Journal of Computer Technology and Software

ISSN:2998-2383

Vol. 4, No. 2, 2025

Enhanced Apple Quality Detection Algorithm Based on Improved YOLOv5 with Coordinate Attention Mechanism

Zylaren Cove[1], Lirien Voska[2]

[1]Department of Computer Science, University of Tennessee, Knoxville, USA

[2]Department of Computer Science, University of Tennessee, Knoxville, USA

*Corresponding Author: Zylaren Cove, zylaren58@utk.edu

Abstract: In recent years, the rapid advancement of artificial intelligence, particularly in the field of deep learning, has significantly improved the efficiency and accuracy of agricultural product quality inspection. This paper presents an improved apple quality detection algorithm based on the YOLOv5 model. The proposed approach incorporates a coordinate attention mechanism into YOLOv5, enhancing the model's feature extraction capabilities and improving its robustness across different scenarios. Experimental results demonstrate that the improved model achieves a recall rate of 88.8%, compared to the original 82.7%, and increases the mean average precision (mAP) from 83.7% to 85.9%. These enhancements not only improve the overall accuracy and recognition rate of apple quality detection but also contribute to reducing manual labor costs and enhancing automation in fruit quality inspection processes. This study provides an effective and efficient solution for intelligent apple quality assessment in modern agricultural applications.

Keywords: Yolov5; Coordinate Attention; mAP; Apple Quality Detection.

1. Introduction

In recent years, the development of science and technology makes artificial intelligence become popular, especially deep learning has been applied in various fields. Fruit quality detection in fruit processing plants, orchards and other places are useful, preliminary identification of the quality of fruit, and then a more detailed division of the pros and cons, can improve the efficiency of artificial detection, make more automatic intelligent. Hui Wang [1] has proposed a modified you only look once (YOLO) fruit recognition model, where group normalization (GN) replaces batch normalization, BN) method to optimize operational parameters. However, the feature fusion effect of feature pyramid networks (FP Net) in YOLOv3 is poor, and fruit features cannot be fully extracted, with an average recognition rate of only 85.91%. Bargoti [2] etc. a multi-scale perceptron and convolution is put forward neural network (convolutional neural network, CNN) fusion segmentation fruit checking calculation method, and then use the watershed segmentation and the Hough transform algorithm for segmentation image into line detection and counting, effectively improves the model segmentation and detection performance. However, the generalization of detection under different kinds of fruits and different scenes needs further study. Fenggang Sun [3] et al. detected and recognized apple fruit diseases based on the improved YOLOv5s model and combined with the transfer learning method, which improved the detection accuracy by 8.5% compared with the results of the original model and realized the rapid and accurate identification of apple diseases while occupying fewer computing resources. Hui Gao [4] et al. used automatic brightness correction technology and weighted vector machine to further improve the accuracy and speed of fruit defect detection. However, this method requires lighting system configuration and costs a lot.

This paper proposes an improved algorithm based on yolov5 model for apple quality inspection, which makes the trained model more generalization ability and robustness.

2. Basic Network Structure

2.1 Yolov5 Network Structure

Firstly, yolov5[5] network structure is composed of input, Backbone, Neck and Head. Yolov5s is the network with the smallest depth and smallest width of feature graph in the series. The Backbone consists of CBS, BotteneckCSP or C3, SPP or SPPF. The Neck part mainly consists of CBS, Up sample, Concat and CSP. The Head part is three Detect detectors, which are the process of target detection based on the anchor of the grid on the feature graph of different scales.

2.2 Introduction of Related Modules

CBS module is composed of Conv, BatchNorm and SiLu. BotteneckCSP module is composed of CBS, Resx and Concat. After the output of CBS, primary pooling, double pooling and tertiary pooling. SPPF module is first spliced, and then features are extracted by CBS. The Concat module is the function of concatenation. The concatenation object is passed in from. The concatenation dimension is determined by the args parameter.

3. Improved yolov5 Network Model Algorithm

3.1 Coordinate Attention Module

Coordinate Attention [6] it is simple, flexible and efficient under little additional computation, can be inserted into the classic lightweight network model, raise the accuracy. Relevant experiments show that Coordinate Attention not only has excellent performance in classification tasks, but also has significant effects on such intensive prediction tasks as target detection and instance segmentation.

It has some advantages. First, it captures not only cross-channel information, but also direction-aware and location-sensitive information, which will help network models locate and identify objects of interest more accurately. Secondly, the method is flexible and lightweight, which can be easily plugged into the classic modules of the mobile network. Furthermore, this attention can bring significant performance improvements to the downstream tasks of the mobile network.

3.2 Introduce the yolov5 Model of Coordinate Attention Module

Firstly, CA attention modules are added to the fourth and fifth layers of Backbone, and the sixth and seventh layers of backbone. Finally, CA attention modules are added to the previous line of SPPF module code. Finally, the number of layers was modified in the corresponding code in the Head section, and the final result was an improved program based on the yolov5 network model. The network structure diagram of the improved model is shown in Figure. 1.

4. Experiment and Analysis

4.1 Collection and Production of Data Sets

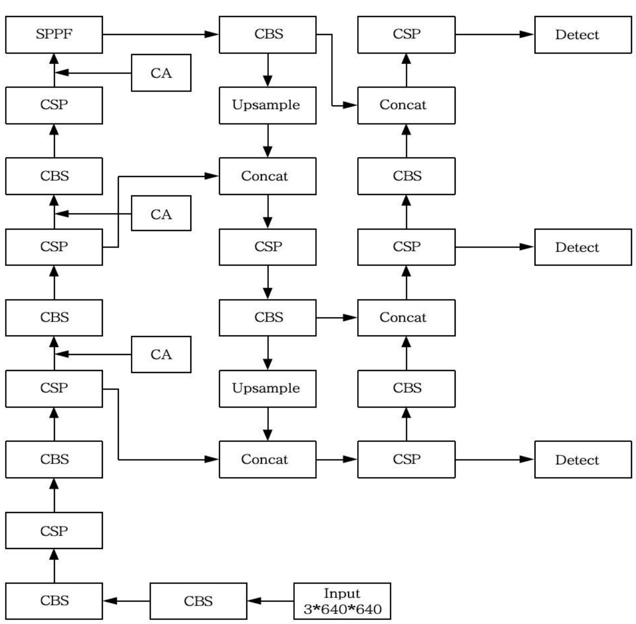

Fig 1. Improve the network structure of the model

This paper uses Fruits fresh and rotten for classification data set from Kaggle dataset website and some images from Fruits360 data set. At the same time, some data enhancement operations were carried out to improve the robustness of the experimental model. Marks were made for the sorted apple data set, and the ratio of training set and verification set was 7:3, with a total of 740 pieces.

4.2 Evaluation Index

The key to evaluate the performance of a model is the authenticity of the evaluation index. The higher the accuracy ratio P, the higher the proportion of correct result samples in the prediction results and the lower the false detection. The calculation formula is as follows.

$$P = \frac{TP}{TP+FP} \times 100\%$$

The higher the recall rate R is, the more positive samples are correctly detected in the prediction results, and the lower the missing samples are. The calculation formula is as follows.

$$R = \frac{TP}{TP+FN} \times 100\%$$

The higher the mean mAP of average accuracy is, the better the average detection effect of each category of the target detection model is. The calculation formula is as follows.

$$mAP = \frac{1}{n}\sum_{i=1}^{n} AP_i$$

4.3 Model Comparison Analysis

The model was trained for a total of 200 epochs. The accuracy, recall rate and mAP pairs of the original model and the improved model are shown in the following Fig.2 and Fig.3.

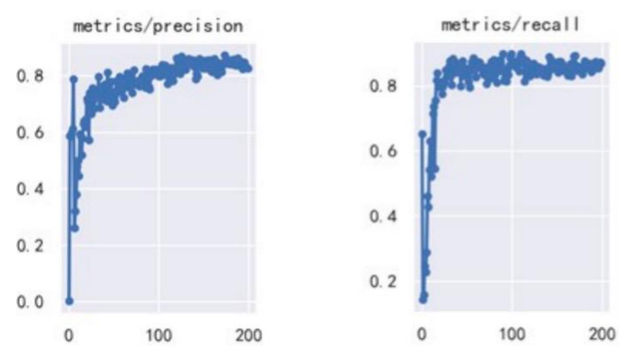

Fig 2. Original model

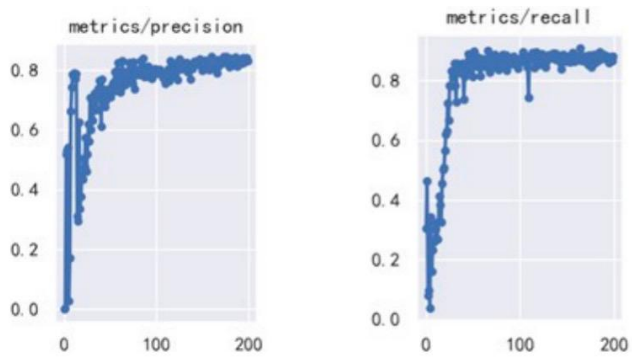

Fig 3. Improved model

It can be seen from the above figure that the improved model tends to be smoother. Next, compare the PR curves of the two models, as shown in the Fig.4 and Fig.5.

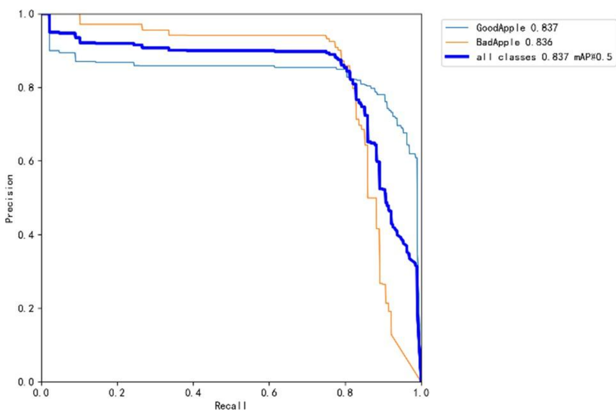

Fig 4. PR curve of the original model

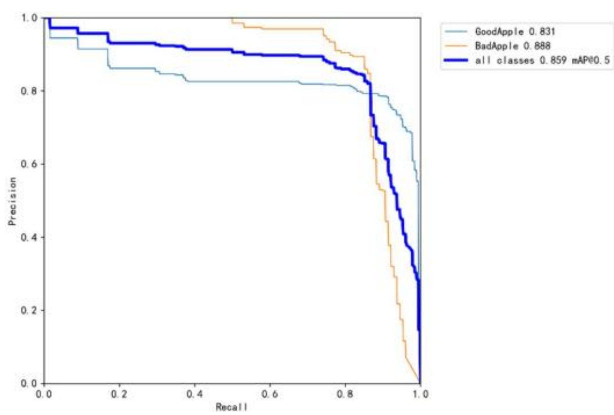

Fig 5. PR curve of the improved model

As can be seen from the above PR graph, the mAP of the original model reaches 83.7%, while that of the improved model reaches 85.9%. The following is an analysis of the data performance of the evaluation indicators in the two models, as shown in the Table 1.

Table 1. Comparison of model evaluation index data

Class	P	R	mAP
all (Original)	0.849	0.827	0.837
GoodApple (Original)	0.789	0.878	0.837
BadApple (Original)	0.908	0.775	0.836
all (Improved)	0.811	0.888	0.859
GoodApple (Improved)	0.73	0.947	0.831
BadApple (Improved)	0.891	0.829	0.888

As can be seen from the above table, the improved recall rate of the improved model indicates that the missed rate is low. The precision is a bit lower, but the mean average precision is better.

5. Conclusion

This paper proposes an improved apple quality detection algorithm based on YOLOv5 model. The improved model adds the coordinate attention mechanism to yolov5, and the improved model recall rate is from 82.7% to 88.8%, and the mAP is from 83.7% to 85.9%. The experiment proves that the improved yolov5 model improves the accuracy and recognition rate, which increases the efficiency of apple quality detection and saves part of the labor force.

References

[1] WANG H, ZHANG F, LIU X F. Fruit image recognition based on DarkNet-53 and YOLOv3[J]. Journal of Northeast Normal University(Natural Science Edition), 2020,52(4):60-65.

[2] BARGOTI S, UNDERWOOD J. Image segmentation for fruit detection and yield estimation in apple orchards [J]. Journal of Field Robotics, 2017,34(6):1039-1060.

[3] Fenggang Sun, et al. Identification of apple fruit diseases using improved YOLOv5s and transfer learning [J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38 (11): 171-179.

[4] Hui Gao, et al. Research on rapid defect detection Method of Apple based on Machine vision [J]. Food and machinery,2020,36(10):125-129+148.

[5] YANG G, FENG W, JIN J, et al. Face mask recognition system with YOLOV5 based on image recognition

[6] [C] //2020 IEEE 6th International Conference on Computer and Communications (ICCC). IEEE, 2020: 1398-1404.

[7] Qibin Hou, Daquan Zhou and Jiashi Feng. “Coordinate Attention for Efficient Mobile Network Design.” 2021

IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2021): 13708-13717.