphid population (Hemiptera: Aphididae) in winter cereals (e.g. wheat, barley, triticale, and oat) might cause a significant reduction in grain yield due to their feeding habits. Aphids are also vectors of pathogens, such as barley yellow dwarf virus (BYDV-PAV, BYDV-MAV, BYDV-PAS) and cereal yellow dwarf virus (CYDV-RPV, CYDV-RPS) (Savaris et al., 2013). Monitoring pest population fluctuations, whether in the field or laboratory experiments, allows researchers to track the variation in infestation levels and support integrated pest management programs (Hodgson et al., 2012). According to Brabec et al. (Brabec et al., 2014), aphid population fluctuation is determined by seasonal periodic sampling and can vary significantly monthly or annually. Furthermore, aphids have an overdispersed population dynamic and can vary by order of magnitude within 24 h (Davis et al., 2014; Howard and Dixon, 2008; Jarošová et al., 2019).

Regarding pest monitoring programs, Brazilian Agricultural Research Corporation, Embrapa Trigo[1] coordinates a winter cereal National Pest Monitoring Network since 2015. This network uses Moericke-type yellow traps in experiments to observe migrations of aphid populations and their natural enemies, the parasitoid wasps (Hymenoptera: Aphelinidae and Braconidae, Aphidiinae). This method is widely used and easy to implement, allowing researchers to capture aphids from a specific area for monitoring and understanding population oscillations (Döring, 2014; Morris, 2018). The experiments allow associate population levels to economic damage and generating relevant data to warning

* Corresponding author.
E-mail addresses: telmojunior@ifsul.edu.br (T. De Cesaro Júnior), rieder@upf.br (R. Rieder), jessicadomenico.pf016@academico.ifsul.edu.br (J.R. Di Domênico), douglas.lau@embrapa.br (D. Lau).

[1] Embrapa unit that develops wheat and other winter cereals.

https://doi.org/10.1016/j.ecoinf.2021.101516
Received 5 August 2021; Received in revised form 1 December 2021; Accepted 1 December 2021
Available online 6 December 2021

systems in the decision-making process related to crop management (Embrapa Trigo, 2017; Embrapa Trigo, 2019; Engel et al., 2021).

The process of sorting elements captured in the traps requires manual tasks such as collecting, removing debris, and relocating aphids and parasitoids on a test plate for visual analysis by a laboratory specialist. Each trap can contain hundreds of insects of different species. During the identification stage, dichotomous keys that specify the morphology of each species are used, such as the number of antenna segments, wing branches, and body shape. Due to the small size of the insects, a stereoscopic microscope (magnifying glass) is commonly used for analysis (Santos et al., 2019).

The variety and quantity of insects captured in the traps significantly increase during the hottest times of the year, interfering with the detection of aphids and parasitoids. Thus, due to seasonal effects, the challenge of analyzing a sample to accurately recognize and count the elements of interest varies throughout the year. Therefore, manual identification is a time-consuming task, may generate fatigue, and its accuracy might be reduced due to the number of insects of interest (aphids and parasitoids), debris (soil, leaves, grains), and other insects (flies, bees, thrips, beetles). Samples with many elements cause connections, partial overlaps and changes in pose, which also hinder the analysis process. When there are large concentrations of objects, usually the manual counting of objects of interest is only estimated based on spatial proportions and the experience of the laboratory specialist (Embrapa Trigo, 2017; Embrapa Trigo, 2019; Nazri et al., 2018).

Many studies address the use of computational resources to automate the process of identifying and counting objects in digital images, such as insect vectors and plant diseases. These studies also highlight that the manual identification process is considered an intensive, slow, tedious, error-prone, and low-precision activity, making its large-scale execution unfeasible (Akintayo et al., 2018; Ärje et al., 2020; Ding and Taylor, 2016; Nazri et al., 2018; Pang et al., 2020; Partel et al., 2019; Sun et al., 2018). In doing so, solutions based on Computer Vision (CV) and Artificial Intelligence (AI) have enabled the automation of repetitive tasks (Kamilaris and Prenafeta-Boldú, 2018). These applications have also reached good results for counting and classifying insects in traps, arachnids in plants, leaf diseases in plants, and insect monitoring by sensors, helping experts and decreasing time in the decision-making process (Fuentes et al., 2017; Høye et al., 2021; Li et al., 2019; Zhong et al., 2018).

For example, Lins et al. (Lins et al., 2020) partially automated the lab identification of aphid morphs (nymphs, wingless and winged adults) from the *Rhopalosiphum padi* L. species. In the study, a set of procedures for the collection of possible specimens retained in the traps was defined, and the CV tool called *AphidCV* (Lins et al., 2019) was developed. The software has many features, such as counting, classification, and measurement of aphids from digital images, using an AI model.

The results about precision in classification, location, and segmentation tasks highlighted in these recent studies and the FCIS challenge (Li et al., 2017), demonstrated the potential of the Mask R-CNN method (He et al., 2017) to solve the problem reported in this paper. As the problem considers identifying aphids and parasitoids in the images, and there may be cases of overlapping, pose variation, debris, and other insects with similar morphology, instance segmentation must be applied. Thus, the Mask R-CNN could handle these cases more accurately when compared to other methods, such as Faster R-CNN. (Ren et al., 2016). Mask R-CNN also makes it possible to use segmentation to compute the areas of all objects in samples, predicting levels of biological activity around a trap. Besides, it is possible to correlate the reduction in the number of insects of interest with the increase of residues and other insects.

The aim of this work is to present the **InsectCV**, a computational system to detect aphids and parasitoids regardless of species from trap images, using CV and AI techniques. The system includes the selection and pre-processing of a set of images, labeling of the objects of interest in each image, training, validation of the intelligent model based on the Mask R-CNN method, and the development of web services for integration into the Trap System (Lazzaretti et al., 2021) monitoring platform. By this system, we intend to support warning systems collecting and processing data, facilitating and expanding the detection of variations in population levels of these insects.

This study is structured as follows: Section 2 presents related work about Faster R-CNN and Mask R-CNN; Section 3 describes the materials and methods used to develop the proposed solution and define the creation of the AI model for object detection; Section 4 shows the training results and validation experiments; Section 5 discusses the results to assess the model's accuracy; and, finally, Section 6 presents the final considerations and future work.

2. Related work

Advances in artificial intelligence and computer vision have enabled the creation of software to solve many problems in different areas, including Agriculture (Kamilaris and Prenafeta-Boldú, 2018). The evolution of Deep Convolutional Neural Network (DCNN) architecture (LeCun et al., 2015) also provided a significant increase in performance and accuracy of these applications. According to Chen et al. (Chen et al., 2018), the accuracy levels of the DCNNs indicate this technique represents the state-of-the-art for image classification tasks (Krizhevsky et al., 2012), object detection (Ren et al., 2016) and segmentation (Long et al., 2015), with better results than traditional methods, in which the characteristics of the objects of interest are manually extracted (hand-crafted features) (Fischer et al., 2014).

During last five years, literature reviews have been highlighting the use of DCNNs to automate repetitive tasks in agriculture, such as plant disease detection and insect identification, among other purposes. Kamilaris and Prenafeta-Boldú (Kamilaris and Prenafeta-Boldú, 2018) presented a selection of 40 studies from 2010 to 2017 that applied different AI techniques in agriculture; of these, 42% used DCNNs.

In a recent survey by De Cesaro Júnior and Rieder (De Cesaro Júnior and Rieder, 2020), approximately 300 studies published in journals between 2015 and 2019 were found, corroborating the trend towards the use of DCNN's for the object detection task (Kamilaris and Prenafeta-Boldú, 2018). In the research, 33 studies were selected from the digital databases: ACM, IEEE, IET, DBLP, ScienceDirect, Scopus, SpringerLink, and Web of Science. These studies were investigated in order to verify the use of techniques for detecting small objects and handling connections and overlaps.

From selected studies by this survey for the object detection task, we verified that most of them used Faster R-CNN two-stage detector. Mask R-CNN extends Faster R-CNN by adding a branch for predicting an object mask in parallel with the existing branch for bounding box recognition (He et al., 2017). No selected study used Mask R-CNN. Considering only studies based on Faster R-CNN for insect recognition, four works proposed changes in the base implementation of this algorithm.

Shen et al. (Shen et al., 2018) presented a Faster R-CNN-based method to identify six species of grain insects. Aiming at improving the extracted feature maps, the authors proposed a change in the creation network (inspired by He et al. (He et al., 2016)) to generate layered feature maps with 1/8 of the original image, making the network sensitive to small objects. The specimens and grains were manually positioned to generate 739 images with a dimension of 1944 × 2592 pixels, enabling the control of connections and overlaps. In comparison with four other convolutional networks, the proposed approach presented an increase of 3.34% mean average precision (mAP) in insect identification.

Yue et al. (Yue et al., 2018) proposed a super-resolution model based on the deep recursive residual network for image reconstruction (the image could be restored and up-sampled according to a scale only the target insect could adapt to the applied object detection method at the pixel scale). Experimental results show that the method improved the recall rate by 202%. In this work, Faster R-CNN was only used to

validate the proposed image interpolation model. Therefore, the objective was to reduce the density of the camera layout of the agricultural Internet of Things (IoT) monitoring systems and reduce infrastructure costs.

Liu et al. (Liu et al., 2019) proposed *PestNet*, a solution to assist in pest management, with pest detection and classification based on Faster R-CNN. The authors used a dataset with more than 80 k images and approximately 580 k insects, categorized in 16 classes (species). The experimental results showed the proposed approach performed well with a 75.45% (mAP). A three-stage architecture was developed: feature extraction (CSA), regions search (RPN, PSSM), and prediction (RoIs). According to the authors, a better accuracy was not possible due to the similarity between some species of insects. Because the trap has a cleaning system performed at 15-s intervals after image acquisition, cases of connections or overlapping objects were significantly reduced.

Kalamatianos et al. (Kalamatianos et al., 2018) presented a system to automate the identification and counting of the olive fruit fly (*Bactrocera oleae*) based on images of the commonly used McPhail trap's contents. The dataset used for model training was relatively small, containing only 542 images. Several deep learning methods were evaluated, achieving a performance of 91.52% (mAP). The images look similar to those captured in yellow traps by the aphid monitoring network (Engel et al., 2021), representing a "pest soup images". This scenario makes it hard to identify submerged insects and thus indistinguishable even by in situ experts, much less by remote observers through digital images. The authors also customized a trade-off between image attributes affecting detail, file size and complexity of approaches and mAP performance that can be selectively used to better tackle the needs of each usage scenario.

Research has demonstrated the use of Mask R-CNN algorithm in agriculture to automate manual tasks, reduce costs and increase accuracy in object identification, counting, and segmentation. To analyze the use of this algorithm, a literature review was carried out in the scientific database Science Direct. The search considered only studies published in English, from 2019 to 2020, that focused on agriculture and used two-dimensional images. The keyword "mask rcnn" was applied to find the studies published in journals. Approximately 250 studies were found, of which 10 were related to the focus of this paper (Bhattarai et al., 2020; Ganesh et al., 2019; Jia et al., 2020; Pang et al., 2020; Reyes-Yanes et al., 2020; Ruiz-Santaquiteria et al., 2020; Vo et al., 2020; Wang et al., 2020; Xu et al., 2020; Yu et al., 2019). Table 1 shows a summary of these studies.

Table 1
Selected studies that applied Mask R-CNN for agricultural purposes.

Authors	Dataset	Object	Overlapping	Customizations
Reyes-Yanes et al. (Reyes-Yanes et al., 2020)	1350	Romaine lettuce	Not	–
Xu et al. (Xu et al., 2020)	750	Cattle	Yes	–
Pang et al. (Pang et al., 2020)	1000	Crop-row maize	Not	branch MaxArea_MaskIoU
Vo et al. (Vo et al., 2020)	3000	Rock lobster	Not	–
Jia et al. (Jia et al., 2020)	1020	Apple	Yes	–
Wang et al. (Wang et al., 2020)	1500	Waxberry	Yes	dilated convolution
Ruiz-Santaquiteria et al. (Ruiz-Santaquiteria et al., 2020)	126	Algae	Yes	–
Yu et al. (Yu et al., 2019)	1900	Strawberry	Yes	–
Ganesh et al. (Ganesh et al., 2019)	150	Orange	Yes	–
Bhattarai et al. (Bhattarai et al., 2020)	205	Blossom apple	Not	–

The analysis of the methodology and the results presented in these 10 recent studies showed that, initially, the image sets used for training and tests were relatively small, ranging from 126 to 3000 images. Most authors applied the technique of data augmentation and transfer learning to avoid overfitting in training. The most used feature extractors were the ResNet-50 and the ResNet-101. Regarding customizations, only two studies (Pang et al., 2020) (Wang et al., 2020) highlighted changes in the default implementation of He et al. (He et al., 2017). Regarding the techniques for treating connections or overlaps, we noted that none of the methods mentioned solve this problem.

3. Materials and methods

For the image acquisition for the training of the model, samples retained in the traps (yellow tray, 45 cm long x 30 cm wide x 4.5 cm tall) were used, collected according to Engel et al. (Engel et al., 2021). Each sample was filtered through two strainers. The resulting material was deposited in a 14 cm-diameter Petri dish and later digitized using a flatbed scanner. We used the same protocol developed by Lins et al. (Lins et al., 2020), generating and selecting 90 images. In addition, other 119 new images were generated in the laboratory to create the image sets, contained only parasitoids (Hymenoptera: Aphelinidae, Braconidae, Aphidiinae) and winged aphids (Hemiptera: Aphididae), without debris.

Two sets of images were defined for model training and testing. The first set consisted of 167 images and 14,809 labeled insects, 12,354 for training, and 2455 for model testing. To create the second set, the images from the first set were used and new images were included, resulting in 209 images and 17,908 labeled insects.

In each set, the images were organized into the following subsets: (i) winged aphids; (ii) parasitoids; (iii) winged aphids and parasitoids; (iv) winged aphids, parasitoids and debris; (v) winged aphids and debris. All objects of interest were labeled with the graphical tool LabelImg.[2] The objects were classified into two classes: aphid or parasitoid. This task was supervised by an Embrapa researcher.

In Table 2, each row represents a subset of images/insects. The Aphids and Parasitoids columns indicate the amounts of these labeled insects in each subset. The column Insects/Train. Images represents the number of insects and images used for training the models. The Insects/Images Test. represents the number of insects and images used to test the models. The set contained 167 images, 134 for training and 33 for testing. Although this set is considered small for deep learning, a significant number of objects were labeled (14,809). Subsets (iv) and (v) contained only images generated by the screening process. The images were captured in field traps located in the cities of Passo Fundo, Coxilha, and Cruz Alta, RS, Brazil, from 2018 to 2020. The other subsets were created artificially with the manual insertion of specimens into the test plate without the presence of debris.

Table 2
First set of images for model training.

Set	Aphids	Parasitoids	Insects/Train. Images	Insects/Test. Images	Total
1	4934	0	4688/ 44	246/ 2	4934/ 46
2	0	6137	5942/ 43	195/ 2	6137/ 45
3	1745	1611	1506/ 13	1850/ 15	3356/ 28
4	77	15	10/ 1	82/ 5	92/ 6
5	290	0	208/ 33	82/ 9	290/ 42
Total	7046	7763	12,354/ 134	2455/ 33	14,809/ 167

[2] Available in: https://github.com/tzutalin/labelImg

Table 3 lists the second set with 209 images. In this new set, to reach the proportion of 80% of images for training and 20% for testing, the five subgroups were balanced with the inclusion of new images. The test images were chosen according to the number of objects so the same proportion between images with different densities of insects could be achieved. They were classified as images of easy recognition (less dense) and medium recognition (dense). Thus, the number of labeled insects increased to 17,908, with 14,669 insects used for training and 3239 used for testing.

For the generation of the intelligent model, this study used the implementation of Mask R-CNN developed by Abdulla (Abdulla, 2017), Python programming language, and OpenCV libraries,[3] TensorFlow[4] for machine learning and Keras,[5] a neural network library. The hardware resources used were a computer with an Intel Core I7-6950× processor, 32 GB of RAM, and the GeForce GTX Titan X graphics processing unit (GPU) with 12 GB of VRAM.

Three web services were developed through the Django REST framework[6] to make the intelligence model available. The registration of new images, deletion, and detection of insects (inference) were developed with Python. These services were integrated into the Trap System website,[7] using the *libcurl*[8] library and the PHP programming language.

The registration service converts the original color image (RGB) to *GrayScale* format, applying the filter *Gaussian Blur* for noise elimination and clipping with the algorithm *Hough Circles* for the elimination of irrelevant areas. After this processing, the final image has a dimension of 6156 × 6156 pixels, a horizontal and vertical resolution of 96 dpi, and an intensity of 8 bits. Finally, the original and final images are stored in a local agronomy database named AgroDB (Lazzaretti et al., 2016).

The detection service receives the code of an image via the Trap System, performs the inference through the intelligence model, applies thresholds by class, and generates a new demonstrative image with the location of the *bounding-boxes* and the classification of objects. Next, the service returns a message in *JSON* with the image in Base64 format and the counting of the insects of interest. Finally, the Trap System stores the detection result in AgroDB tables.

The deletion service removes a selected image. All services are hosted on a server at the Instituto Federal Sul-Rio-Grandense (IFSul), in Passo Fundo. This server has approximately 14 GB of RAM and 8 vCPUs. The implementation in the form of services allows other systems to access these functionalities. In addition to these services, a routine was developed to process the images already stored in AgroDB.

According to Abdulla (Abdulla, 2017), the training of an intelligent model can be started with random weights *starch*, which implies changing the weights of all layers of the network or reusing *transfer learning* weights, where weights from an external model are used.

Table 3
Second set of images used for model training.

Set	Aphids	Parasitoids	Insects/Train. Images	Insects/Test. Images	Total
1	6101	0	5035/ 36	1066/ 9	6101/ 45
2	0	6311	5154/ 36	1157/ 9	6311/ 45
3	1950	1646	2884/ 23	712/ 6	3596/ 29
4	1093	168	1042/ 34	219/ 8	1261/ 42
5	639	0	554/ 38	85/ 10	639/ 48
Total	9783	8125	14,669/ 167	3239/ 42	17,908/ 209

[3] Available in: https://opencv.org/
[4] Available in: https://tensorflow.org
[5] Available in: https://keras.io
[6] Available in: https://www.django-rest-framework.org/
[7] Available in: http://gpca.passofundo.ifsul.edu.br/traps/
[8] Available in: https://curl.haxx.se/libcurl/

Moreover, when applying *transfer learning*, only a subset of network layers can be trained, using fewer hardware resources, and reducing processing time. As this study predicted the use of grayscale input images, the following subsets of layers were adjusted:

- **heads**: conv1.*, mrcnn_.*, rpn_.* e fpn_.*;
- **5+**: conv1.*, res5.*, bn5.*, mrcnn_.*, rpn_.* e fpn_.*,
- **4+**: conv1.*, res4.*, bn4.*, res5.*, bn5.*, mrcnn_.*, rpn_.* e fpn_.*;
- **3+**: conv1.*, res3.*, bn3.*, res4.*, bn4.*, res5.*, bn5.*, mrcnn_.*, rpn_.*, fpn_.*;
- **all**: ".*",

Based on the available hardware resources, the input images could not be fully processed, due to their original size and dimension (6156 × 6156 pixels). Therefore, this study resized the images to 1024 × 1024 and 2048 × 2048 (*square* mode), and applied the random cropping technique (*crop* mode) of samples in the dimension of 2048 × 2048 pixels. These modes are mentioned by Abdulla (Abdulla, 2017) and can be activated by parameters.

The feature extractor used was the ResNet50 due to hardware restrictions. The *Batch Size* was set at 1. Three sets of values were applied for the lengths of the anchor side of RPN (hyperparameter *RPN_ANCHOR_SCALES*). The DETECTION_MAX_INSTANCES was set at 500, considering the possibility of images with hundreds of objects of interest. The MAX_GT_INSTANCES was set at 100, due to hardware restrictions. The other parameters set by Abdulla (Abdulla, 2017) were not changed.

Initially, five experiments were performed with a set of 167 images. Each experiment ran from 40 to 140 epochs (training), and each epoch had 623 steps. Due to the limited number of images, two compensatory techniques were applied to avoid the *over-fitting* problem. First, all layers of the network were started with the weights of the MS COCO (*transfer learning*) model, except the layers: *mrcnn_class_logits*, *mrcnn_bbox_fc*, *mrcnn_bbox*, *mrcnn_mask*, and *conv1*, that were initialized with random weights. The subset of retrained layers was 4+. The random geometric transformations of rotation (−90, +90), horizontal (0.5) and vertical (0.5) inversion were applied in experiments 2, 4 and 5, to triple the number of images (*data augmentation*).

The objective of these experiments was to find the configuration that achieved the highest mean precision (mAP) with the test images. Details of the experiments are shown in Table 4. The "Mode" column indicates the use of resizing of the original image: *square* or *crop*.

According to the data in Table 4, the fifth experiment reached the highest average precision (68.3%), after running 40 epochs. Therefore, based on the best model of this experiment, two additional experiments were performed with the second set of images, involving 3+ layers and value variations for the hyperparameter RPN_ANCHOR_SCALES. The aim was to reduce the loss values and increase the precision in the classification and location of the insects of interest. Details are shown in Table 5.

Experiments 6 and 7 performed another 60 epochs, resulting in 100 epochs. From the 80th epoch, the hyperparameter RPN_ANCHOR_SCALES received the values (4,8,16,32,64) in experiment 6 and (8,16,32,64,128) in experiment 7. The maximum number of instances per image was set to 466, using the hyperparameter MAX_GT_INSTANCES. This value corresponds to the largest amount of labeling performed on a single image of the training set. The number of regions of interest was reduced in the second stage, from 200 to 100, in order to reduce the computational cost. The hyperparameter used was the TRAIN_ROIS_PER_IMAGE.

The last two experiments were evaluated by the loss value generated at each epoch. The tool *TensorBoard*[9] was used for this scenario, which showed that experiment 7 reached the lowest overall loss value, close to

[9] Included in TensorFlow.

Table 4
Results from the initial experiments.

Exp.	Mode	Set of Images	Resolution	Expanded	Time	Prec. train.	Prec. test
1	square	1	1024 × 1024	No	1d 4 h 47 m	87.0%	49.8%
2	square	1	1024 × 1024	Yes	1d 15 h 39 m	71.1%	53.6%
3	square	1	2048 × 2048	No	3d 19 h 16 m	84.5%	59.4%
4	square	1	2048 × 2048	Yes	16d 1 h 15 m	85.2%	60.4%
5	crop	1	2048 × 2048	Yes	15d 4 h 15 m	70.8%	**68.3%**

Table 5
Details of additional experiments 6 and 7.

Exp.	Mode	Set of Images	RPN_ANCHOR_SCALES	Resolution	Expanded	Precision
6	crop	2	4,8,16,32,64	2048 × 2048	Yes	76%
7	crop	2	8,16,32,64,128	2048 × 2048	Yes	79%

1.4 after the 100th epoch. According to Abdulla (Abdulla, 2017), the reference value for overall loss is 0.5. Therefore, experiment 7 determined the choice of the best model for this study.

After choosing the model, approximately 580 images stored in AgroDB were used to validate the selected model. The images were from the 2019 and 2020 wheat crops, from the Embrapa stations in Passo Fundo and Coxilha. We used the same protocol to generate images applied by Lins et al. (Lins et al., 2020). These images were organized into fifteen series: ten series of images to analyze the identification and counting of aphids, and five series to validate the model considering identification and counting of parasitoids.

Thus, with the data obtained from the inferences of the 580 images processed by the detection service, three approaches were applied to analyze the model's performance: statistical coefficients, *precision/recall/F1 score*, and cut-off points. Moreover, customizations were developed with the inclusion of hyperparameters and variations of confidence thresholds.

The weekly sums obtained by manual counting performed by Embrapa's researcher were compared with the data generated by InsectCV to define the coefficients. The coefficient of determination (R^2) was adopted, which indicates how well the model can explain the collected data; the slope, and intercept, which define the linear relationship between two variables, and can be used to estimate an average rate of change. Thus, the intercept must be zero (0), and the slope must be one (1), showing that the model accurately estimates the number of insects of interest in the image.

The *precision* metric shows the correct level of the model for objects classified as parasitoids. The *recall* metric shows how often the model finds examples of a given class. Finally, the *F1 score*, which is the ratio between *recall* and *precision*, generates a combined percentage. These metrics were only applied to identify parasitoids. The labeling process was performed with the researcher's assistance.

Finally, to analyze the applicability and assertiveness of the model developed for the warning systems, we evaluated the temporal coincidences between the visual assessment and the model regarding the achievement of cut-off points (or thresholds) pre-established as critical for decision-making in aphid pest management. These cut-off points were based on historical series of winged aphid captures in traps in the Southern Brazil (Engel et al., 2021).

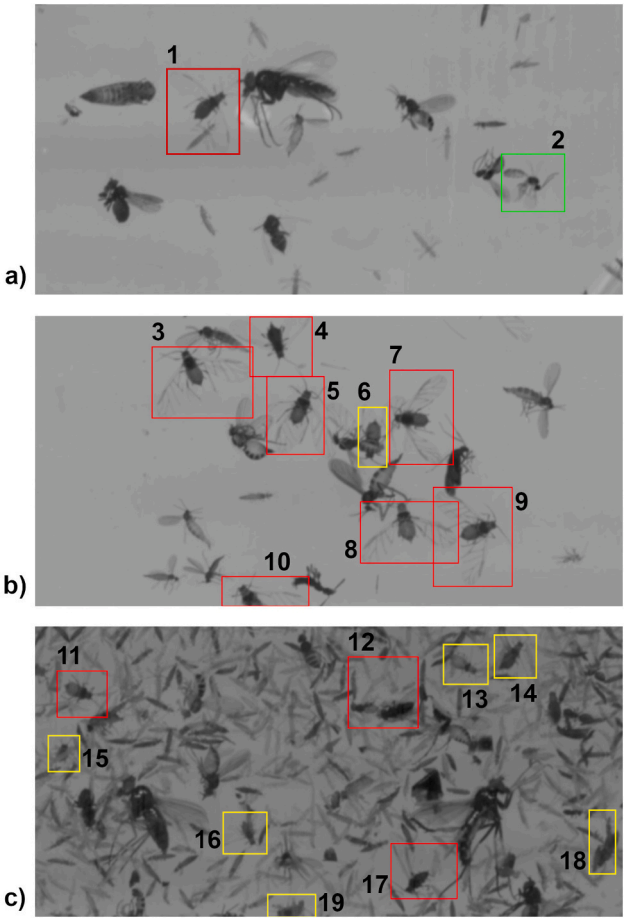

Fig. 1. Examples of low (a), medium (b), and high (c) complexity cases in the detection of winged aphids and parasitoids from samples from field traps scanned in the lab.

4. Results

Since the traps were located in field conditions, there was a large amount of debris in the samples. During peak periods of aphid infestation, high concentrations of insects that occupied all the space on the image capture slide were identified, mainly in traps type 1 located in Coxilha. Fig. 1 shows three image situations used for validation. Low complexity cases (Fig. 1 (a)) present a few insects and debris, facilitating the detection process (objects 1 and 2 were identified correctly as aphids and parasitoids, respectively). Medium complexity cases (Fig. 1 (b)) consider many insects, some connected or overlapped instances (objects 3, 4, 5, 7, 8, 9, and 10 were identified correctly as aphids, and object 6 is a false positive). High complexity cases (Fig. 1 (c)) contain a significant presence of thrips (Thysanoptera) result in numerous connected or overlapped instances (only objects 11 and 17 were identified correctly as aphids, object 12 is a false positive, and objects 13, 14, 15, 16, 18, and 19 are false negatives).

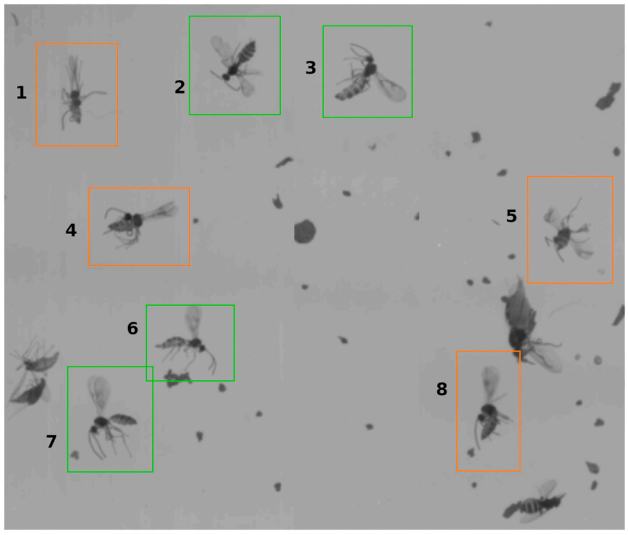

Fig. 2. An example of scanned sample containing parasitoids in different poses and tail angles. Green bounding boxes represent true positives, and orange bounding boxes indicate false negatives. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 2 shows the complexity during the identification of parasitoids. In this image, there are eight specimens. Objects 2, 3, 6, and 7 were correctly identified as parasitoids. Objects 1, 4, 5, and 8 are false negatives in different poses and variations in the angle of the tail about the rest of the body.

Due to the three situations identified in the images used for validation, and the additional complexity in detecting parasitoids, two hyperparameters were created to define the acceptance threshold (confidence) by class. Therefore, this study established different values for the classes of parasitoids and aphids. With the assistance of the researcher, empirical tests were performed to discover the acceptance thresholds that could be applied in this study. The defined acceptance thresholds were 0.985 and 0.976 for the respective classes. For an object to be considered a parasitoid or an aphid, the confidence level obtained by the model must be 98.5% and 97.6%, respectively. Nonetheless, high thresholds might cause false negatives and their reduction increases the probability of false positives.

4.1. Coefficients

In Table 6 and Table 7 ten validation series are listed, referring to the 2019 and 2020 crops. In Table 8, five series are listed, referring to the 2020 crops (07/01/2020 to 10/30/2020). Columns marked with the suffix VL represent the use of a threshold variation parameter. This study used images from different crops, since a significant amount of parasitoids was not identified in 2019.

The greater the number of aphids in the sample, the greater the deviation between the actual number of insects and the number detected by the model. Consequently, at times of peak population of aphids, deviations are greater (Table 6). For trap 4 of 2019, the model achieved the best fit, with a slope coefficient of 0.67 and the $R^2 = 0.99$. However, in trap 3, where 630 aphids were identified at the peak of infestation, the slope coefficient was only 0.39, and the coefficient $R^2 = 0.91$. When counting the five series, the $R^2 = 0.89$ and the slope coefficient of 0.50 were achieved. In the second and third columns, the X value represents the manual identification.

For the five series of 2020, the slopes and the R^2 were lower than in 2019, which indicates less accuracy from the model (Table 7). The line graph shown in Fig. 3 represents the relationship between the manual counting (blue line) and the data generated by InsectCV (red line) for aphids. Considering the time series of the relationship between manual counting and counting obtained by InsectCV, the results showed that the model correctly identified population fluctuations in the two analyzed crops (Fig. 3). For aphids, in the scatter plot shown in Fig. 4,considering the ten series, the slope coefficient was 0.35 and $R^2 = 0.81$.

According to the data shown in Table 8, the accuracy of parasitoid detection was not related to the peak of infestation. For this class, when counting the five series, the slope coefficient was 0.40 and the $R^2 = 0.78$, according to the scatter plot presented in Fig. 6.

The line graph shown in Fig. 5 illustrates the relationship between the manual counting (blue line) and the data generated by InsectCV (red line) for parasitoids. The model correctly identified the increases in population levels that occurred between weeks 28 and 34, the decline between weeks 35 and 38, as well as the small increase in week 41. However, due to the number of parasitoids found during the 2020 crop in the traps located in Coxilha, the analysis of the model's fit was hindered.

A significant amount of false negatives in images with high complexity cases was identified from the results, as shown in Fig. 1 (c). To reduce this characteristic and offer another option for the researcher's analysis, a new parameter (threshold variation) was developed so that the user could inform the image type in each request made to the detection service. The image types were: (i) default image; (ii) high aphid population; (iii) high population of parasitoids; and (iv) high population of aphids and parasitoids. Thus, the threshold was reduced to 75% in option (ii), and it was set to 70% in option (iii). In option (iv), both thresholds were reduced. With this new feature, the fifteen validation series were reprocessed, applying thresholds of 75% and 70% in the images referring to the peak weeks of aphid or parasitoid infestation, respectively.

The coefficients generated after reprocessing the series are also listed in Tables 6, 7, and 8, in the columns with the suffix (VL), representing the threshold variation. A significant increase in the slope coefficient can be observed, which represents a more accurate fit to detect variations in high aphid populations.

Table 6
Aphid coefficients in 2019.

Trap	Angular coefficient	Angular coefficient (VL)	Intercept	Intercept (VL)	R^2	R^2 (VL)	Peak
4	0.67 * X	0.92 * X	+ 1.22	- 0.84	0.99	0.99	144
0	0.67 * X	0.93 * X	+ 3.00	- 1.34	0.99	0.99	271
2	0.61 * X	0.98 * X	+ 2.01	- 3.26	0.98	1.00	276
1	0.40 * X	0.81 * X	+ 10.77	+ 1.58	0.90	0.99	614
3	0.39 * X	0.75 * X	+ 10.71	+ 5.63	0.91	0.98	630
All	**0.44 * X**	**0.81 * X**	**+ 8.7**	**+ 2.45**	**0.89**	**0.98**	**1735**

Table 7
Coefficients for aphids in 2020.

Trap	Angular coefficient	Angular coefficient (VL)	Intercept	Intercept (VL)	R^2	R^2 (VL)	Peak
0	0.54 * X	0.68 * X	+2.99	+2.13	0.91	0.83	79
1	0.25 * X	0.64 * X	+11.27	+6.17	0.65	0.95	324
3	0.26 * X	0.56 * X	+9.27	+6.27	0.77	0.82	484
4	0.28 * X	0.59 * X	+8.95	+5.12	0.77	0.92	604
2	0.30 * X	0.59 * X	+11.47	+ 4.03	0.87	0.91	686
All	**0.28 * X**	**0.58 * X**	**+9.52**	**+5.40**	**0.80**	**0.90**	**1954**

Table 8
Coefficients for parasitoids in 2020.

Trap	Angular coefficient	Angular coefficient (VL)	Intercept	Intercept (VL)	R^2	R^2 (VL)	Peak
0	0.53 * X	0.73 * X	+ 0.12	+ 0.44	0.89	0.93	24
1	0.49 * X	0.63 * X	+ 0.64	+ 0.60	0.80	0.85	14
4	0.32 * X	0.67 * X	+ 0.67	- 0.23	0.64	0.92	29
3	0.41 * X	0.72 * X	+ 0.21	- 0.03	0.92	0.85	30
2	0.40 * X	0.77 * X	+ 0.13	−0.03	0.75	0.92	38
All	**0.40 * X**	**0.73 * X**	**+ 0.45**	**+ 0.07**	**0.78**	**0.90**	**129**

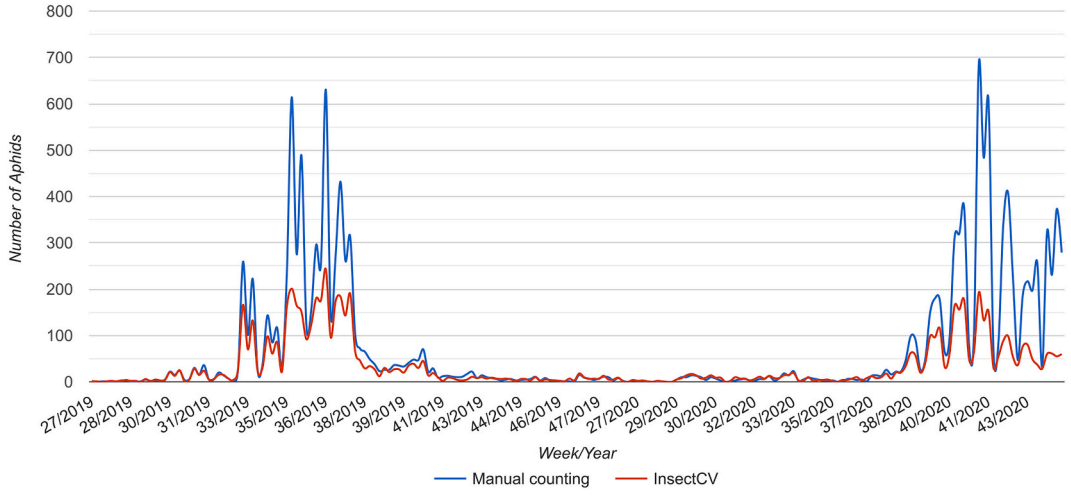

Fig. 3. Adjustment for aphids. Periods: week 27 to 48 of 2019, and week 27 to 44 of 2020. $N = 200$.

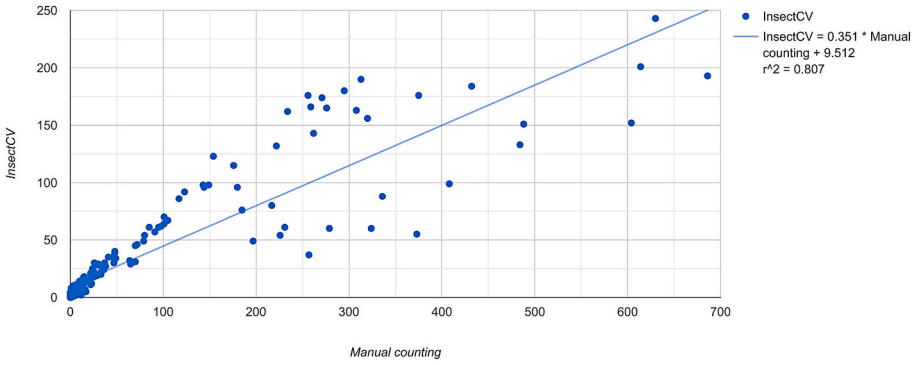

Fig. 4. Relation between the number of aphids obtained by manual and automatic counting using InsectCV.

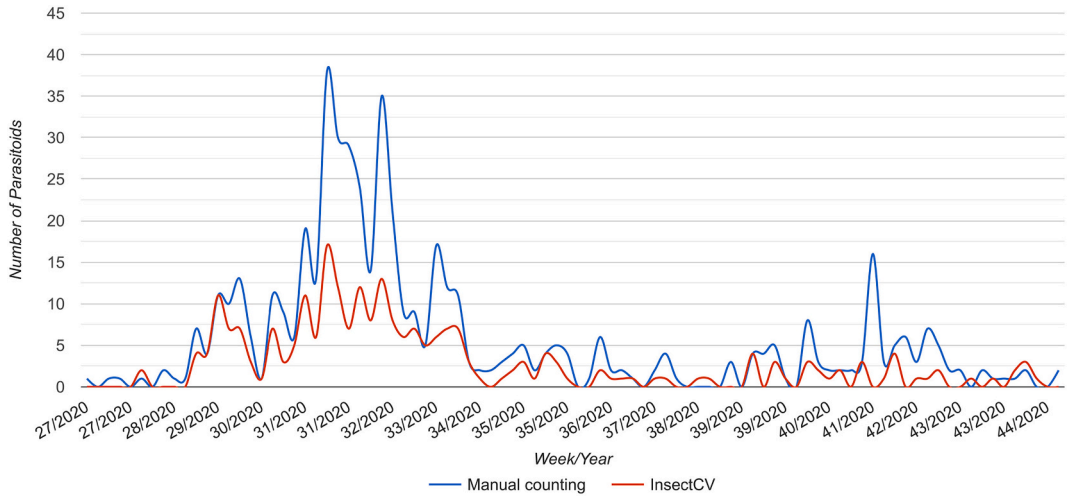

Fig. 5. Adjustment for parasitoids, 2020 crop. $N = 90$.

Table 9
Metrics for parasitoids.

Trap	True Positives (TP)	False Positives (FP)	False Negatives (FN)	Precision	Recall	F1 score
1	32	9	29	0.78	0.52	0.63
2	59	8	103	0.88	0.36	0.52
3	46	6	63	0.88	0.42	0.57
4	36	7	57	0.84	0.39	0.53
All	173	30	252	0.85	0.41	0.55

used for the integrated pest management. Therefore, the assessment of accuracy and precision of the detection of epidemiologically established population critical points must be performed. Delays in detecting populations at these levels imply late adoption of management measures. With this in mind, three cut-off points were tested: 10, 20 and 50 aphids per trap. These cut-off points were established considering the historical oscillation patterns of populations of winged aphids trapped in Southern Brazil (Engel et al., 2021). Empirically, populations between 10 and 20 aphids per trap indicate 10% of plants infested by aphids. This indicative is one of the parameters used for chemical application in the wheat crop

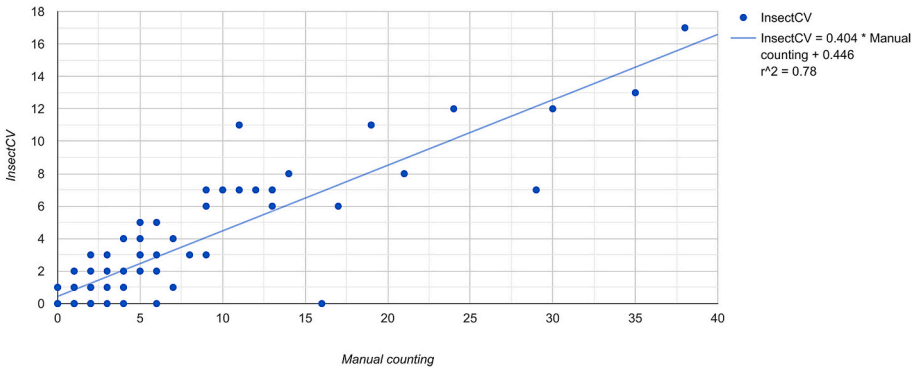

Fig. 6. Relation between the number of parasitoids obtained by manual and automatic counting using InsectCV.

4.2. Metrics

The results obtained with the application of the metrics *precision*, *recall*, and *F1 score* are presented in Table 9.

The analysis of the metric values showed that, for objects classified as a parasitoid, the model was right in 85% of the tested cases. Nevertheless, when an insect is indeed a parasitoid and the model did not detect them (false negatives), the frequency at which the model correctly classified the object was 41%. According to this finding, the strategy of using high thresholds reduced false positives. However, 60% of the parasitoids were not identified. By combining metrics of *precision* and *recall*, the value obtained for the *F1 score* was 0.55.

4.3. Cut-off points

One of the main applications of automated counting aphids in traps aims to expand the monitoring network, supporting warning systems in Southern Brazil, depending on the crop phenological stage.

According to the dotted gray and solid black lines in Fig. 7, considering aphid population dynamics in the 2019 crop, in the five validation series the peak was reached in week 35 (S35), with 347 aphids/trap/week for the manual counting and 154 aphids/trap/week for InsectCV.

Considering intermediate thresholds for management decision-making until reaching the peak, the threshold of 10 aphids/trap/week was exceeded in S30 for manual counting (14 aphids) and similarly for InsectCV (12.8 aphids). From that, the number of captured aphids increased rapidly and the threshold of 20 aphid/trap/week aphids was reached in S33 (127 aphids in the manual counting and 84 for the InsectCV), also exceeding the threshold of 50 aphids.

In the 2020 wheat crop (weeks 27 to 44), considering the set of five series under analysis, as shown in the yellow and gray lines in Fig. 7, the peak was detected in S41 by both methods, with 391 aphids/trap/week for manual counting and 136 aphids/trap/week for InsectCV. Considering intermediate thresholds for management decision-making until

reaching the peak, the threshold of 10 aphids/trap/week was exceeded in S29 for manual counting (10 aphids) and similarly for InsectCV (13 aphids).

From that moment on, there was a decrease in the number of captured winged aphids, reaching the level again in S33 (13 aphids in manual counting and 14 by InsectCV). Finally, there was a drop and then the curve resumed with sustained growth with the threshold of 10 aphids in S37 (16,4 aphids in manual counting and 11 by InsectCV), and 20 aphids in S38 (Manual 56 aphids, InsectCV: 43 aphids). Only for the threshold of 50 aphids per trap, there was a one-week delay between detection and counting methods (S38 for manual and S39 for InsectCV).

5. Discussion

The problem presented in this study was more challenging compared to the cited studies, as in the reported scenarios there was no possibility of false positives with similar morphology, pose variation and the presence of significant amounts of other objects or debris. After capturing the aphids and parasitoids, there was a natural process of decomposition and reaction with the detergent used in traps, which changed the insect's color. For this reason, grayscale images were used, which eliminated color variations, thus reducing the number of features available for model training. The analysis of this study was also more complex due to the significant confusion with other insects in the sample that were not objects of interest, compared with *AphidCV* (Lins et al., 2019) that used images containing only the objects of interest (aphids).

Compared to the studies listed in Table 1, InsectCV is similar to those that presented situations of occlusion and overlapping objects, marked as "Yes" in the "Overlapping" column, as well as related work that used the Faster R-CNN method. In addition, we also highlight the study of Kalamatianos et al. (Kalamatianos et al., 2018) because the use of images with the "pest soup" aspect. However, in these studies, there was no significant presence of debris and/or similar insects, which could reduce the accuracy of the model. About the data set used for training, the number of images used in InsectCV was smaller compared to most of the other studies. The dimension of the images for training and inference was greater concerning the studies listed in Table 1.

The option for Mask R-CNN, instead of Faster R-CNN, is also justified by the possibility of using the instance segmentation in future work, enabling to sum the areas of all objects in the sample to measure the levels of biological activity around the trap. Furthermore, we could correlate the reduction in the identification of insects of interest accuracy to the number of residues and other insects increasing.

The results obtained with the validation of the model showed that through a small set of images (209), a significant amount of labeled objects (17,908), and the application of customizations, a model was generated in this study, using the Mask R-CNN method. This model is capable of identifying variations in population levels of aphids and parasitoids in traps based on patterns found in images.

For the detection of parasitoids, two additional factors related to aphids were identified, which reduced the model accuracy. According to the detection example presented in Fig. 2, the parasitoids are similar to other insects, such as flies and thrips (Thysanoptera), and may also have variations related to the angle of the tail, joining it with the insect body.

Notwithstanding the low number of parasitoids in the images of 2020, compared to aphids, this situation still allowed us to evaluate the model's performance based on the coefficients. However, the values found for the *precision* and *recall* metrics emphasize the need to balance this relationship with threshold adjustments. These metrics were not applied to aphids, as Embrapa performs the manual identification of physical samples with the aid of a magnifying glass or microscope to more quickly locate the morphological characteristics of the specimens, therefore, not labeling objects in digital images. In the case of parasitoids, due to the smaller number, it was possible to manually count true positives, false positives, and false negatives, enabling the generation of metrics.

Despite not having this information, we could conclude that the model can identify the variation in population levels of insects in traps according to the angular coefficient, distortion in the intercept, and coefficient of determination. These coefficients were obtained comparing the manual identification and at the cut-off points (Fig. 7: low deviation in the initial phases of the growth curve and the maintenance of the shape of this curve in the analysis of two wheat crop periods). Furthermore, considering the coefficients presented in Fig. 4 for aphids and in Fig. 6 for parasitoids, we did not notice significant variation, therefore, we understand that the values of *precision* (0.85), *recall* (0.41) and *F1 Score* (0.55) for parasitoids tend to be equivalent/close for aphids.

The model was able to detect connected aphids, as shown in Fig. 1. However, in cases of dense agglomerations with several insects, the model was not able to recognize most of the insects in these regions. In general, considering the three types of images used for validation, the model obtained satisfactory accuracy based on the coefficients presented, especially in the 2019 crop. According to Embrapa, in 2019 there was less proliferation of other insects compared to 2020, due to the drought periods that occurred in 2020.

According to the results, the deviations between manual counting and InsectCV escalated with increasing population density. Despite these deviations, the shape of the winged aphid capture curve over time was the same for the two wheat crops analyzed. Examining the low deviation in the initial phases of the population growth curve and the maintenance of the shape of this curve, the use of InsectCV allowed the identification of the critical decision-making points for pest management. Moreover, InsectCV showed the epidemic peak with the same precision as the manual counting, even in weeks of high insect counting where the model underestimated the population.

Due to the hardware resources available to perform the experiments (Intel Core I7-6950×, 32 GB RAM, and GeForce GTX Titan X GPU with 12 GB VRAM) and the image size (6156 × 6156), the number of trained layers could not be increased, due to the need for more VRAM memory.

The image inference process for model validation was also hindered by the lack of RAM and GPU memory in the IFSul server (Intel Xeon E31240, 8 vCPUs, 14 GB RAM). Consequently, the researchers had to resize the image to 5568 × 5568 pixels, which may have reduced the amount of information analyzed by the model. The time required to perform the inference of each image through the detection service ranged from 3 to 6 min. This time variation is usual in two-stage convolutional approaches.

The time to perform each training epoch was approximately 9 h, resulting in 35 days to complete experiment 7. As an alternative to reduce this time, studies can use the strategy mentioned by Bobba (Bobba, 2019). This strategy consists of using smaller image clippings (512 × 512) for faster updating of weights and applying larger clippings (2048 × 2048) in the final phase to adjust the parameters of the final model. Another possibility is to split the original image into smaller images and reposition the labels. To reduce inference time and eliminate resizing, the sliding window technique can be applied, performing the inference in smaller parts and repositioning the detected objects in the original image at the end of this process.

By using the feature to adjust the thresholds based on the type of image, the angular coefficient (VL), Intercept (VL), and coefficient of determination R^2 (VL), this study had a significant increase in the accuracy of the model for aphids and parasitoids in situations of high population density (Tables 7 and 8). This feature can be enhanced by using a color image to represent the detection result so that the researcher has more information to determine the proper threshold.

Looking into the coefficients highlighted in bold in Tables 6 and 7, and examining the cut-off analysis, data shows the model correctly identified variations in population levels, based on the 580 images used for validation. However, at times of greater proliferation of insects and low levels of rainfall, the accuracy of the model tends to be lower, due to the presence of other insects and connections between aphids. Thus, the

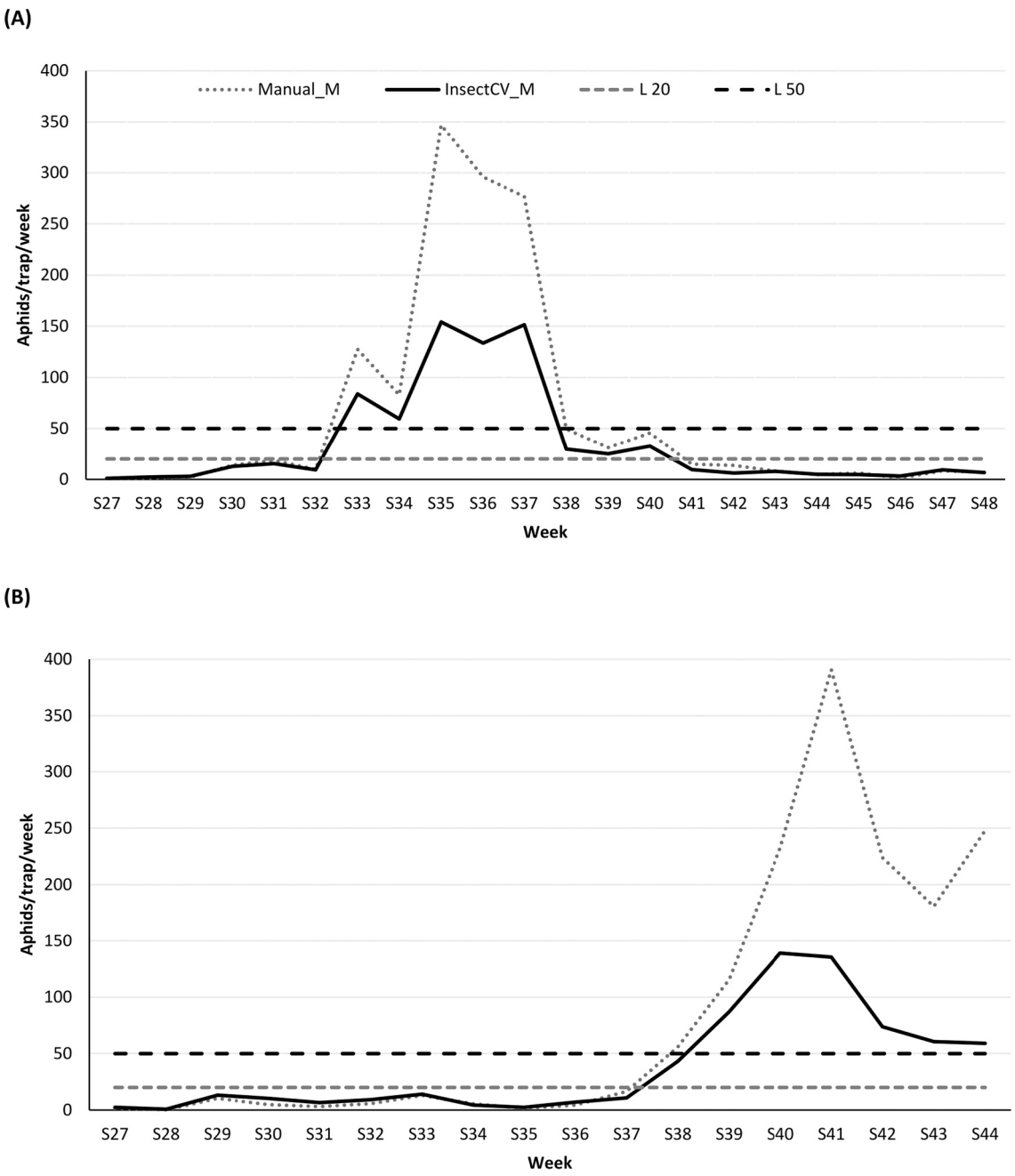

Fig. 7. Aphid population progress curve for wheat crops in two years and two counting methods. (A) 2019 and (B) 2020. Data correspond to an average from 5 traps. Dotted gray: manual counting. Solid black: computer vision count (InsectCV). L20 gray dashed line, refers to the preset threshold of 20 aphids per trap per week. L50 black dashed line, refers to the preset threshold of 50 aphids per trap per week.

importance of evaluating management procedures is emphasized to further improve the developed models.

The reduction in accuracy when counting aphids in dense agglomerations might be mitigated by adopting management procedures such as (i) splitting the sample into two or more images; (ii) relocation or elevation of traps; (iii) use of a larger Petri dish. These changes can contribute to the reduction of connections and the amount of debris, increasing the accuracy of the model.

Systems such as InsectCV enable the availability of monitoring networks on a larger scale. What would the implications be if this becomes reality? In theory, the use prophylactically or scheduled of insecticides would be reduced. On the other hand, showing data on aphid populations can raise awareness about phytosanitary problems not explicitly perceived or diagnosed by farmers. So, the system itself would not reduce the number of applications but make them more assertive and efficient.

Another aspect concerns the degree of development of monitoring tools. In the current version, InsectCV only allows the counting of total aphid populations, not discriminating between species that have different relevance in terms of damage to crop yield and, as a result, not discerning between viruliferous and non-viruliferous aphids. Thus, although there are similarities between the oscillation patterns of total populations and oscillations of some species (Engel et al., 2021), these are, without a doubt, limitations to be addressed in future advances needed for more accurate warning systems.

In addition, the case study was demonstrated for winter cereal aphids. Concerning other winter cereals, such as barley and triticale, our experience indicates dynamics are very similar to wheat. Obviously, the

entire calibration and validation process needs to be carried out for different cultures and at other times of the year.

6. Conclusion

This study presented InsectCV, a system for identifying and counting aphids and parasitoids from trap images. We generated the image dataset from digitized samples captured by Moericke-type traps at Embrapa's experimental stations. For the development of the intelligent model, we present an approach based on the Mask R-CNN method. The model demonstrated the ability to estimate the fluctuation of population levels for these insects in traps. In addition, it was able to distinguish cases of connections and overlaps in images that contained hundreds of objects.

The main contribution of our work is related to the applicability of InsectCV as a means to facilitate the counting of insects in traps, automatic storage into an agricultural database, and contribute with data sources to support decision-making for aphid pest management. Likewise, automated parasitoid counting is fundamental to estimate by models the suppression that this biological control agent exerts on aphids predicting aphid population growth.

About the evolution of the developed system, one of the improvements concerns the precision adjustment of the acceptance threshold for parasitoids and aphids, which were initially defined as 0.985 and 0.976, respectively, through empirical tests performed with the researcher. We intend to label all images used for model validation and add new images from previous and later crop seasons, aiming at generating metrics and analysis of model performance with the presence of other insect species.

Thus, new studies will be able to evaluate different threshold values using a precision-recall curve or a ROC curve (Davis and Goadrich, 2006; Lee et al., 2018). Another suggestion is to increase the graphics processing capabilities with new GPUs. In addition, other algorithms should be tested to assess the ability to recognize connected and/or overlapping insects of the Mask R-CNN through the Intersection over Union (IoU) metric.

Regarding the alternatives to increase model accuracy, the results of this study showed that new images containing parasitoids in different poses and with different tail positions should be added to the set of images for training. Considering trap screening and image collection, field procedures should be reviewed to reduce the generation of dense images, which represent a convoluted environment for analysis.

Another future resource to the system could be counting the amount of material present in the sample, considering the sum of the area of objects in the image. This calculation could be performed with mask data, which was not explored in this study. The calculation would measure the levels of biological activity around the trap, and would also correlate the reduction in the accuracy of pest insects identification with the increase in the amount of debris and other insects.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abdulla, W., 2017. Mask R-CNN for Object Detection and Instance Segmentation on Keras and Tensorflow. URL. https://github.com/matterport/Mask_RCNN.

Akintayo, A., Tylka, G.L., Singh, A.K., Ganapathysubramanian, B., Singh, A., Sarkar, S., 2018. A deep learning framework to discern and count microscopic nematode eggs. Sci. Rep. 8 (1) https://doi.org/10.1038/s41598-018-27272-w.

Ärje, J., Melvad, C., Jeppesen, M.R., Madsen, S.A., Raitoharju, J., Rasmussen, M.S., Iosifidis, A., Tirronen, V., Gabbouj, M., Meissner, K., Høye, T.T., 2020. Automatic image-based identification and biomass estimation of invertebrates. Methods Ecol. Evol. 11 (8), 922–931 arXiv: https://doi.org/10.1111/2041-210X.13428.

Bhattarai, U., Bhusal, S., Majeed, Y., Karkee, M., 2020. Automatic blossom detection in apple trees using deep learning. IFAC-PapersOnLine 53 (2), 15810–15815, 21th IFAC World Congress. https://doi.org/10.1016/j.ifacol.2020.12.216. URL. https://www.sciencedirect.com/science/article/pii/S2405896320304857.

Bobba, R., 2019. Taming the Hyper-parameters of Mask RCNN, Analytics Vidhya - Community of Analytics and Data Science Professionals. URL. https://medium.com/analytics-vidhya/taming-the-hyper-parameters-of-mask-rcnn-3742cb3f0e1b.

Brabec, M., Honěk, A., Pekár, S., Martinková, Z., 2014. Population dynamics of aphids on cereals: digging in the time-series data to reveal population regulation caused by temperature. PLoS One 9 (9), e106228.

Chen, J., Fan, Y., Wang, T., Zhang, C., Qiu, Z., He, Y., 2018. Automatic segmentation and counting of aphid nymphs on leaves using convolutional neural networks. Agronomy 8 (8). https://doi.org/10.3390/agronomy8080129.

Davis, J., Goadrich, M., 2006. The relationship between precision-recall and ROC curves. In: Proceedings of the 23rd International Conference on Machine Learning, 06. ACM, p. 8. https://doi.org/10.1145/1143844.1143874.

Davis, T.S., Abatzoglou, J.T., Bosque-Pérez, N.A., Halbert, S.E., Pike, K., Eigenbrode, S. D., 2014. Differing contributions of density dependence and climate to the population dynamics of three eruptive herbivores. Ecol. Entomol. 39 (5), 566–577 arXiv.

De Cesaro Júnior, T., Rieder, R., 2020. Automatic identification of insects from digital images: a survey. Comput. Electron. Agric. 178, 105784. https://doi.org/10.1016/j.compag.2020.105784. URL. http://www.sciencedirect.com/science/article/pii/S0168169920311224.

Ding, W., Taylor, G., 2016. Automatic moth detection from trap images for pest management. Comput. Electron. Agric. 123, 17–28. https://doi.org/10.1016/J.COMPAG.2016.02.003.

Döring, T.F., 2014. How aphids find their host plants, and how they don't. Ann. Appl. Biol. 165 (1), 3–26.

Embrapa Trigo, 2017. Treinamento ajuda pesquisadores na identificação de insetos em cereais de inverno. URL. https://www.embrapa.br/busca-de-noticias/-/noticia/26089616/treinamento-ajuda-pesquisadores-na-identificacao-de-insetos-em-cereais-de-inverno.

Embrapa Trigo, 2019. Treinamento para aprimorar técnicas de identificação de insetos em cereais de inverno. URL. https://www.embrapa.br/busca-de-noticias/-/noticia/43111123/treinamento-para-aprimorar-tecnicas-de-identificacao-de-insetos-em-cereais-de-inverno.

Engel, E., Lau, D., Godoy, W.A.C., Pasini, M.P.B., Malaquias, J.B., Santos, C.D.R., Pivato, J., Pereira, P.R.V.D.S., 2021. Oscillation, synchrony, and multi-factor patterns between cereal aphids and parasitoid populations in southern Brazil. Bull. Entomol. Res. 1–8. https://doi.org/10.1017/S0007485321000729.

Fischer, P., Dosovitskiy, A., Brox, T., 2014. Descriptor matching with convolutional neural networks: a comparison to SIFT. CoRR abs/1405.5769, 1–10, 1405.5769. URL. http://arxiv.org/abs/1405.5769.

Fuentes, A., Yoon, S., Kim, S.C., Park, D.S., 2017. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. Sensors (Switzerland) 17 (9). https://doi.org/10.3390/s17092022.

Ganesh, P., Volle, K., Burks, T., Mehta, S., 2019. Deep Orange: Mask R-CNN based orange detection and segmentation, IFAC-PapersOnLine 52 (30) (2019) 70–75. In: 6th IFAC Conference on Sensing, Control and Automation Technologies for Agriculture AGRICONTROL. https://doi.org/10.1016/j.ifacol.2019.12.499. URL. https://www.sciencedirect.com/science/article/pii/S2405896319324152.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in. IEEE Conf. Comp. Vis. Patt. Recog. (CVPR) 2016, 770–778. https://doi.org/10.1109/CVPR.2016.90.

He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask R-CNN, in. IEEE Int. Conf. Comp. Vis. (ICCV) 2017, 2980–2988. https://doi.org/10.1109/ICCV.2017.322.

Hodgson, E.W., McCornack, B., Tilmon, K., Knodel, J., 2012. Management recommendations for soybean aphid (Hemiptera: Aphididae) in the United States. J. Integrated Pest Manage. 3 (1), E1–E10.

Howard, M., Dixon, A., 2008. Forecasting of peak population density of the rose grain aphid Metopolophium dirhodum on wheat. Ann. Appl. Biol. 117, 9–19. https://doi.org/10.1111/j.1744-7348.1990.tb04190.x.

Høye, T.T., Ärje, J., Bjerge, K., Hansen, O.L.P., Iosifidis, A., Leese, F., Mann, H.M.R., Meissner, K., Melvad, C., Raitoharju, J., 2021. Deep learning and computer vision will transform entomology. Proc. Natl. Acad. Sci. 118 (2) arXiv:

Jarošová, J., Żelazny, W.R., Kundu, J.K., 2019. Patterns and predictions of barley yellow dwarf virus vector migrations in Central Europe. Plant Dis. 103 (8), 2057–2064 pMID: 31241012. arXiv:

Jia, W., Tian, Y., Luo, R., Zhang, Z., Lian, J., Zheng, Y., 2020. Detection and segmentation of overlapped fruits based on optimized mask R-CNN application in apple harvesting robot. Comput. Electron. Agric. 172, 105380. https://doi.org/10.1016/j.compag.2020.105380. URL. https://www.sciencedirect.com/science/article/pii/S0168169919326274.

Kalamatianos, R., Karydis, I., Doukakis, D., Avlonitis, M., 2018. DiRT: The DACUS image recognition toolkit. J. Imag. 4 (11) https://doi.org/10.3390/jimaging4110129.

Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: a survey. Comput. Electron. Agric. 147, 70–90. https://doi.org/10.1016/j.compag.2018.02.016.

Krizhevsky, A., Sutskever, I., Hinton, G., 2012. ImageNet classification with deep convolutional neural networks. Neural Inform. Proces. Syst. 25 https://doi.org/10.1145/3065386.

Lazzaretti, A.T., Fernandes, J.M.C., Pavan, W., Toebe, J., Wiest, R., 2016. AgroDB–integration of database management systems with crop models. In: Proceedings of the 8th International Congress on Environmental Modelling and Software. iEMSs Society, pp. 194–201.

Lazzaretti, A.T., Lau, D., Fernandes, J.M.C., Wiest, R., Bavaresco, J.L.B., Schaefer, F., Rieder, R., De Cesaro Júnior, T., 2021. Trap System. URL. http://gpca.passofundo.ifsul.edu.br/traps/.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444. https://doi.org/10.1038/nature14539.

Lee, U., Chang, S., Putra, G.A., Kim, H., Kim, D.H., 2018. An automated, high-throughput plant phenotyping system using machine learning-based plant segmentation and image analysis. PLoS One 13 (4), e0196615.

Li, W., Chen, P., Wang, B., Xie, C., 2019. Automatic localization and count of agricultural crop pests based on an improved deep learning pipeline. Sci. Rep. 9 (1) https://doi.org/10.1038/s41598-019-43171-0.

Li, Y., Qi, H., Dai, J., Ji, X., Wei, Y., 2017. Fully convolutional instance-aware semantic segmentation, in. IEEE Conf. Comp. Vis. Pattern Recogn. (CVPR) 2017, 4438–4446.

Lins, E.A., Rieder, R., Rodriguez, J.P.M., Scoloski, S.I., Lau, D., Pivato, J., Lima, M.B., Fernandes, J.M.C., Pereira, P.R.V.S., 2019. AphidCV, Patent: Computer program. Register number: BR512019000884–7, Brazilian National Institute of Industrial Property.

Lins, E.A., Rodriguez, J.P.M., Scoloski, S.I., Pivato, J., Lima, M.B., Fernandes, J.M.C., da Silva Pereira, P.R.V., Lau, D., Rieder, R., 2020. A method for counting and classifying aphids using computer vision. Comput. Electron. Agric. 169, 105200. https://doi.org/10.1016/j.compag.2019.105200. URL. http://www.sciencedirect.com/science/article/pii/S0168169919306039.

Liu, L., Wang, R., Xie, C., Yang, P., Wang, F., Sudirman, S., Liu, W., 2019. PestNet: an end-to-end deep learning approach for large-scale multi-class pest detection and classification. IEEE Access 7, 45301–45312. https://doi.org/10.1109/ACCESS.2019.2909522.

J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation: 2015 IEEE Conf. Comp. Vis. Patt. Recog. (CVPR), 2015, pp. 3431–3440.

Morris, R., 2018. First experiences with water traps. Leicestershire Entomological Society. Occasional Publication Series [LESOPS]. 36, 1–8.

Nazri, A., Mazlan, N., Muharam, F., 2018. PENYEK: Automated brown planthopper detection from imperfect sticky pad images using deep convolutional neural network. PLoS One 13 (12). https://doi.org/10.1371/JOURNAL.PONE.0208501.

Pang, Y., Shi, Y., Gao, S., Jiang, F., Veeranampalayam-Sivakumar, A.-N., Thompson, L., Luck, J., Liu, C., 2020. Improved crop row detection with deep neural network for early-season maize stand count in UAV imagery. Comput. Electron. Agric. 178, 105766. https://doi.org/10.1016/j.compag.2020.105766. URL. https://www.sciencedirect.com/science/article/pii/S0168169920311376.

Partel, V., Nunes, L., Stansly, P., Ampatzidis, Y., 2019. Automated vision-based system for monitoring asian citrus psyllid in orchards utilizing artificial intelligence. Comput. Electron. Agric. 162, 328–336. https://doi.org/10.1016/j.compag.2019.04.022.

Ren, S., He, K., Girshick, R., Sun, J., 2016. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell. 39 (6), 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031.

Reyes-Yanes, A., Martinez, P., Ahmad, R., 2020. Real-time growth rate and fresh weight estimation for little gem romaine lettuce in aquaponic grow beds. Comput. Electron. Agric. 179, 105827. https://doi.org/10.1016/j.compag.2020.105827. URL. https://www.sciencedirect.com/science/article/pii/S0168169920309546.

Ruiz-Santaquiteria, J., Bueno, G., Deniz, O., Vallez, N., Cristobal, G., 2020. Semantic versus instance segmentation in microscopic algae detection. Eng. Appl. Artif. Intell. 87, 103271. https://doi.org/10.1016/j.engappai.2019.103271. URL. https://www.sciencedirect.com/science/article/pii/S0952197619302398.

Santos, C., Sampaio, M., Lau, D., Redaelli, L., Jahnke, S., Pivato, J., Carvalho, F., 2019. Taxonomic status and population oscillations of Aphidius colemani species group (Hymenoptera: Braconidae) in southern Brazil. Neotrop. Entomol. 48 (6), 983–991.

Savaris, M., Lampert, S., Salvadori, J., Lau, D., Pereira, P.D.S., Smaniotto, M., 2013. Population growth and damage caused by Rhopalosiphum padi (l.)(Hemiptera, Aphididae) on different cultivars and phenological stages of wheat. Neotrop. Entomol. 42 (5), 539–543.

Shen, Y., Zhou, H., Li, J., Jian, F., Jayas, D.S., 2018. Detection of stored-grain insects using deep learning. Comput. Electron. Agric. 145, 319–325. https://doi.org/10.1016/j.compag.2017.11.039.

Sun, Y., Liu, X., Yuan, M., Ren, L., Wang, J., Chen, Z., 2018. Automatic in-trap pest detection using learning for pheromone-based Dendroctonus valens monitoring. Biosyst. Eng. 176, 140–150. https://doi.org/10.1016/j.biosystemseng.2018.10.012.

Vo, S.A., Scanlan, J., Turner, P., Ollington, R., 2020. Convolutional neural networks for individual identification in the southern rock lobster supply chain. Food Control 118, 107419. https://doi.org/10.1016/j.foodcont.2020.107419. URL. https://www.sciencedirect.com/science/article/pii/S0956713520303352.

Wang, Y., Lv, J., Xu, L., Gu, Y., Zou, L., Ma, Z., 2020. A segmentation method for waxberry image under orchard environment. Sci. Hortic. 266, 109309. https://doi.org/10.1016/j.scienta.2020.109309. URL. https://www.sciencedirect.com/science/article/pii/S0304423820301370.

Xu, B., Wang, W., Falzon, G., Kwan, P., Guo, L., Chen, G., Tait, A., Schneider, D., 2020. Automated cattle counting using mask R-CNN in quadcopter vision system. Comput. Electron. Agric. 171, 105300. https://doi.org/10.1016/j.compag.2020.105300. URL. https://www.sciencedirect.com/science/article/pii/S0168169919320149.

Yu, Y., Zhang, K., Yang, L., Zhang, D., 2019. Fruit detection for strawberry harvesting robot in non-structural environment based on mask-RCNN. Comput. Electron. Agric. 163, 104846 https://doi.org/10.1016/j.compag.2019.06.001. URL. https://www.sciencedirect.com/science/article/pii/S0168169919301103.

Yue, Y., Cheng, X., Zhang, D., Wu, Y., Zhao, Y., Chen, Y., Fan, G., Zhang, Y., 2018. Deep recursive super resolution network with laplacian pyramid for better agricultural pest surveillance and detection. Comput. Electron. Agric. 150, 26–32. https://doi.org/10.1016/j.compag.2018.04.004.

Zhong, Y., Gao, J., Lei, Q., Zhou, Y., 2018. A vision-based counting and recognition system for flying insects in intelligent agriculture. Sensors 18 (5), 1489. https://doi.org/10.3390/s18051489.