RESEARCH ON FARMLAND PEST IMAGE RECOGNITION BASED ON TARGET DETECTION ALGORITHM

Shi Wenxiu and Li Nianqiang

School of Information Science and Engineering,
University of Jinan, Jinan, China

ABSTRACT

In order to achieve the automatic identification of farmland pests and improve recognition accuracy, this paper proposes a method of farmland pest identification based on target detection algorithm .First of all, a labeled farm pest database is established; then uses Faster R-CNN algorithm, the model uses the improved Inception network for testing; finally, the proposed target detection model is trained and tested on the farm pest database, with the average precision up to 90.54%.

KEYWORDS

Object detection algorithm, Faster R-CNN, Inception network

1. INTRODUCTION

The harm of farmland pests is one of the main factors affecting agricultural production, and its economic impact has spread all over the world. In this case alone, the annual agricultural economic losses in Europe reach 28.2%, North America reaches 31.2%, and the economic losses in Asia and Africa are as high as 50%[1]. For a long time, we have mostly adopted manual identification and counting methods for Integrated Pest Management (IPM)[2], which is closely related to the comprehensive quality of professionals. The subjective factors have a great influence on the accuracy and timeliness of pest control, which is not conducive to the effective work of farmland pest control. Therefore, it is necessary to study a fast and low-cost automatic detection method of farmland pests and diseases.

Li Wenbin [3] used a single pixel background segmentation method based on RGB color to segment rice pest images, and used SVM support vector machine to identify and classify. Compared with traditional recognition methods, the recognition rate is better, but the number of pest image samples is small, and the identification type is single. Pan Chunhua [4] and others used a grid search algorithm to improve the search efficiency of the image target pest area, and used SVM to train multiple classifiers to identify four major vegetable pests, with an average recognition rate of 93%.Yang Guoguo [5] used Otsu algorithm to find threshold adaptively to complete image segmentation. The SVM classifier was used to classify and identify the Chinese rice locust, with an accuracy rate of 88.3%.Yang Wenhan [6] used Canny operator and Otsu threshold segmentation method to segment 15 cotton pests, and used binary tree classification and support vector machine for classification and recognition, and the recognition effect was good. The actual environment of real-life farmland pest classification is very complicated, and there are many types of pests. In order to achieve more effective and wider application of farmland pest

Natarajan Meghanathan et al. (Eds): SIPP, BIGML, DaKM, SOEN, AISC - 2020
pp. 111-117, 2020. © CS & IT-CSCP 2020 DOI: 10.5121/csit.2020.100210

detection technology, this paper combines deep learning and field pest detection, and proposes a field pest identification method based on target detection algorithm, which greatly improves the accuracy of field pest detection.

2. MODEL BUILDING

2.1. Farmland Pest Database

A good sample set is the basis of image recognition research [7]. Because no public data for farmland pest detection is currently available, this article has collected 2,472 farmland pest image samples through the Internet for the goals and tasks of pest recognition. In order to prevent overfitting due to insufficient data during training, this paper expands the training samples. The main methods of image data expansion include: resize, scale, and noise noise, rotate, flip, zoom, zoom, shear, shift, contrast, random channel shift, Principal Component Analysis (PCA), etc. Finally, a data set of farmland pests including 10 categories was compiled, as shown in Table 1, with a total of 12,474 images.

2.2. Object Detection Model Design

2.2.1. Target Detection Algorithm Faster R-CNN

In order to realize the end-to-end operation of the entire network, Faster R-CNN unifies the region suggestion algorithm on the convolutional neural network [8]. This method does not need to manually select the suggestion region and can fully utilize the features extracted by the neural network.

Regional suggestion network (RPN) [9] is a set that takes an image of any size as input and outputs a rectangular target recommendation box, as shown in Figure 1. The regional recommendation network is connected to the last layer of the feature convolution layer. A small 3×3 network sliding window is used on the feature map of this layer. Only one sliding traversal can extract candidate windows for the entire image, reducing the network calculation burden. The point in the center of the sliding window is called the anchor, and the position of each anchor point can be mapped to the original image, corresponding to the target suggested area on the original image. In order to make the recommendation window meet the target needs without size, the network adopts a multi-scale method, so each sliding window has three scale ratios, and the aspect ratio is 1: 1, 1: 2, 2: 1, and 9 types of suggestion windows are generated with three scales, and there are K suggestion windows on the right side of the figure.

As shown on the left side of Figure 1, each sliding window generates a recommendation window, and the recommendation window is mapped to a low-dimensional vector, which is output to two fully connected layers at the same level, bounding box regression layer (reg) and bounding box classification layer (els). The bounding box regression layer contains the position information of each window. A single suggestion window has four coordinate values to determine the accuracy of the suggestion window generated by RPN. The purpose of the bounding box classification layer is to output the score of the target category of the recommendation window. Each recommendation window has two outputs, corresponding to the probability that the target recommendation window is the foreground / background. For the bounding box generated by the region suggestion network, the part outside the target suggestion region will be discarded, and the remaining regions will be assigned multiple binary labels (target or background). If the highest overlap ratio of the prediction area overlapping the ground truth box (Intersection-over-Union, IoU) is greater than the defined threshold, a positive label is assigned to it. If the IOU ratio of the prediction area is lower than the defined threshold, a negative label is assigned to it. The definition of IoU is as follows:

$$IoU = \frac{area(B_{in\sec t} \cap B_{group})}{area(B_{in\sec t} \cup B_{group})} \quad (1)$$

$area(B_{in\sec t} \cap B_{group})$ represents the overlapping area of the target recommendation area and the ground truth area, $area(B_{in\sec t} \cup B_{group})$ represents the union of the target recommendation area and the ground truth area.

RPN and Fast R-CNN use the same loss function. The calculation formula for this multi-tasking loss function is as follows:

$$L(\{p_i\},\{p_i\}) = \frac{1}{N_{cls}}\sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}}\sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (2)$$

N is the total amount of anchors, i is the index of anchors in a mini-batch, and p_i is the predicted probability of the i-th anchor. If anchor is positive, the ground truth label p_i* is 1, and if anchor is negative, p_i* is 0.

t_i represents the four parameterized vectors of the coordinates of the predicted rectangular frame, and t_i* is the coordinate vector of the ground truth corresponding to the positive anchor. Classification loss (L_{cls}) is the log loss of two categories (foreground and background).

$$L_{cls}(p_i, p_i^*) = -\log[p_i^* p_i + (1 - p_i^*)(1 - p_i)] \quad (3)$$

Return loss is L_{reg} :

$$L_{reg}(t_i, t_i^*) = R(t_i - t_i^*) \quad (4)$$

2.2.2. Inception Network

This article considers the use of Googlenet Inception [10] structure to build an Inception network. A simple Inception network contains 22 layers of deep network. Because of the problem of blocked information flow, the optimization ability of the deep model is greatly reduced. To solve this problem, Googlenet [11] added two additional Softmax layers to calculate the new loss value, and then recalculated the network gradient based on the new loss value. Ballester [12] proposed the use of Shortcut Connection to reduce the effect of gradient disappearance. This structure can also reduce the Degradation phenomenon. However, some farmland pests have smaller targets. Using the simple concept structure can not solve the problem of pest detection and recognition in specific situations. Therefore, this paper uses an improved Inception network [13] for farmland pest detection,the improved neural network is shown in Figure 2.

As can be seen from Figure 2, the improved Inception [13] network uses the Shortcut Connection, and a deconvolution layer is connected behind the second Inception structure, and the feature map is doubled to the original one. At the same time, after extracting the feature map of the seventh perception structure, connecting a full connection layer will reduces the dimension of the feature map to the feature vector of 1024 dimension, Then the two feature vectors are stitched into a 2048-dimensional feature vector, and then the 2048-dimensional feature vector is reduced to a 1024-dimensional feature output as a feature output of the picture. And the Inception network can directly transfer the gradient from the deep layer of the network back to the shallow layer. At the same time, the Shortcut Connection can extract the shallow feature map. The scale of the shallow feature map is 1/8 of the original image. After the deconvolution layer, the feature map is increased to 1/4 of the original image. This kind of network combining multiple layers of features and different scales has better ability to detect and recognize targets, and can get better results in farmland pest recognition.

2.3. Experimental Results and Analysis

2.3.1. Experimental Results

When using the farm pest database detection model to train, in order to ensure that all samples can be used for training and testing, this paper uses K-fold cross-validation [14], where K is selected 10 and 9 subsets are selected for training Data, 1 subset as test data. The results of this experiment were evaluated using mean Average Precision (mAP). The formula is as follows:

$$mAP = \frac{\sum_{q=1}^{Q} AP(q)}{Q} \times 100\% \tag{5}$$

Q is the total number of pest categories, and AP(q) is the average accuracy rate of the detection results of category.

The average accuracy of the experimental results and the final average of the average accuracy are shown in Figure 3. From Figure 3, it can be seen that the average accuracy of detection of most types of farm pests is high, and the average average accuracy can be seen from the broken lines in the figure mAP reached 90.54%.

2.3.2. Experimental Comparison

In this paper, two detection methods are selected for comparison. The first is a method based on SVM [15] for agricultural pest detection proposed by Mundada et al. This paper does not preprocess the test data for the convenience of training data and the uniformity of comparison experiments. The second is a pest detection model proposed by Liu et al[16]. Which is modified on the native AlexNet network structure. In order to ensure the uniformity of training data, this paper replaces the fully connected layer in AlexNet with the Global Average Pooling (GAP) [17] can ensure that the input image data does not need to be of a fixed size. The same farmland pest data set was used to test with the method proposed in this paper and SVM and AlexNet models, and the comparison results are listed in Table 2. It can be found that the classification accuracy of the monitoring model used in this paper is improved by about 17% compared with AlexNet, and the overall detection classification accuracy has been greatly improved.

2.4. Figures and Tables

Table 1. Heading and text fonts.

Category	Quantity	Proportion
cutworms	1213	9.72%
aphidoidea	1220	9.78%
armyw	1202	9.64%
chafer	1205	9.66%
locust	1225	9.82%
plantho	1297	10.40%

psilogramma menephron	1315	10.54%
ostrinia furnacalis	1207	9.68%
clanis bilineata	1285	10.30%
helicoverpa armigera	1305	10.46%

Table 2. Comparison of test results

Method	**Mean precision /%**
SVM	52.89
AlexNet	73.57
Method of this article	90.54

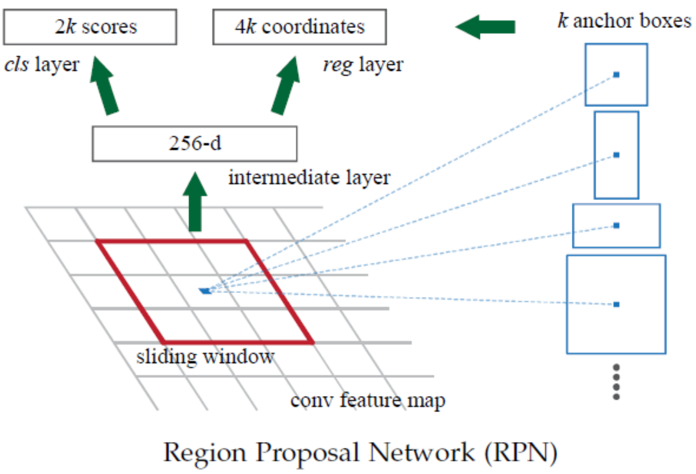

Figure 1. Faster R-CNN framework

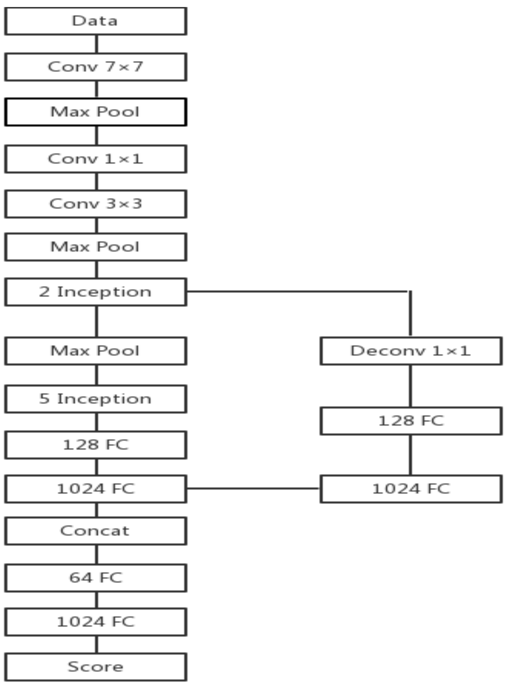

Figure 2. Inception network structure diagram

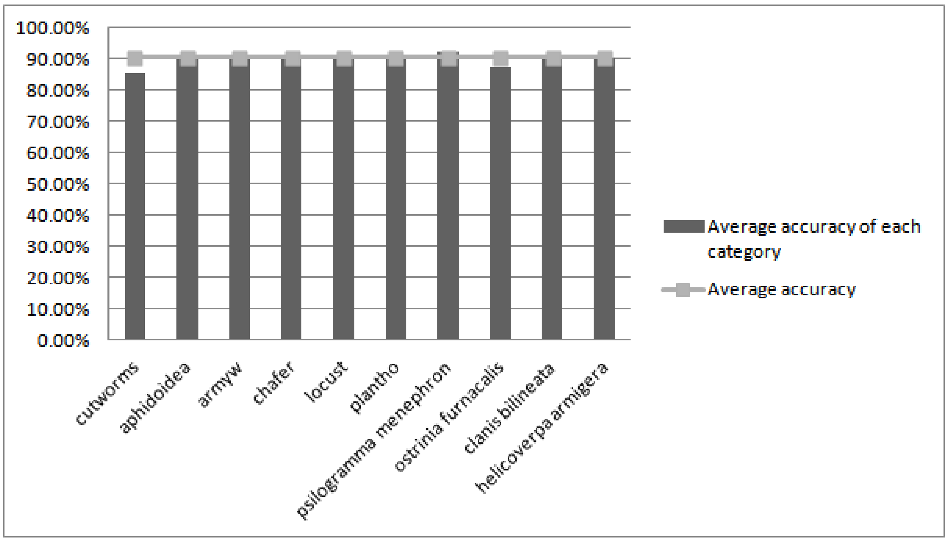

Figure 3. Farm pest detection results

3. CONCLUSIONS

Based on the idea of target detection algorithm in deep learning, this paper proposes a farmland pest detection model based on the target detection algorithm. This model combines a faster R-CNN algorithm and uses an improved Inception network, and test this model with a established farmland pest database. The detection accuracy rate of 90.54% was obtained. The experimental results show that the detection model proposed here can well perform the detection of farmland pests, and the detection results are accurate and the detection speed is fast. However, the method of this paper still has some shortcomings. The improved Inception network also needs to design a lot of hyperparameters. In the experiments, it is impossible to avoid the complicated hyperparameter tuning process, which brings the risk of overfitting to the recognition model.

ACKNOWLEDGEMENTS

The successful completion of this article is inseparable from the meticulous guidance and supervision of my tutor, as well as the help and encouragement of the lab brothers and sisters. Thank everyone.

REFERENCES

[1] Li Y , Xia C , Lee J . Detection of small-sized insect pest in greenhouses based on multifractal analysis[J]. Optik - International Journal for Light and Electron Optics, 2015, 126(19):2138-2143.

[2] Parsa S , Morse S , Bonifacio A , et al. Obstacles to integrated pest management adoption in developing countries[J]. Proceedings of the National Academy of Sciences, 2014, 111(10):3889-3894.

[3] Li Wenbin. Research on rice pest image recognition technology based on SVM [D]. Hangzhou University of Electronic Science and technology, 2015.

[4] Pan Chunhua, Xiao Deqin, Lin Tanyu, et al. Classification and identification of major vegetable pests in South China based on SVM and regional growth algorithm [J]. Journal of agricultural engineering, 2018 (8): 192-199

[5] Yang Guoguo. Identification and detection of early locust pupae of Chinese rice locust based on machine vision [D]. Zhejiang University, 2017.

[6] Yang Wenhan. Research on cotton pest identification system based on digital image processing [D]. Sichuan Agricultural University, 2015.

[7] Qin Fang. Insect image recognition based on deep learning [D]. Southwest Jiaotong University,2018:31-34

[8] Krizhevsky A , Sutskever I , Hinton G . ImageNet Classification with Deep Convolutional Neural Networks[C]// NIPS. Curran Associates Inc. 2012:25(2)

[9] Xia Denan. Research on agricultural insect image recognition based on deep learning [D]. Anhui University, 2019:26-30.

[10] Szegedy C , Liu W , Jia Y , et al. Going Deeper with Convolutions[J]. 2014.

[11] He K , Zhang X , Ren S , et al. Deep Residual Learning for Image Recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society, 2016.

[12] Ballester P L , Araujo R M . On the performance of GoogLeNet and AlexNet applied to sketches[C]// AAAI. AAAI Press, 2016.

[13] Shen Yufeng.Research on Stored Grain Pest Detection Algorithm Based on Deep Learning[D].2018:36-45.

[14] Hu Juxin,Zhang Gongjie.Selective ensemble classification algorithm based on K-fold cross validation[J].Science and Technology Bulletin,2013(12):123-125.

[15] Rupesh G. Mundada Rupesh G. Mundada.Detection and Classification of Pests in Greenhouse Using Image Processing[J].IOSR Journal of Electronics and Communication Engineering,2013,5（6）：57-63.

[16] Yang G , Bao Y , Liu Z . Localization and recognition of pests in tea plantation based on image saliency analysis and convolutional neural network[J]. Transactions of the Chinese Society of Agricultural Engineering, 2017, 33(6):156-162.

[17] He Kaiming, Gkioxari Georgia, Dollar Piotr,et. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence:1-1.

AUTHOR

Shi Wenxiu, 1995.03.01, School of Information Science and Engineering, University of Jinan, Postgraduate, Mainly engaged in image processing research.