2021 IEEE 4th International Conference on Electronics Technology

Agricultural Pest Detection System Based on Machine Learning

Shanshan Zhang
School of Information Science and Engineering
University of Jinan
Jinan, China
sliviawt@163.com

Junsheng Zhu
Plant Protection Department
Shandong Provincial Plant Protection Station
Jinan, China
120976967@qq.com

Nianqiang Li
School of Information Science and Engineering
University of Jinan
Jinan, China
ise_linq@ujn.edu.cn

2021 IEEE 4th International Conference on Electronics Technology (ICET) | 978-1-7281-7673-4/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICET51757.2021.9451034

***Abstract*—This paper designs an agricultural pest detection system based on machine learning. The system consists of pest detection algorithm and PC terminal system. The algorithm uses resnet50 as the backbone network, uses Feature Pyramid Network (FPN) to extract features, and optimizes them by Stochastic Gradient Descent (SGD) and Non-Maximum Suppression (NMS). Finally, the method is implemented by HALCON machine vision software. The PC side uses C # as the development language and C / S three-tier architecture for development, which is realized by visual studio 2015 combined with MySQL database. The system can detect and count the uploaded pest images, and save the detection results to MySQL database. The system constructed 27 agricultural common pest detection data sets, with an average precision of 92.5%. The experimental results show that the system can be effectively applied to the actual detection.**

Keywords—agricultural pest detection, machine learning, convolutional neural network, feature pyramid network, HALCON

I. Introduction

As a large agricultural country with a population of 1.4 billion, the yield of crops is particularly important. Crops are often affected by diseases and pests in the growth process. If the control is not timely or improper, the threat of plant diseases and insect pests will be more serious. This will also affect the yield of crops and cause huge losses to agricultural income. Under the traditional methods of pest detection, people mainly rely on their experiences. There are many kinds of agricultural pests with similar morphology and morphology, which are difficult to distinguish. It is time-consuming and inaccurate to judge only by human eye observation. Non-professional agricultural workers will be affected by their own experience, unable to accurately identify the pest species, thus affecting the late pest control work. Therefore, pest control is an important part of agricultural production, detection of agricultural pests is of great significance to promote the development of agriculture in China.

In recent years, with the development of artificial intelligence technology, image recognition technology has made a great leap. The use of advanced image recognition technology can quickly and accurately determine the species and number of pests, eliminate the subjective factors of manual detection, and reduce the detection error. However, most existing recognition methods use network images as training data sets. Although it has a good recognition rate, the images collected on the Internet can only detect one side of the pest, so there is a big gap in practical application.

In this paper, an agricultural pest detection system based on machine learning is proposed. The pest detection algorithm is written by HALCON machine vision software [1, 2], and the system is established by C# [3] and MySQL [4] database. The system can detect several forms of 27 common agricultural pests and count their numbers. The detection speed is fast and the accuracy is high, which reduces the time of artificial identification of pests.

II. Algorithm Design

Based on HDevelop, an interactive programming platform for HALCON machine vision software developed by MVtec in Germany, a pest detection algorithm was written. Since the 2017 version, HALCON has added support for deep learning [5], mainly using CNN algorithm to achieve deep learning, and provides a variety of pre-training CNN classifiers. Using HALCON for software development can shorten the development cycle and save the development cost.

A. Pretreatment

The establishment of pest image database is the basis of deep learning. Due to the lack of public pest data set, the sample images used in this paper are the images of insects killed by high temperature in the actual scene, and the collection location is Shandong. The image is real and effective, and has high application values. Deep learning needs a large number of sample data for training. Firstly, the collected image is filtered, and then a targeted program is written to segment and expand the collected samples.

Because the original image is large and contains a large number of insects, it is difficult to identify them, so it is necessary to segment them. This paper presents a practical image segmentation method, which can automatically obtain the image target area, and remove the impurities that affect the image recognition, so as to prepare for the later recognition.

Firstly, the image is grayed, and the image is scaled according to the maximum gray value to enhance the image

Shandong Province Agricultural Major Applied Technology Innovation Project: SD2019ZZ007.

contrast and make the background became more conducive to remove. The gray scale calculation formula is as follows:

$$g' = g * Mult + Add \tag{1}$$

$$Mult = \frac{255}{GMax - GMin} \tag{2}$$

$$Add = -Mult * GMin \tag{3}$$

g is the current gray value, $Mult$ is the multiplied coefficient, Add is the added offset, $GMax$ is the maximum gray value, $GMin$ is the minimum gray value.

In HALCON, there are three kinds of image segmentation methods: global threshold segmentation, automatic global threshold segmentation and local threshold segmentation. Here, the background is simply separated by threshold method, and then the image is segmented by the combination of open and close operations of morphology. Then find the target insect and cut a certain range around the target. Using texture propagation to remove the impurities in the clipped image [6], the specific process is shown in Fig. 1.

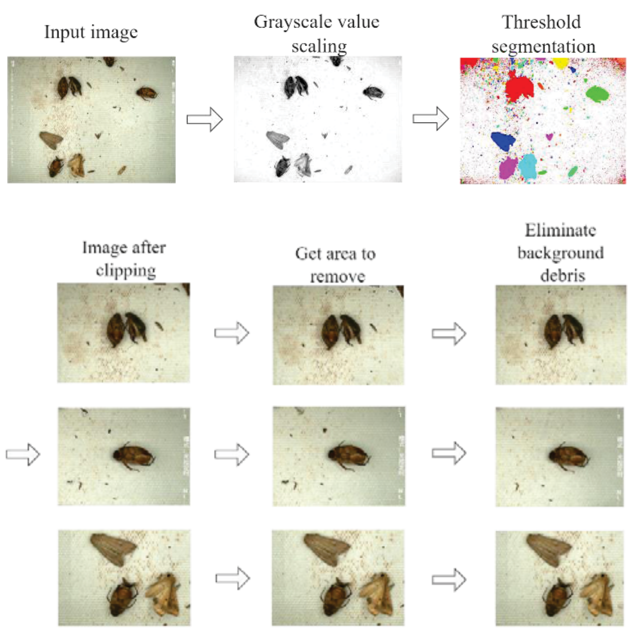

Fig. 1. Data set preprocessing.

Use the HALCON special data set markup software Deep Learning tool to mark the enhanced image [7], and the exported file after the mark can be directly imported into HLCON training and recognition to prepare for later training. This paper collected a total of 6064 basic images, including 10803 samples in 27 categories [8]. The specific categories and data distribution are shown in Table I.

TABLE I. TYPES AND QUANTITY OF PEST PICTURES

Pest number	Pest name	Number of samples
1	Pyrausta nubilalis	542
2	Helicoverpa armigera	1560
3	Anomala corpulenta Motschulsky	1417
4	Macdunnoughia crassisigna	144
5	Proxenus lepigone	926
6	Mamestra brassicae	723
7	Proceras venosatus	264
8	Hermonassa cecilia	70
9	Mythimna separta	619
10	Latoia sinica Moore	189
11	Spodoptera exigua	355
12	Agrotis segetum	182
13	Coccinellidae	374
14	Amata emma	494
15	Gryllotalpa	167
16	Semiothisa cinerearia	221
17	Eupolyphaga sinensis	256
18	Callambulyx tatarinovi	340
19	Scarites	228
20	Cricket	179
21	Diaphania quadrimaculalis	238
22	Agrius convolvuli	200
23	Smerinthus planus	200
24	Parum colligata	204
25	Marumba gaschkewitschi	190
26	Serraca punctinalis conferenda	147
27	Holotrichia parallela Motschulsky	374

B. Preprocessing and Setting Training Parameters

Preprocessing is an important part of target detection. The preprocessing part is mainly divided into six aspects: setting parameters, setting input and output paths, reading and splitting data sets, determining model parameters according to data, creating target detection model and data set preprocessing.

Set super parameter batch size, learning rate, momentum, Epoch, etc. Appropriate batch size can improve the utilization of memory, improve the training speed, and make the direction of gradient descent more accurate. The learning rate determines how to weight the gradient during training, whether the objective function converges to the local minimum value and when it can converge to the local minimum value. Different pre-training networks and pre training data need to adjust the learning rate. If the learning rate is too small, the convergence process will become very slow. If the learning rate is too large, the gradient may oscillate back and forth near the minimum value and cannot converge. Here we set the learning rate to decrease by 10 times at the specified epoch. Momentum can be used to modify the retrieval direction and speed up the convergence. Epoch determines the number of times the algorithm loops on the training set.

C. Model Establishment and Training

There are two main tasks in object detection, one is to find the location of the instance, the other is to judge its category. For pest image detection, convolution neural network algorithm is mainly used, and the training network is composed of three main parts. The first part is called backbone network, which is composed of pre-trained classifiers. The task is to generate various feature predictions and remove the classification layer. The deep learning version of Halcon provides four pre-trained neural networks for target detection:

Alexnet [9], compact, enhanced and Resnet50 [10]. These networks have been pre-trained on large image data sets. Alexnet is designed for simple classification tasks. The convolution kernel of its first convolution layer is larger, which is more conducive to feature extraction. The structure of Compact network is relatively simple, and the recognition speed is the fastest, which can be used for ordinary recognition training. Enhanced is an enhanced network. The neural network has more hidden layers than the pre-training classifier and has the best recognition rate. The structure of Resnet50 network is complex, which is suitable for dealing with complex recognition training, and the training is more stable and robust.

The second part is to combine different backbone networks, also known as the feature pyramid, which together with the first part forms the FPN [11]. FPN uses the traditional convolution network to extract features, and the spatial resolution gradually decreases with the deepening of convolution. Then the high-level features are sampled twice, and then combined with the feature pixels of the previous layer, the final feature map is obtained by iteration.

The third part is the additional network for each selected level, called heads. They get the corresponding feature map as input and learn how to locate and classify potential targets respectively. At the same time, the third part can reduce the bounding box produced by overlapping prediction. An overview of these three parts is shown in Fig. 2.

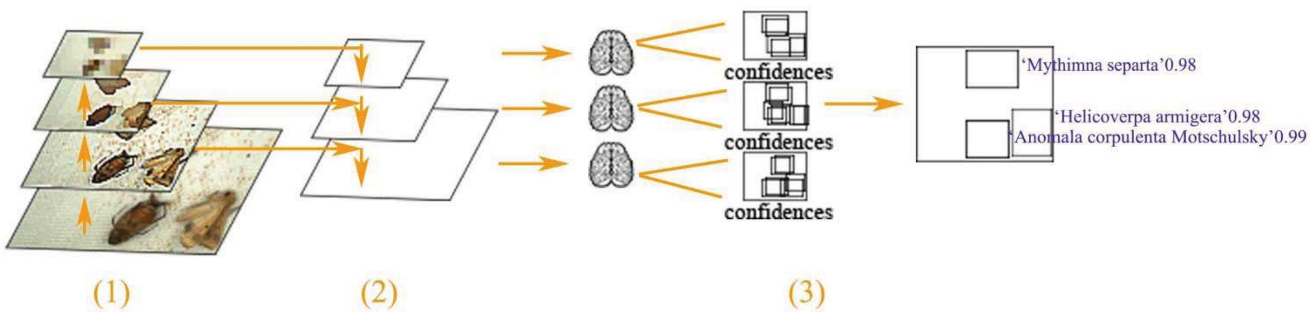

Fig. 2. Object detection algorithm: (1) The backbone. (2) Backbone feature maps are combined and new feature maps generated. (3) Additional networks.

In this paper, Resnet50 is selected as the pre-training network, which is used for transfer learning. A trained network is used to retrain it as a specified task, which can make up for the situation of less data. The network model convolutes images through convolution layer and uses ReLU function as activation function. ReLU function increases the nonlinear relationship between neural networks and accelerates the training speed. The feature dimension reduction of the image is realized by down-sampling in pooling layer, which can prevent overfitting in training process. The fully connected layer mainly maps the feature representation learned by the sample to the output space, which acts as a classifier, and finally the output layer outputs the result.

$$ReLU(x) = Max(0, x) = \begin{cases} x & x > 0 \\ 0 & x < 0 \end{cases} \tag{4}$$

The most widely used optimization algorithm in machine learning is gradient descent algorithm. In the training process, a nonlinear optimization algorithm is adopted to minimize the loss function. The algorithm for optimization is SGD, which uses only one sample to iterate each time, so the training speed is very fast, and it is suitable for the case of large sample size. It updates the layer weights of the previous iteration t and w^t to the new value w^{t+1} at iteration $t+1$, as follows:

$$v^{t+1} = \mu v^t - \lambda \nabla_\omega L \tag{5}$$

$$w^{t+1} = w^t + v^{t+1} \tag{6}$$

Here λ is the learning rate, μ is the momentum, L is the total loss, $\nabla_\omega L$ is the gradient adjustment of the total loss relative to the weight. Too large learning rate will lead to divergence of the algorithm, and a small learning rate will lead to unnecessary multi-step operation. As a result, it usually starts with a higher learning rate and reduces it during training.

In the prediction of bounding box, we may find many bounding boxes with similar positions but different sizes and directions. Therefore, NMS is added to the algorithm [12]. By setting the "max_overlap" parameter of the model to reduce the overlapping bounding boxes with the same category, and setting the "max_overlap_class_agnostic" parameter to reduce the overlapping bounding boxes with different categories. The set value determines the maximum number of overlapping points of the two bounding boxes. When the score is exceeded, it will be regarded as a different instance. Overlap calculation is calculated by the smallest closed rectangle of the instance.

D. Evaluation Model

For agricultural pest detection, the performance of the network needs to be evaluated. In this paper, Precision, Mean Average Precision (mAP) and Average Precision (AP) are used to evaluate the network after training [13].

Precision indicates how many of the predicted positive samples are true positive samples. This can be divided into two types: one is determined to be a positive sample, which is in fact a positive sample (TP), and the other is a positive sample determined to be a negative sample (FP).

$$P = \frac{TP}{TP+FP} \tag{7}$$

AP namely average precision, is the average value of the maximum precision under different recall values, which represents whether the prediction object is correctly detected. mAP is the average value of AP. The calculation formula of AP is as follows:

$$mAP = \frac{1}{|Q_R|}\sum_{q\in Q_R} AP(q) \tag{8}$$

For a Class A, first, all the output prediction boxes of category A are sorted according to the confidence level. Select the K prediction boxes with the highest confidence and calculate their TP and FP so that the recall rate is equal to 1. Then calculate the precision rate and repeat the previous steps to make the recall rates equal to 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0. Finally, the 11 precision rates obtained are averaged to obtain AP, and AP of each class is averaged to get mAP.

III. SYSTEM DESIGN

The system uses C# language as the development language, based on the development environment of Microsoft Visual Studio 2015, uses MySQL database, and adopts C/S three-tier architecture to develop and implement the agricultural pest detection system [14].

C/S structure is the structure of client and server [15]. Choosing C/S architecture can make full use of the advantages of hardware environment, reduce system overhead and allocate tasks reasonably.

The three-tier architecture includes presentation layer (UI), business logic layer (BLL) and data access layer (DAL) [16]. UI is the user interface layer that users can see intuitively. BLL makes logical judgment and analysis of the transmitted data, and transmits the correct value. DAL is mainly used to store the access to data classes, and it is the basic operation of adding, deleting, modifying and updating the database. The system architecture is shown in Fig. 3.

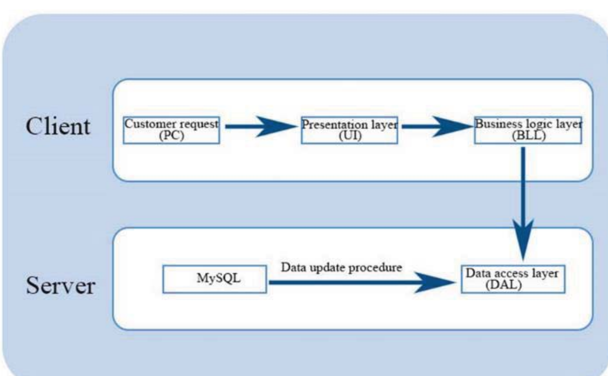

Fig. 3. C/S three-tier system architecture.

IV. EXPERIMENTAL RESULTS

In this experiment, 70% of the data set is set as the training set, 15% as the verification set and 15% as the test set. 100 epoch iterative training is carried out on the samples. In the preprocessing, the samples in the dataset are expanded, and the total number of samples is increased to 100% by randomly selecting images for mirror flipping. The other parameters are set as shown in Table II. Using Resnet50 as the pre-training network, calculate and record the average precision and average recall of each sample, and plot the recorded data into a table.

TABLE II. SUPER PARAMETER SETTING TABLE

Parameter name	value
epoch	100
learning_rate change epoch	70
learning_rate	0.0005
momentum	0.9
Weight_prior	0.00001
max_overlap	0.2
max_overlap_class_agnostic	0.8
min_confidence	0.8

In order to make the experiment more convincing, we compare the four pre-training models in Halcon: Alexnet, Compact, Enhanced and Resnet50, and test them in the test set. It can be seen from the Table III that the accuracy of various types obtained from the experiment is shown in the table. After 100 epoch, the average precision is 92.5%, and the average recall is 88.5%.

TABLE III. THE AVERAGE PRECISION AND RECALL OF THE FOUR MODELS

Model	Precision	Recall
pretrained_dl_classifier_alexnet.hdl	91.4%	79.8%
pretrained_dl_classifier_compact.hdl	92.0%	81.9%
pretrained_dl_classifier_enhanced.hdl	91.5%	86.0%
pretrained_dl_classifier_resnet50.hdl	92.5%	88.5%

As can be seen from Table III that the mAP of the model using Resnet50 as the pre-training network is 0.5%-1.1% higher than other pre-training networks, and the recall is increased by 2.5%-8.7%. It can be seen that using Resnet50 as a pre-training network can obtain the best recognition precision and the best recall rate. The experimental results can be used in practical applications. Finally, this article tests the model and the test result is as follows. As shown in Fig. 4:

Fig. 4. Model test results.

The system consists of three parts: real-time detection, image detection and data statistics. The system can read the pest images collected on the day, count the pest species and the number of pests in the images, and then store the statistical data into the database. The system test results are shown in Fig. 5-6.

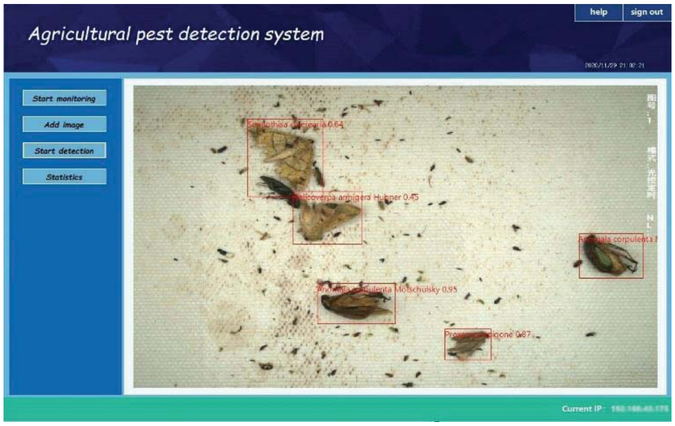

Fig. 5. Test results of pest detection function of the system.

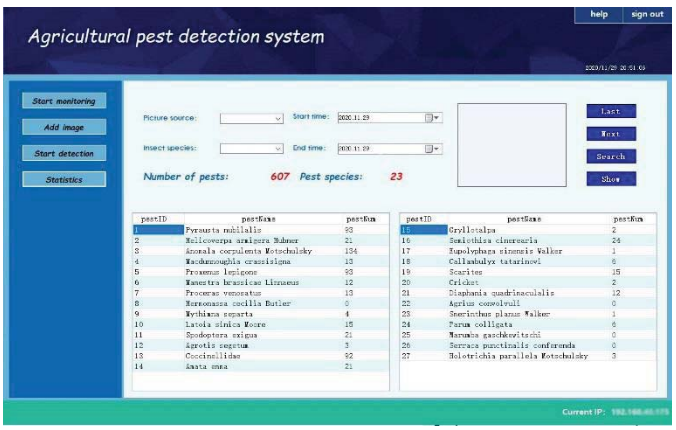

Fig. 6. Results of functional test on the number of 100 pest pictures.

V. CONCLUSION

In this paper, an agricultural pest detection system based on machine learning is proposed, which combines feature pyramid network and convolution neural network, and uses C/S architecture to build the system. The system can identify 27 kinds of common agricultural pests with an average precision of 92.5%. Compared with the traditional recognition method based on sample features, the robustness and accuracy of the method are improved. This method can help agricultural workers quickly detect agricultural pests, take corresponding control measures in advance, and improve crop yield.

ACKNOWLEDGMENT

This work is supported by the Shandong Province Agricultural Major Applied Technology Innovation Project (SD2019ZZ007).

REFERENCES

[1] Steger C , Ulrich M , Wiedemann C . Machine Vision Algorithms and Applications[M]. 2007.

[2] Xiang Huiyu, Xue Zhen, Leng Chongjie, et al. Experimental study on visual inspection of apple quality based on Halcon[J]. Food and machinery, 2016, 032 (010): 123-126.

[3] Barnett M , Rustan K , Leino M , et al. The Spec# Programming System: An Overview[C]// Proceedings of the 2004 international conference on Construction and Analysis of Safe, Secure, and Interoperable Smart Devices. Springer, Berlin, Heidelberg, 2005.

[4] Scott W. Ambler, Pramod J Sadalage. Refactoring Databases: Evolutionary Database Design[J]. 2006.

[5] Lin yunsen, fan Wenqiang, Jiang Jialiang. Research on fruit recognition technology based on deep learning[J]. Application of photoelectric technology. 2019 (06): 45-48.

[6] Liu, Li, Fieguth, et al. Texture Classification from Random Features[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012.

[7] Marchisotti D , Sala V . Evaluation of the bounding box uncertainty of deep-learning object detection in HALCON software[C]// 2020 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0&IoT). IEEE, 2020.

[8] Ding Jianyun, Zhang Jianhua. Atlas of moths under the lamp in Beijing [M]. China Agricultural Press, 2016.

[9] Krizhevsky A , Sutskever I , Hinton G . ImageNet Classification with Deep Convolutional Neural Networks[C]// NIPS. Curran Associates Inc. 2012.

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

[11] Lin T Y , Dollár, Piotr, Girshick R , et al. Feature Pyramid Networks for Object Detection[J]. 2016.

[12] Berg A C , Malik J . Shape Matching and Object Recognition[M]// Toward Category-Level Object Recognition. IEEE Computer Society, 2007.

[13] Christine, Meschede, Daniel. Anaphora Resolution: Analysing the Impact on Mean Average Precision and Detecting Limitations of Automated Approaches[J]. International Journal of Information Retrieval Research, 2018.

[14] Zhou Xiaobo. Research and implementation of intelligent identification system for tea garden pests based on machine vision [J]. Modern agricultural science and Technology,2019(18):229-230.

[15] Chen Meixiang, Guo Jiying, Xu Jianping, et al. Development of automatic monitoring, identification and counting system for Grapholitha molesta [J]. Acta Environmental Entomology, 2018,40 (5): 1164-1174.

[16] Jing D, Dechun D . Optimized WebGIS application and research based on three-tier under. Net Environment and distributed Browser/Server architecture[C]// International Conference on Information Management. IEEE, 2014.