Computers and Electronics in Agriculture 145 (2018) 319–325

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Detection of stored-grain insects using deep learning

Yufeng Shen[a], Huiling Zhou[a,*], Jiangtao Li[a], Fuji Jian[b], Digvir S. Jayas[b]

[a] *Beijing University of Posts and Telecommunications, PO Box 137, Road Xitucheng 10, Haidian District, Beijing 100876, China*
[b] *Department of Biosystems Engineering, University of Manitoba, Winnipeg, MB R3T 5V6, Canada*

ARTICLE INFO

Keywords:
Stored-grain insect
Object detection
Insect classification
Convolutional neural network
Trap images
Faster R-CNN

ABSTRACT

A detection and identification method for stored-grain insects was developed by applying deep neural network. Adults of following six species of common stored-grain insects mixed with grain and dockage were artificially added into the developed insect-trapping device: *Cryptoleste Pusillus*(S.), *Sitophilus Oryzae*(L.), *Oryzaephilus Surinamensis*(L.), *Tribolium Confusum*(Jaquelin Du Val), *Rhizopertha Dominica*(F.). Database of Red Green and Blue (RGB) images of these live insects was established. We used Faster R-CNN to extract areas which might contain the insects in these images and classify the insects in these areas. An improved inception network was developed to extract feature maps. Excellent results for the detection and classification of these insects were achieved. The test results showed that the developed method could detect and identify insects under stored grain condition, and its mean Average Precision (mAP) reached 88.

1. Introduction

Image recognition to detect and identify insects in a stored product is the critical component of a stored-grain insect monitoring system. The main challenges in the image recognition of these insects are to identify areas of the image containing insects in grain mixed with other materials (mostly the fine materials and broken grain kernels) and to classify the small body size insects conglutinated with other insect species and/or the same species and/or the other materials in the target area.

Object detection system such as pedestrian detection and vehicle detection applies the region proposal algorithms to infer locations of objects (Girshick et al., 2013). The early developed region proposal algorithms include Selective Search, Sliding Window, Rigor, Superpixels and Gaussian (Hosang et al., 2015). The Region Proposal Network (RPN) Ren et al., 2017 was proposed in 2016, which applied convolutional neural network method to get the areas of interest more quickly and accurately through the acceleration of Graphics Processing Unit (GPU).

In the field of the insects classification based on computer vision, most of researches focused on the extracting of insects' features including texture, shape, and local characteristics (Qiu et al., 2003; Zhang et al., 2009, 2005; Wu et al., 2015; Jayas, 2017). Procedure of this feature extraction was complex under in-situ situations, and those extracted features might not accurately represent the image characteristics of insects. In the practical application, the variation of image background, impurities, illumination and insect's gestures will also increase the difficulty of feature extraction.

Krizhevsky et al. (2012) used deep convolutional neural networks got first on ImageNet Large Scale Visual Recognition Challenge in 2012. Ding and Taylor (2016) used the Sliding Window method to obtain regions of interest and applied a 5-layer convolutional neural network to determine whether the regions contained a moth. They got a higher recall rate using convolutional neural network than that using LogReg algorithm. Liu et al. (2016) applied the GrabCut for the segmentation of paddy field pests and classified the pests using a 8-layer convolutional neural network. The accuracy of the convolutional neural network was higher than that of the Histogram of Oriented Gradient (HOG), Speeded Up Robust Features (SURF). However, the network used by Liu et al. was too shallow, and could not extract more effective features when the targets were similar in appearance. In this study, we used an Online Insect Trapping Device (OITD) (Wang et al., 2016) to capture the images of live insects with or without fines, foreign materials, dockages and broken grains (referred to as FFDB) under laboratorial conditions. The insects imaged were: *Cryptoleste Pusillus*(S.), *Sitophilus Oryzae*(L.), *Oryzaephilus Surinamensis*(L.), *Tribolium Confusum*(Jaquelin Du Val), *Rhizopertha Dominica*(F.) and *Lasioderma Serricorne*(F.). A dataset with 739 images of the insects with or without FFDB was established, and artificially marked. To increase the accuracy of the convolutional neural network, we applied a 27-layer convolutional neural network to extract the features from the images of stored-product insects, and adopted the Softmax as classifier to identify the insects. The Faster R-CNN method was used to locate and classify the insects. First, the inception network (Szegedy et al., 2015) extracted the feature maps of

* Corresponding author.
E-mail address: huiling@bupt.edu.cn (H. Zhou).

https://doi.org/10.1016/j.compag.2017.11.039
Received 18 June 2017; Received in revised form 16 October 2017; Accepted 26 November 2017
0168-1699/ © 2018 Elsevier B.V. All rights reserved.

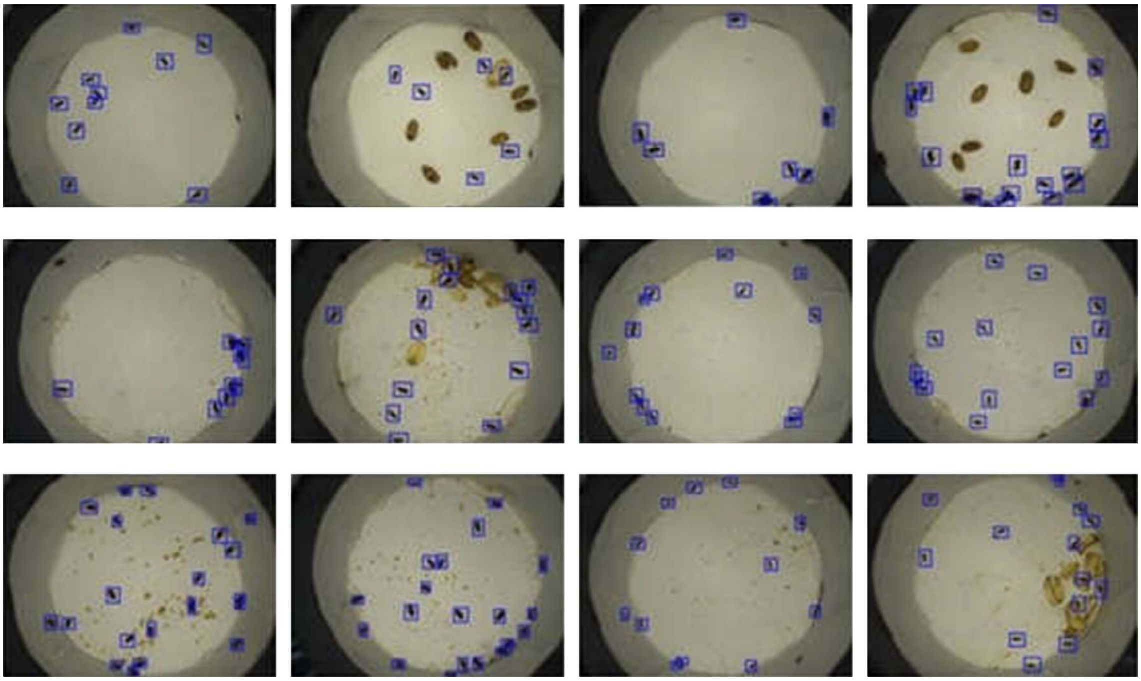

Fig. 1. Ground truth.

each image, then the RPN returned the coordinates of the areas which might have insects in the feature maps. These coordinates were merged by Non Maximum Suppression (NMS) Neubeck and Van Gool, 2006 and mapped to the feature maps. These target areas were classified by the improved inception network. The coordinates of the target areas were also corrected at the same time (Girshick, 2015). Finally, NMS was used to merge the overlapping target boxes.

In this study, we developed a method based on the Faster R-CNN, which can be used to identify the insects mixed with FFDB under different illumination conditions. This method improved the accuracy of the insect detection. The rest of this article is organized to introduce this image processing procedure.

2. Image acquisition and preprocessing

The resolution of the images taken by the OITD system was 1944 × 2592 pixels. These images were collected under the conditions of multi live insects mixed with artificially added FFDB in wheat in order to increase the difficulty of detection and simulate the real situation in grain warehouses. The FFDB included fine materials and broken grains (Fig. 1). Multiple live insects of single species were added into OITD system every time and multiple pictures were snapshotted. Some pictures were randomly chosen as the testing set, and the rest were used as the training set. This procedure was repeated and new insects were added for the next replicate. Table 1 shows the number of the images and insects in the training and testing sets.

To enrich the training set, extract image features accurately, and generalize model to prevent overfitting, the image dataset was augmented by flipping and color jittering. To account for the change of illumination level and insects' posture, the color jittering was conducted by randomly adjusting the saturation, contrast, brightness, and sharpness of the image. After augmentation, the size of training set was increased by 12 folds of the original training set. The original images had a high resolution. To improve the training and testing speed and to reduce the GPU consumption, the image resolution was lowered to 600 × 800 pixels. Each insect in the images was artificially marked by a blue bounding box (Fig. 1) as ground truth. The marked blue bounding box was used for training.

Table 1
Number of images and insects used for training and testing.

		SO[a]	LS[b]	TC[c]	RD[d]	OS[e]	CP[f]
Training	Images	77	54	98	114	63	117
	Insects	1206	1088	1423	1908	830	2957
Testing	Images	33	28	42	21	23	69
	Insects	465	514	446	335	358	978
Total	Images	110	82	140	135	86	186
	Insects	1671	1602	1869	2243	1188	3935

[a] SO = *Sitophilus Oryzae*.
[b] LS = *Lasioderma Serricorne*.
[c] TC = *Tribolium Confusum*.
[d] RD = *Rhizopertha Dominica*.
[e] OS = *Oryzaephilus Surinamensis*.
[f] CP = *Cryptolestes Pusillus*.

3. Object detection network

The detection steps for the target object were: acquire the region proposals of the image by RPN, merge these proposals as candidate boxes by NMS, map these candidates boxes to the feature maps, classify these regions of candidate boxes by classification network, use the NMS to merge these overlapping candidate boxes. The detection process is shown schematically in Fig. 2 and details are provided in the following sections.

3.1. Feature extraction network

In order to obtain high-quality model, the width (different sizes of kernels were used to extract the same feature maps) and depth of the neural network model should be increased. The inception structure (Szegedy et al., 2015) was adopted to extend the convolutional layers of the model. The main purpose of the inception structure was to find a simple dense component to replace an optimal local sparsity structure and repeated this structure spatially. This procedure would cluster together the units with high relative correlation to form the next layer and this next layer would connect to its top layer. The adopted inception structure is shown in Fig. 3. To reduce the number of parameters

Download English Version:

https://daneshyari.com/en/article/6539803

Download Persian Version:

https://daneshyari.com/article/6539803

Daneshyari.com