Fabric Defect Detection Based on Improved YOLOv7 Network

Yuwen Shi

Wuhan Textile University, Wuhan 430200, China.

Abstract: In order to achieve real-time defect detection technology, detection accuracy, prediction speed, and lightweight deployment models are important issues. Traditional object detection methods often fail to achieve a balanced effect on all aspects. Therefore, a detection model based on lightweight convolutional neural network YOLOv7 is proposed. Firstly, lightweight convolutional Ghost conv is introduced to lighten the backbone network; Secondly, adding CBAM attention mechanism to suppress invalid information and enhance feature extraction ability; Finally, a new measurement method is introduced at the regression loss function α- Replacing IoU with SIoU accelerates algorithm convergence and improves detection efficiency for defect targets. The experiment shows that the accuracy P of the detection model reaches 96.27%, the mAP index is 83.84%, the detection speed is 23.83ms, and the model size is only 19.10MB, effectively balancing the accuracy, real-time performance, and lightweight deployment of defect detection.

Keywords: effect detection; YOLOv7; CBAM Attention Module; Ghost Convolution; -SIoU

Introduction

With the continuous development of textile industry, fabric products are more and more widely used in daily life. However, due to various factors in the textile process, such as yarn quality, production environment, weaving technology, etc., there are often various defects on fabric products, such as broken yarn, holes, stains and wrong flowers. These defects not only affect the appearance quality of the fabric, but also may affect its service life and functional performance. Therefore, the accurate detection and classification of fabric defects is very important to ensure product quality and improve production efficiency.

Traditional fabric defects detection methods usually rely on manual visual inspection, which is time-consuming, labor-intensive and easy to be affected by subjective factors, and can not meet the needs of large-scale production. In recent years, with the rapid development of deep learning technology, fabric defects detection methods based on deep learning have gradually become a research hotspot. Deep learning is a machine learning method based on artificial neural network, which has the advantages of strong learning ability, high feature expression ability and high degree of automation. Through deep learning, features can be learned and extracted from a large number of fabric image data, so as to realize automatic detection and classification of fabric defects.

Considering that fabric defect detection requires very high real-time and accuracy, some inherent deep learning networks may be inadequate in terms of practicality. To solve the above problems, this paper proposes a lightweight YOLO-v7 network structure. By improving the deep network, adding the attention mechanism, improving the loss function and other operations, the improved network can significantly improve the detection speed, reduce the number of parameters, and effectively improve the practicality and reliability of fabric defect detection while maintaining the detection accuracy as much as possible.

1. Target detection network YOLO-v7 improved

1.1 YOLO-V7 Network Structure

The YOLO-v7 object detection network architecture consists of three main parts. The Backbone module is the backbone feature extraction network module, which is responsible for extracting the three initial feature layers. The second part corresponds to the Neck module on the graph, which is used to strengthen feature selection and fusion, and integrates these three initial feature layers to extract more high-quality features. Finally, the prediction network module corresponds to the Head module on the graph, which is used to generate the final target detection prediction result.

1.2 Improvement of YOLO-V7 network structure

Due to the large data volume of YOLO-V7, model training becomes time-consuming, and the generated model requires a lot of storage space, which makes it not suitable for resource-constrained devices or scenarios with high real-time requirements. In addition, for some simple object detection tasks, the detection speed of the YOLO-V7 can introduce a large amount of data redundancy. Therefore, in order to balance detection speed and detection accuracy, an ultra-lightweight YOLO-V7 object detection network structure is proposed.

1.2.1 Improvement of backbone extraction network

Considering that the backbone feature extraction network of YOLO-V7, CSPDarknet53[1], has a large number of parameters, multiple convolution nuclei are superimposed on each other, resulting in high computational costs. In order to solve the problem of model bloat, some scholars have proposed some network optimization schemes, such as ShuffleNet[2] and MobileNet[3]. Although these schemes reduce the size of the model to some extent, there is still a lot of data redundancy in extracting features. Therefore, a good alternative is to use Ghost conv convolution . As shown in Figure 3, Ghost conv is a lightweight convolutional extraction network. Its core idea is to reduce the computational load by introducing Ghost feature mapping, and to decompose the traditional convolution operation into primary convolution and additional convolution. The main convolution uses fewer convolution kernels to generate the initial feature map, and then reduces the computational cost of non-critical features through simple linear transformations. Convolution uses additional convolution cores to process the remaining key features and operates in series with the result of the main convolution to generate the final feature map.

Ghost Conv converts traditional convolution operations into two steps. In the first step, a small amount of the original feature map is generated.

In the second step, the final feature map is generated by a linear low-cost transformation. The experimental results show that this algorithm has better performance than some networks that reduce the number of layers for lightweight, and can be used to solve the lightweight modeling problem. In this way, Ghost conv convolution can significantly reduce the number of parameters and computations while maintaining model performance.

1.2.2 Improvement of backbone extraction network

The lightweight design of YOLO-v7 will lead to some loss of accuracy, especially in the detection of small targets and complex scenes, so it is necessary to add some network structures to make up for the loss caused by lightweight. Therefore, this paper adds CBAM attention mechanism between the feature extraction network and the feature fusion network. CBAM innovatively integrates channel attention and spatial attention, enabling the network to focus on important features and suppress unnecessary ones. Compared with other modules, CBAM achieves the best efficiency in both channel dimension and spatial dimension.

1.2.3 Improved loss function (-SIOU)

In this paper, -SIOU loss function is proposed as an alternative to boundary frame loss function in YOLOv7 model. -SIOU loss function is improved on the basis of SIOU[10] loss function. Parameters are introduced to improve the gradient convergence rate of loss function, and the influence of distance, Angle and area on model boundary regression is fully considered. The original SIOU loss function is composed of Angle loss (Langle), distance loss (Ldis), shape loss (Lshape) and cross loss (IOU) to consider multiple factors.

2. Model training and analysis

2.1 Data Set

In this paper, the data set used for fabric defect detection in a textile factory in Jiangsu province is adopted, which contains a large number of fabric defect image samples, including stains, holes, broken yarn, etc. The data set is labeled, including five characteristics, such as holes, broken warp and weft, stains, wrong colors, and double warp and weft. In order to ensure the consistency of training and verification results, the data set is processed, including cropping, scaling, flipping and so on. After processing, the new dataset included 6658 images of fabric defects. Finally, the experimental data set is divided into verification set, test set and training set in the ratio of 1:1:8.

2.2 Test environment and parameter setting

The research test environment is 64-bit Win10 operating system, NVIDIA GeForce RTX3080 graphics card, CPU is Inter (R) Core (TM) i7-7700. The programming language is Python and the deep learning framework is Pytorch. The initial learning rate is set to 0.0002, the minimum learning rate is set to 0.00002, the momentum parameter is set to 0.942, the batch size is 16, and 160 epochs are trained.

2.3 Evaluation index

In order to evaluate the accuracy and practicability of the improved YOLO-v7 model, a number of comprehensive evaluation indicators were adopted in this study. The parameters include P (Precision), R (Recall), Average Precision mAP (Average Precision), parameter number (Params), and calculation Time (Time). These evaluation indexes are used to measure the effect of defects detection.

3. Results and analysis

3.1 Model comparison test

In order to further evaluate the performance of the improved algorithm in object detection, The mainstream target detection models EfficientNet, Faster RCNN, YOLOv5, YOLOv7, YOLOX, and SSD were trained to perform comparative experiments on fabric defects in the test set. Seven algorithms were used to carry out partial detection results of defects recognition in the same environment. It can be seen that the typical network EfficientNet, which is lightweight, has inaccurate positioning of small targets, and it also fails to detect cases where the ratio of wrong color to texture of the fabric is not large. The detection accuracy of SSD network is the lowest, and due to the influence of prior frame, the positioning accuracy of defects is not high, resulting in a high error rate. In addition, SSD networks mainly generate candidate frames through intensive sampling, which will produce a large amount of data redundancy when processing multi-target detection, including many overlapping or redundant frames, delaying the detection time. Also, the Faster RCNN network cannot fully adapt to defects of various scales due to the influence of fixed prior frames, thus reducing the detection effect of the model, and the defects of small targets are easy to be missed. YOLOV7 performs well in the detection of large-scale targets, and the algorithm performs well in the optimization of network structure, but the detection performance of small-scale targets is weak, and positioning errors are prone to occur in small defects problems. Under the premise of improved lightweight, the proposed algorithm maintains the detection accuracy of the original YOLOV7 network, improves the detection ability of small target defects, and also has a good recognition effect for defects with low contrast of fabric color and texture, which further proves the practicability of the method.

4. Knots

In order to further evaluate the performance of the improved algorithm in all aspects of target detection, the baseline YOLOv7 model is compared with the improved network. Experiments show that YOLOV7 performs well in the detection of large-scale targets, and the algorithm performs well in the optimization of network structure, but the detection performance of small-scale targets is weak, and positioning errors are prone to occur in small defects problems. Under the premise of improved lightweight, the proposed algorithm maintains the detection accuracy of the original YOLOV7 network, improves the detection ability of small target defects, and also has a good recognition effect for defects with low contrast of fabric color and texture, which further proves the practicability of the method.

References

[1] Zeng Huafu, Yang Jie, Li Linhong. Clothing image classification algorithm based on improved ShuffleNet v1 [J]. Modern Textile Technology, 2019,31(2):23-35. (in Chinese) DOI:10.19398 / j.AT.202208003.

[2] Zhang Chun-chun, Luo Ruilin, Lu Lin, et al. Image recognition of microstrip green tobacco leaf based on MobileNet and transfer learning [J]. Journal of Yunnan Normal University (Natural Science Edition), 2023,43(4):46-48. DOI:10.7699/j.nnu. ns-2023-052.

[3] Zhang Qingchun, Wu Zheng, Zhou Ling, et al. Vehicle and Pedestrian target recognition method based on improved YOLOv5 [J]. China Tests and Tests,2023,49(7):82-88. (in Chinese) DOI:10.11857/j.issn.1674-5124.2022060008.