WEP Warwick Evans Publishing

International Journal of Computer Science and Information Technology

ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 4, Number 3, Year 2024
DOI: https://doi.org/10.62051/ijcsit.v4n3.40
Journal homepage: https://wepub.org/index.php/IJCSIT/index

Research and Application of Multi-Objective Workpiece Classification and Recognition Based on YOLO v5 Model

Li Xia [1, 2], Yuming Qi [1, 2]

[1] Robot and Intelligent Equipment Research Institute of Tianjin University of Technology and Education, Tianjin 300222, China

[2] Tianjin Enterprise Key Laboratory of Intelligent Robot Technology and Application, Tianjin 300350, China

ABSTRACT

Workpiece recognition is an important application of machine vision in industry. The article proposes a deep learning image recognition method for workpiece recognition and classification. First, the workpiece sample library is established, the model is built using the YOLOv5 algorithm, and the optimal model is obtained after training and parameter tuning; in the process of solving the model, it is tested for the workpiece dataset, and the results show that the model achieves an accuracy of more than 94% in the validation set, and finally, after the validation, the model can be accurately recognized and classified, and it provides effective decision-making support for the actual recognition of the workpiece.

KEYWORDS

Artifact recognition; YOLOV5; Deep learning

1. INTRODUCTION

Machine vision is a comprehensive technology that encompasses many disciplines such as image processing, control science, and mechanical engineering technology. It has an important role in astronomy, medicine, manufacturing and other fields, including sorting, measuring and testing in industry. The application of machine vision is quite popular in foreign countries, mainly focusing on electronics, automotive, metallurgy, food and beverage, spare parts assembly and manufacturing. At the same time, China is in a critical period of transformation from a manufacturing power to a manufacturing power, "Made in China 2025" clearly points out that we should accelerate the deep integration of a new generation of information technology and manufacturing industry as the main line, and promote intelligent manufacturing as the main direction of attack, and the demand for intelligent production is more urgent. To this end the authors of this paper combine machine vision technology with improved YOLO v5s modeling [1-4] Intelligent algorithms are applied to the positioning and identification of multi-targeted workpieces in the intelligent factory to realize the intelligent sorting of multiple workpieces, so as to solve the problems of low efficiency, high cost, and inability to guarantee the accuracy rate of the traditional manual sorting of workpieces.

Some researchers have already carried out some related studies on defect recognition. zhen liu et al. [5] Aiming at the problem of a large number of small and dense objects and complex background noise interference in high-altitude shooting, an improved target detection algorithm for YOLOv5 UAV shooting scene is proposed, which improves the feature extraction ability of small targets through the feature enhancement block (FEBlock) and self-centered feature expansion board (SCEP).Burchan Aydin et al. [6] proposed an improved target detection algorithm for YOLOv5 UAV filming scene, which significantly improves the model's ability to detect small and dense objects through the introduction of Feature Enhancement Block (FEBlock) and Self-Contained Feature Expansion Plate (SCEP). Many scholars have studied the defect recognition in localized situations but less on the recognition of multiple types of machine parts and surface defects in complex scenes, therefore, from the practical needs of machine parts surface damage type recognition, the article proposes a workpiece classification recognition algorithm based on the improved YOLOv5, and discusses in depth the application and development prospects of deep-learning-based machine vision inspection technology in machine parts recognition.

2. YOLOV5S

2.1. Principle of the Algorithm

The initial core model of the multi-target workpiece recognition and detection method used in this study is YOLOv5, whose structure is divided into four parts: the first part is the Input, i.e., the input side, which applies both the image pre-adjustment method and the anchor frame adaptive locking computation method to the input image; the second part is the Backbone, i.e., the backbone network, which realizes the extraction of image features and contains convolutional layers, C3 module and faster SPPF module. The second part is Backbone, i.e. backbone network, which realizes the extraction of image features, and contains convolutional layer, C3 module and SPPF module with faster computation speed, among which C3 module plays a major role in feature extraction and learning; the third part is Neck network, i.e. PANet combinatorial network, which is mainly used for fusing the features from different scales of the Backbone, and the paths of which adopt top-down and bottom-up; the fourth part is Prediction, i.e. output side, which is mainly used for detecting the features in the Backbone, and the output side is used for detecting the features in the Backbone. The fourth part is Prediction, i.e., the output side, whose main function is detection, and its specific form is to realize the function of anchoring frame on the feature map output from the Neck network, and finally output the effective prediction category. Figure 1 shows the structure of YOLOv5 algorithm.

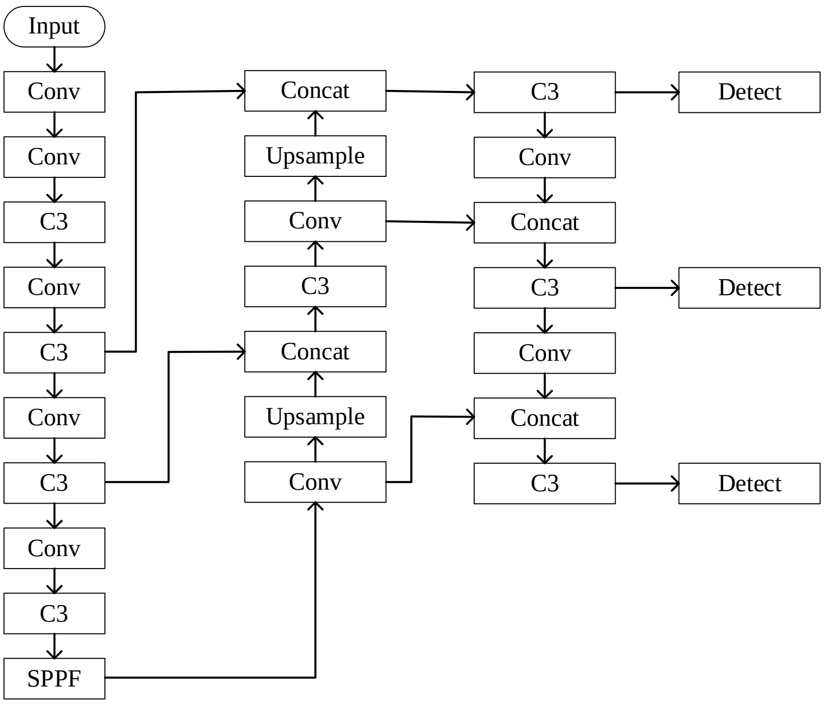

Figure 1. YOLOv5 structure

2.2. Improving the YOLOv5s Algorithm

2.2.1. Enhanced capacity for model training

Because there are many types of models to be trained, in order to avoid the situation where a single model recognizes multiple types, after the program is trained, when the model accuracy is judged on the test set, additional training is done on the parts that recognize multiple types. The features of the actual types are extracted first, and the features of the other recognized types are discriminated, the features that are the same as the actual types are eliminated, and the remaining features are strengthened to increase the weight of the features to enhance the recognition ability of the model. When using the weights of the trained model for recognition, if there are models that recognize more than one type, they will be extracted and recognized again, and at the same time, the activation function will be changed as follows:

$$\mathrm{Re}\,lux = \max\left(x1, x2, \cdots, xn, 0\right)$$

This activation function compares the confidence levels of the multiple categories that the part is recognized as, selects the category with the highest confidence level for saving, and eliminates duplicates of recognized categories.

2.2.2. Improvement of model accuracy

In order to improve the accuracy, ensure the overall recognition rate and reduce the scrap rate of the training model, we compare the recognition rate of the model types after each training round, and change the training model to improve the recognition rate if the recognition rate is lower than the required one. The basic training structure used in this article is YOLOv5s, and through the comparison after each training round, the models with low recognition rate and high fn and fp values are assigned to a larger training model. The basic training structure used in this article is YOLOv5s, and by comparing the models with low recognition rates and high fn and fp values after each round of training, the data of these types of models are divided into another profile and assigned to a larger training model, e.g., the YOLOv5s model is adjusted to the YOLOv5m model, to improve the recognition rate. Although the training time will be extended by changing to a larger model, it can effectively improve the recognition rate of the original categories that are not high, and reduce the fn and fp values.

3. MATERIALS AND DATA

3.1. Image Acquisition

The original images of the workpiece dataset used in the article study were collected from a large-scale machining factory in Tianjin, and the objects were gears, gear shafts, and bearings. The image acquisition equipment was an iPhone cell phone and CyberTrack H7, and the resolution of the data was 4032×3024 and 1920×1080 pixels, respectively, saved in PNG format. The data were collected indoors in the first half of June 2022 from 18:00 to 22:00, and the distance between the recording device and the workpiece was 5 to 50 cm, and the recording angle was 30° to 90° from the vertical downward direction when the workpiece was captured. A total of 847 raw images were captured, including data of different background complexity, near and far views, different lighting, and the number of targets in a single image.

3.2. Data Set Construction

LabelImg, a rectangular area labeling tool, is used to manually label the captured images to obtain the category and location information of the target artifacts in the images. After the labeled information is saved in the form of a .txt file, the construction of the artifact dataset is completed. In this study, the dataset is randomly divided into a training set (593 images), a validation set (169 images), and a test set (85 images) in the ratio of 7:2:1, and each dataset contains the acquired artifact images and the labeled information.

4. EXPERIMENTS AND ANALYSIS

4.1. Experimental Environment

In this paper, the dataset production, training and testing of the model are all carried out in the same workstation, and the specific experimental environment is shown in Table 1.

Table 1. Experimental environment

Category	Parameters
System ubuntu18.04	ubuntu18.04
CPU	Intel (R) Xeon (R) Platinum 8255C CPU @ 2.50GHz
GPU	RTX 3080, 10GB RAM
Libraries	PyTorch 1.7.0, Python 3.8, Cuda 11.0
Solid State Disk	300GB

In order to adapt to the training, the resolution of the training set of images are adjusted to 640 × 640 pixels, and the Stochastic gradient descent (SGD) optimizer is selected to train 300 epochs, the batch size is set to 32, the number of categories is 3, and the momentum parameter and weight decay parameter are set to 0.937 and 0.0005, respectively. 0.0005. During the experiment, it is necessary to adjust the number of convolutional kernels in the convolutional layer of the prediction module of the network, according to the actual situation, the number of categories in this paper is 3, and the correct number of convolutional kernels can be obtained by bringing in the following formula, which is calculated as follows.

$$\text{filter} = \left(class + 5\right) \times 3 \tag{1}$$

The speed of training to reach the optimal effect is largely affected by the learning rate learning_rate, if the learning rate is set too high, the training will be difficult to converge, or even overfitting; learning rate is set too low, the training can not converge for a long time. Through many experiments and comparisons, it is found that setting the learning_rate to 0.001 has a better effect. Specific parameter configuration as shown in Table 2

Table 2. Table of training parameter settings

Batch	Attenuation Optimizer	Class	filter	Learning_rate	Number of iterations
32	SGD	3	24	0.001	300

4.2. Assessment of Indicators

In this experiment, four evaluation indexes are mainly used to evaluate the model performance, namely precision (P), recall (R), floating point operations (FLOPs) and mean average precision (mAP). mAP).

In the measurement task, precision and recall are important indicators for judging the recognition effect of the networks. In addition, in order to compare the computational complexity of different networks, the article chooses the computational volumes (GFLOPs) and the number of parameters (Params) to represent the performance differences between different networks.

Is the ratio of the number of correctly predicted positive samples in the prediction sample to the number of all predicted positive samples, calculated as shown in equation (2).

$$P = \frac{TP}{TP + FP} \tag{2}$$

Recall R is the ratio of the number of correctly detected positive samples to the total number of actual positive samples, and is calculated as shown in Equation (3).

$$R = \frac{TP}{TP + FN} \tag{3}$$

The average precision AP is the area of the curve between P and R. The formula is shown in equation (4).

$$AP = \int_0^1 P(R)\,\mathrm{d}(R) \tag{4}$$

Mean Average Precision mAP represents the mean value of AP for each category, which is calculated as shown in equation (5).

$$\mathrm{m}AP = \frac{1}{N}\sum_{i=1}^{N} AP_i \tag{5}$$

Where: TP (true positive) is the number of correctly predicted positive samples; FP (false positive) denotes the number of incorrectly predicted positive samples; FN (false negative) denotes the number of positive samples in the sample that were not detected; R denotes the rate of checking completeness, which indicates how many positive samples will be correctly predicted; n denotes the number of detected categories and the unit is one.

5. EXPERIMENTAL RESULTS

Four types of workpieces, namely, bearings, gears, and gear shafts, were detected in the experiment, and the images of the workpieces at different angles were taken for detection, and the detection effect is shown in Fig. 2.

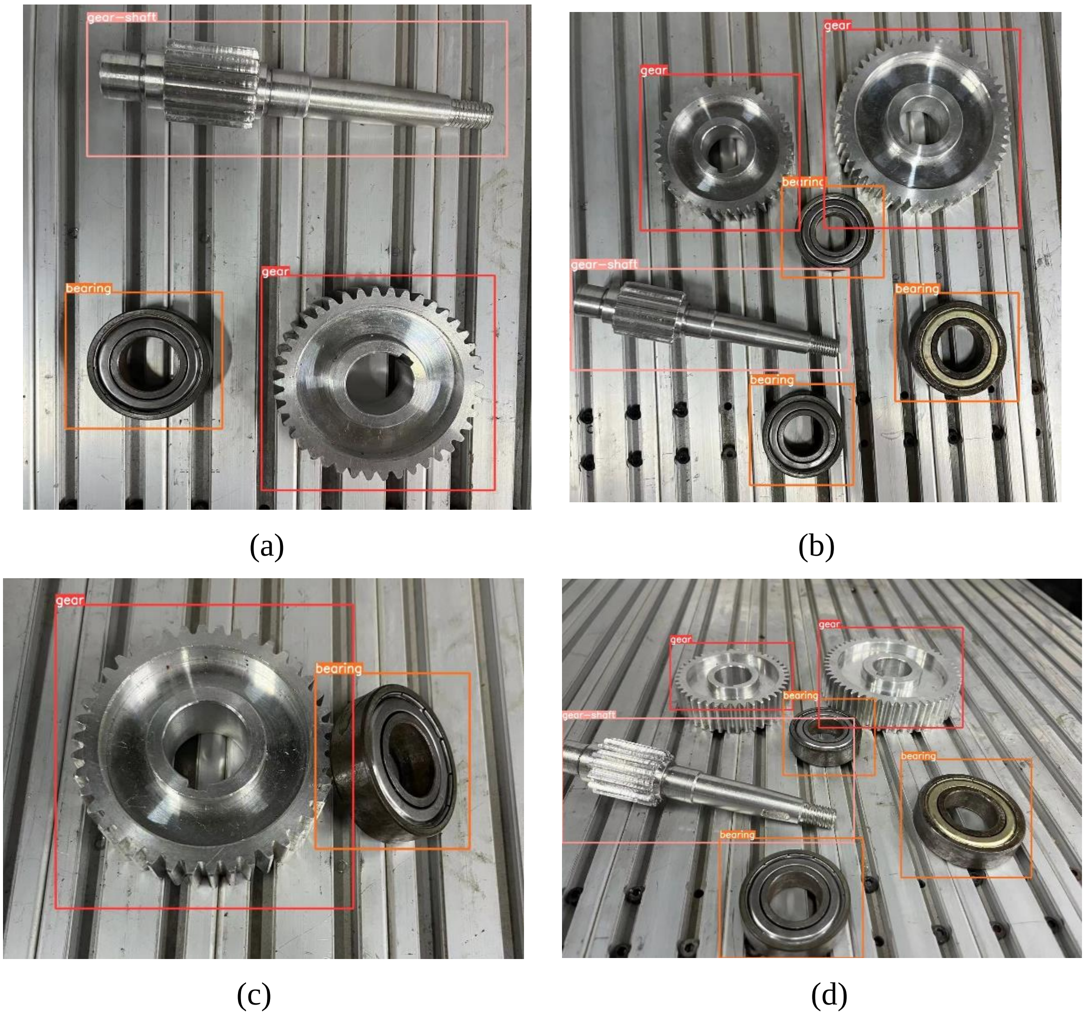

Figure 2. Workpiece inspection results

As can be seen in Figure 2, during the actual detection process, when the position and angle of the workpiece changes, the detection method used in this paper is still able to effectively detect the workpiece and classify the detected workpiece with labels.

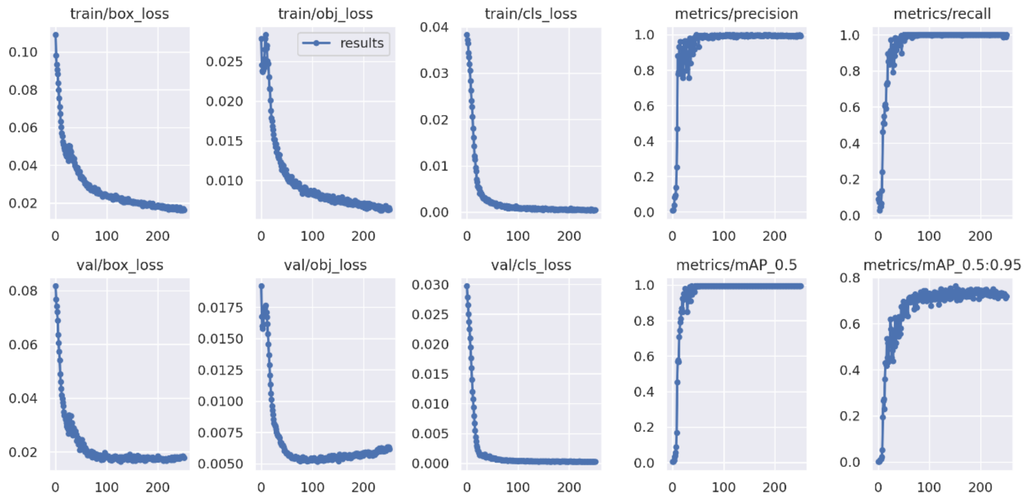

Figure 3. Running results

6. CONCLUSION

Aiming at the traditional workpiece detection template matching algorithm used by the workpiece angle changes, position changes by the influence of the larger, poor robustness of the problem, this paper adopts the YOLO v5s algorithm for workpiece detection. YOLO v5s target recognition algorithm on the workpiece recognition effect is good, and is not subject to the interference of the change of the angle of the workpiece, the change of the position of the workpiece interference. This paper uses several training indicators: Precision (accuracy), Recall (recall rate), Fps (detection rate), the evaluation results as shown in Table 3, the workpiece recognition method used in this paper can be guaranteed under the premise of accuracy, has a high detection rate, can better meet the requirements of the workpiece recognition method in the actual industrial production.

Table 3. Detection model test

Detection Algorithms	Precision / %	Recall / %	Detection rate (frames/sec)
YOLO v5s	94	94	80

The workpiece recognition method used in the article has good accuracy and real-time performance, which can identify and classify the target workpieces, and provides technical support for the important links such as disorderly sorting of workpieces, intelligent assembly, and automated dimensioning in industrial automated production.

REFERENCES

[1] Xue, Z.; Xu, R.; Bai, D.; Lin, H. YOLO-Tea: A Tea Disease Detection Model Improved by YOLOv5. Forests 2023, 14 (2), 415. https://doi.org/10.3390/ f14020415

[2] Xu, L.; Dong, S.; Wei, H.; Ren, Q.; Huang, J.; Liu, J. Defect Signal Intelligent Recognition of Weld Radiographs Based on YOLO V5-IMPROVEMENT. j. Manuf. . Process. 2023, 99, 373-381. https://doi.org/10.1016/j.jmapro.2023.05.058

[3] Wang, L.; Liu, X.; Ma, J.; Su, W.; Li, H. Real-Time Steel Surface Defect Detection with Improved Multi-Scale YOLO-V5. Processes 2023, 11 (5), 1357. https://doi.org/10.3390/pr11051357

[4] Taşyürek, M.; Öztürk, C. A Fine-Tuned YOLOv5 Deep Learning Approach for Real-Time House Number Detection. peerJ Comput. Sci. 2023, 9, e1453. https:// doi.org/10.7717/peerj-cs.1453

[5] Liu, Z.; Gao, X.; Wan, Y.; Wang, J.; Lyu, H. An Improved YOLOv5 Method for Small Object Detection in UAV Capture Scenes. IEEE Access 2023, 11, 14365 -14374. https://doi.org/10.1109/ACCESS.2023.3241005

[6] Aydin, B.; Singha, S. Drone Detection Using YOLOv5. Eng 2023, 4 (1), 416-433. https://doi.org/10.3390/eng4010025