A Review of Object Detection Techniques Applied to Pest Images

Hongshuo Zhai [1*], Fengzhi Dai [1], Lijiang Zhang [2], Qiang Wang [3]

[1] Tianjin University of Science and Technology, Tianjin,300222, China

[2] Xinjiang Shenhua Biotechnology Co., Ltd, Xinjiang, China;

[3] Easy Control Intelligent Technology (Tianjin) Co., Ltd., Tianjin, China

E-mail: [*] 1936130557@qq.com

Abstract

In agricultural information management, crop pest control has always been an important topic, and the image detection technology of small target pests is particularly critical in this process. At present, the technology faces challenges such as difficult data collection and insufficient robustness. This paper first introduces the development of object detection technology and its application in the field of agriculture, then analyzes the challenges of information-based pest control, discusses the research progress of pest dataset construction, image data augmentation technology and object detection algorithm, and finally points out the future research direction in this field.

Keywords: image detection technology, data collection, insufficient robustness, agriculture

1. Introduction

The application of target detection in the field of pest detection is very important. Since the acquisition of pest images can be affected by environmental conditions such as weather, light and occlusion, this increases the difficulty of detection. In addition, the diversity of surface textures and shapes of pests also poses challenges for detection. Even objects of the same class can vary from Angle to Angle and from individual to individual, adding to the complexity of detection1.

In the traditional methods, sliding window based search and artificial feature extraction face the challenge of efficiency and accuracy in practical applications. For example, the sliding window method requires traversing every possible position of the image, which is time-consuming and prone to creating a large number of redundant Windows. To overcome these problems, deep learning techniques have been introduced into the field of pest detection. Deep convolutional neural networks (CNNS) reduce manual intervention and improve the robustness of models by automatically learning features from large amounts of data. Deep learning models such as YOLO and SSD are trained end-to-end to make predictions directly from image pixels to target categories and bounding boxes, simplifying the detection process and increasing efficiency.

This paper mainly introduces the development of target detection in the field of agricultural pests, analyzes the existing problems and solutions, and finally introduces the improvement and lightweight of the model.

2. The Phase Detection Algorithm and Its Application in Agriculture

2.1. Two-stage deep learning algorithm

Two-stage deep learning algorithm plays an important role in the field of object detection, and its workflow usually includes two main steps: First, the candidate region is generated through the region proposal network (RPN); Then, the convolutional neural network (CNN) is used to classify these candidate regions and perform boundary box regression. Representative algorithms of this method include R-CNN, Fast R-CNN, and Faster R-CNN. Fig.1 represents the framework of the two-stage algorithm.

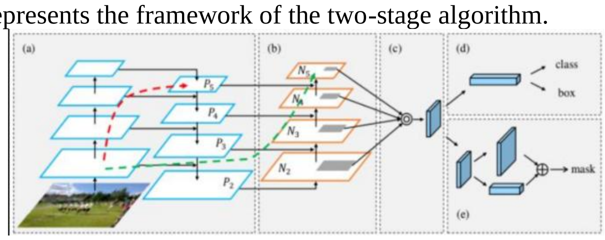

Fig.1 Two-stage algorithm

R-CNN is one of the earliest two-stage object detection algorithms. It identifies multiple candidate regions that may contain target objects in the original image by selective search algorithms, then scales these regions to a fixed size, and extracts features using a pre-trained CNN model. Finally, support vector machine (SVM) is used to classify the extracted features, and the boundary box is adjusted by regression model. Although R-CNN has made significant progress in accuracy, it is less computationally efficient because each candidate region needs to be processed independently[2].

Fast R-CNN significantly improves computational efficiency by feeding the entire image into the CNN for one-time feature extraction. It introduces an RoI Pooling layer to unify the feature sizes proposed by different regions, thereby reducing double counting. However, Fast R-CNN still relies on selective search to generate candidate regions, resulting in higher computational costs. Fig.2 represents the fast R-CNN algorithm.

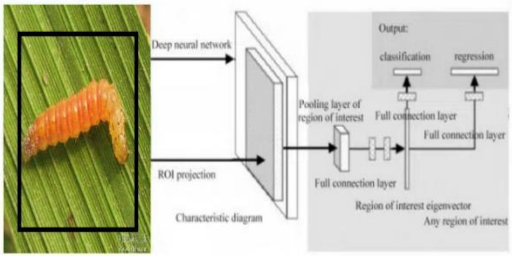

Fig.2 Fast R-CNN algorithm diagram

2.2. *Single-stage deep learning algorithm*

The single-stage deep learning algorithm uses an end-to-end approach to identify the category and location of the object using only one network on a detected image. Compared with the two-stage method, the single-stage method simplifies the detection process, greatly improves the efficiency, and has better real-time performance. A typical single-stage algorithm is YOLO[3].Fig.3 represents the framework of the single-stage algorithm.

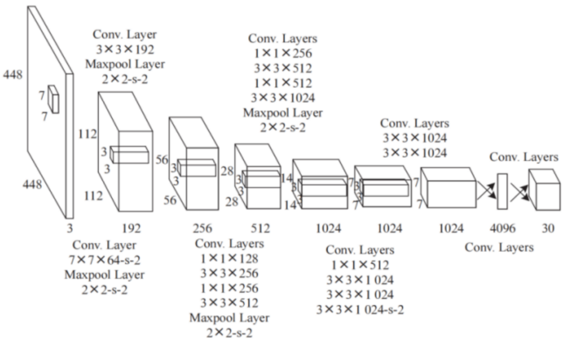

Fig.3 Single-stage algorithm

The YOLO series algorithm works by dividing the image into multiple grids, each of which predicts the border and category information of the target. The YOLO series of algorithms works by dividing the image into multiple grids, each of which predicts the boundary and category information of the target. In pest detection, YOLO algorithm is widely used in real-time monitoring and rapid response. For example, the lightweight YOLO-MobileNet-CBAM model performed well in the detection of small crop pests and diseases, with an accuracy of 92.38%.Fig.4 shows the verification results of IP102 data set after the training of YOLOv10 algorithm.

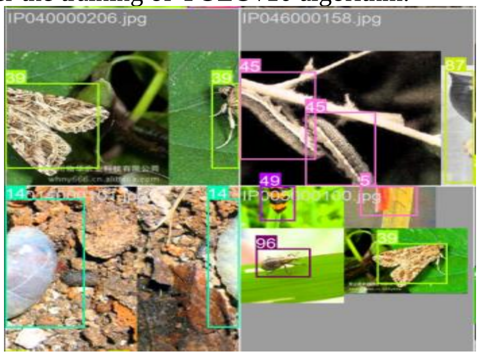

Fig.4 Verification results

3. The Main Existing Problems and Solutions

3.1. *Major problem*

In the process of applying target detection technology to automatic pest identification, although it has improved significantly compared with traditional methods, it still faces multiple challenges in practical application.

Insufficient training samples: Deep learning models, especially CNNS, require large amounts of training data to learn complex features. In a field environment, collecting enough tagged pest images is difficult because shooting in field conditions can be affected by a variety of factors, such as leaf occlusion, lighting changes, and so on. In addition, the diversity and number of insects is limited, resulting in a scarcity of samples available for training.

Pest targets are small in size: Images of pests taken in wild environments tend to be small, perhaps only a few pixels, which makes them difficult to identify in images. With the increase of the number of CNN layers, the feature information of these small targets may be lost, which increases the difficulty of detection.

High similarity between pests: Different pests can be very similar to each other in appearance, which makes it difficult for models to distinguish between them. In addition, pests of the same species may also be difficult to identify due to individual differences, attitudinal changes, or partial occlusion. Fig.5 shows the image of a small sample of insects in the actual agricultural scene.

Fig.5 Small pest targets

3.2. Solution

For the above problems, we usually have some to solve.

Data enhancement and synthesis: Generate composite images through data enhancement techniques (such as rotation, scaling, color adjustment), increase the diversity and number of data sets, and improve the generalization ability of models. Fig.6 shows the image of the original data after partial data enhancement.

Fig.6 Image after data enhancement

Multi-scale feature fusion: The feature pyramid network (FPN) and other technologies are used to fuse features of different scales to enhance the detection ability of the model for small targets.

Deep learning model optimization: single-stage detection algorithms such as YOLO and SSD are adopted to simplify the detection process and improve the detection speed. At the same time, the algorithm structure is iteratively optimized, such as the YOLOv10 model, and the accuracy and efficiency of detection are improved by improving the network structure and loss function.

4. Datasets and Data Enhancement

4.1. Datasets

In the field of deep learning-driven object detection, the quality and diversity of data sets have a decisive impact on model performance. The construction of pest image data set is the first step to achieve efficient target detection. Since the acquisition of pest images is limited by environmental factors and acquisition techniques, it is essential to build a comprehensive, diverse and accurately labeled data set. In recent years, researchers have documented a large number of insect photos, but there are still insufficient samples. For example, the ip102 dataset contains 102 insect species, but some samples have far fewer images than others. Fig.7 visualizes the number of categories in the IP102 dataset.

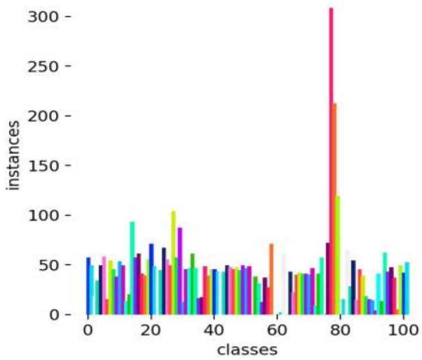

Fig.7 IP102 data set category number visualization.

4.2. Data enhancement

In view of the problem of insufficient samples in pest image datasets, data enhancement technology has become the key to improve model generalization and detection performance. Data enhancement manually extends the training set by applying a series of transform operations, including but not limited to rotation, scaling, color adjustment, blurring, and noise addition. These operations can not only increase the diversity of samples, but also improve the adaptability of the model to pest images under different environmental conditions.

In addition to traditional data enhancement methods, the researchers also explored more advanced techniques such as:

Mixup: By mixing two images and their labels in a certain proportion, a new training sample is generated to enhance the robustness of the model to input perturbations.

CutMix: Similar to Mixup, but cuts and pastes local areas of the image while adjusting the mix ratio of the label. Fig.8 shows the CutMix mixing operation.

Fig.8 Demonstration of CutMix

5. Detection Methods

After the creation of specific pest data sets and the application of data enhancement techniques to form large-scale data sets, the problem of high similarity of pest images remains a challenge in detection. Therefore, enhancing the performance of the algorithm and improving the accuracy of the model is the key to solve the problem of automatic pest image recognition. Although the object detection technology of deep learning has made remarkable application and progress, there are still many challenges in small object detection. The main problem is that the features of small-size targets are gradually lost after passing through the pooling layer of CNN, which may lead to missing detection of some targets in multi-layer networks, resulting in poor effect of existing target detection methods in processing small-size pest images.

5.1. Improvement based on network structure

On the one hand, the detection performance can be improved by increasing the depth of the network. The deep network can learn more complex feature representations, helping to capture the subtle features of small targets, but attention should be paid to solving problems such as gradient disappearance or explosion that may occur. Hu et al. studied the image classification task, based on the traditional convolutional neural network, and improved the performance by increasing the network depth. For example, the depth of the original network is increased appropriately and the residual connection technique is used to detect the tiny deformed digits in handwritten digital images. The results show that the classification accuracy of these small target numbers with subtle differences is improved by about 5%, which proves that increasing network depth combined with appropriate technology can improve the detection effect of small target.

On the other hand, changing the network architecture is also an effective approach, such as using the feature pyramid network (FPN) structure, which can comprehensively utilize different levels of feature information, and fuse shallow high-resolution feature maps with deep semantic feature maps, so that the model can detect small targets with both the required resolution for positioning and semantic information for classification. Lin et al. proposed FPN architecture for target detection. When detecting small objects in the COCO dataset, the Faster R-CNN model based on FPN architecture is used. This model integrates the feature maps of different levels, which improves the positioning accuracy of small objects by about 8% compared with the original model without FPN, fully demonstrating the advantages of FPN in the improvement of network architecture in the detection of small objects.

5.2. Improvements based on feature enhancement

The first is to carry out multi-scale feature fusion. Since small targets have different performance at different scales, it is difficult to describe a single scale feature comprehensively. By fusing feature maps output by different convolutional layers, which contain multiple features ranging from details to semantics, the model can obtain more comprehensive target information.

The second is the application of attention mechanism, such as spatial attention mechanism can guide the model to focus on the region where the small target is located, and channel attention mechanism can adapt to adjust the importance of channel features, and improve the detection accuracy by highlighting the features of the small target. Hu et al. proposed SENet to apply SENet's channel attention mechanism to the target detection model when detecting small animals in ImageNet data sets. By increasing the channel weights that contain key features of small targets, the detection accuracy of small animals is improved from about 60% to about 68% without using this mechanism, highlighting the improvement effect of attention mechanism on small target detection.

5.3. Lightweight model

The application scope of automatic pest identification should not be limited to large-scale farms, small-scale family farms are also in urgent need of it. To this end, it is crucial to create a lightweight system that can directly photograph and recognize pest images, thereby helping farmers quickly and accurately identify pest species4.

The realization of system lightweight can be started from many aspects. On the one hand, by reducing the parameter scale of the model, pruning is carried out on those unimportant weights and convolutional kernels in the CNN model to remove redundant parts and make the model "slim". Fig.9 shows the schematic diagram of pruning weight reduction.

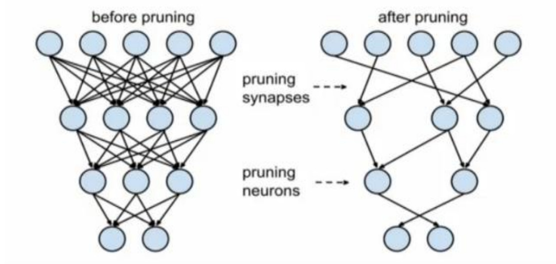

Fig.9 Pruning schematic diagram

Distillation technology can also be introduced to use the soft label information of the teacher model to guide the learning of the student model, so that the student model can maintain high performance while the number of parameters is greatly reduced and the model is further lightweight.

References

1. Popescu Dan, Dinca Alexandru, Ichim Loretta & Angelescu Nicoleta. (2023). *New trends in detection of harmful insects and pests in modern agriculture using artificial neural networks. a review* . Frontiers in Plant Science 1268167-1268167.
2. Kai Run Li, Li Jun Duan, Yang Jun Deng, et al.,(2024).*Pest Detection Based on Lightweight Locality-Aware Faster R-CNN*.Agronomy(10),2303-2303.
3. Ajantha Vijayakumar & Subramaniyaswamy Vairavasundaram. (2024).*YOLO-based Object Detection Models: A Review and its Applications.Multimedia Tools and Applications*(35),83535-83574.

4. Tingyao Jiang & Shuo Chen. (2024).*A Lightweight Forest Pest Image Recognition Model Based on Improved YOLOv8.*Applied Sciences(5),

Authors Introduction

Mr. Hongshuo Zhai

He is a second-year master candidate in Tianjin University of Science and Technology.

Dr. Fengzhi Dai

He received M.E. and Doctor of Engineering (PhD) from the Beijing Institute of Technology, China in 1998 and Oita University, Japan in 2004 respectively. His main research interests are artificial intelligence, pattern recognition and robotics. He worked in National Institute of Technology, Matsue College, Japan from 2003 to 2009. Since October 2009, he has been the staff in College of Electronic Information and Automation, Tianjin University of Science and Technology, China.

Mr. Lijiang Zhang

He has been chairman of Xinjiang Shenhua Biotechnology Co., Ltd, China and Zhongnong Guoyou Technology (Tianjin) Co., Ltd, China since 2019. His research interest is deep processing of agricultural products.

Mr. Qiang Wang

He was graduated in Tianjin University of Science and Technology in 1997. Since 2022, he has been the chairman of Easy Control Intelligent Technology (Tianjin) Co., Ltd., China, focusing on the development of KRMotion: A PC-Based Software Motion Control Platform.