

# Research on Tool Detection Algorithm based on YOLOv8 Improved Model

Wanpeng Qi

Tianjin University of Technology and Education, China

**Abstract:** With the continuous improvement of people's living standards, the public's awareness of protecting the right to life safety has become increasingly important, and higher requirements have been put forward for the prevention of public safety. Controlled cutting tools are a part of cutting tools, which are controlled and managed by law to prevent criminals from using them as weapons for illegal activities. To prevent criminals from committing crimes in places with high population mobility, such as subways, campuses, airports, and other enclosed areas with high population mobility. This article proposes a controlled tool detection algorithm based on the improved YOLOv8 model to improve the detection accuracy and efficiency of controlled tools. After experiments, SE attention mechanism was added to the original YOLOv8 algorithm model, and the loss function SIOU was added to the YOLOv8 algorithm. The improved mAP increased from the initial 80.8% to the improved 85%, an increase of 4.2%. This provides a certain reference value for future algorithm models for identifying controlled tools.

**Keywords:** YOLOv8; SE Attention Mechanism; Loss Function SIOU; mAP.

## 1. Introduction

With the increasing prosperity of our country's national strength, the continuous improvement of economic level, and the continuous improvement of people's happiness index, the public's awareness of protecting the right to life and safety is also constantly increasing. Controlled knives play an essential role in daily life, but if they fall into the hands of criminals, they can pose a significant threat to the safety of citizens, especially in densely populated areas such as subways, campuses, airports, and other closed areas with high population mobility. The monitoring and control of knives should be strengthened. This also places higher demands on algorithms for identifying controlled cutting tools.

Object detection is one of the most important domains in computer vision tasks, which is an important branch of artificial intelligence. It aims at finding and locating the accurate position of objects in given pictures or videos. With the development of deep learning techniques, more powerful and robust algorithms have emerged to deal with multi-scale, high-level features to overcome the limitations of traditional pipeline of object detectors[1]. At present, object detection algorithms are mainly divided into two types. One is the two-stage algorithm represented by R-CNN series algorithms. The other one is the one-stage algorithm represented by you only look once (YOLO) series algorithms. The two-stage target detection algorithm is based on the idea of candidate region[2]. The two-stage object detection algorithm has high detection accuracy, but the detection speed is slower compared to the single object detection algorithm. One-stage target detection algorithm is based on the global regression and classification, which directly generates the location and category of the target object. Compared with Two-stage algorithm, One-stage algorithm is more real-time, but the detection accuracy is lost[3]. Due to the fact that identifying controlled tools belongs to small target recognition and requires real-time detection of whether pedestrians are carrying controlled tools, we need a single target detection algorithm with faster detection speed to undertake the detection task in the algorithm for identifying controlled tools.

Due to the small target size of controlled cutting tools and the complexity and diversity of detection scenarios, it is necessary to design efficient, accurate, and diverse algorithms with diverse application scenarios. Therefore, this article will make certain improvements to the YOLOv8 algorithm model to make the improved model more efficient in recognition.

## 2. YOLOv8 Object Detection Algorithm

### 2.1. YOLOv8 Algorithm Model

YOLO has become a leading object detector due to its excellent speed accuracy balance, and YOLOv8 was opened by Ultralytics in January 2023, providing state-of-the-art object detection performance[4]. In the YOLOv8 algorithm, there are five different models: YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l, and YOLOv8x. The depth and width of these five models are constantly increasing. Among the five models, YOLOv8n has the smallest depth and width, because as the depth and width of the models increase, although the detection accuracy increases, the detection speed decreases. As the target of detection is controlled cutting tools and belongs to small target detection, improving detection speed cannot be ignored in improving the algorithm model. Therefore, this article will adopt the YOLOv8n model and improve the algorithm based on it to improve the recognition efficiency of the model.

### 2.2. Overview of YOLOv8 Algorithm Structure:

The YOLOv8 algorithm provides a brand new SOTA model, which also offers models of different sizes based on scaling coefficients to meet different scene requirements, while supporting image classification, object detection, instance segmentation, and pose detection tasks. The YOLOv8 algorithm is divided into three parts: Backbone, Neck, and Head. Backbone network usually refers to the main part of the entire deep neural network, responsible for extracting features from input data. Backbone networks typically consist of multiple convolutional layers or other

feature extraction layers, used to gradually extract advanced features from input data. Neck: Responsible for further processing and integrating the features extracted from the backbone network.

Head is responsible for performing specific tasks such as classification, object detection, or semantic segmentation. In YOLOv8, the C3 structure in YOLOv5 was replaced with a C2f structure with richer gradient flow in the backbone network and Neck section, which improved the performance of the model. The Head section has changed from the original coupling head to the understanding coupling head, and has changed from Anchor Based in YOLOv5 to Anchor Free. In terms of computation, the Task AlignedAssigner positive sample allocation strategy has been adopted, and Distribution Focal Loss has been introduced the network structure of YOLOv8 is shown in Figure 1:

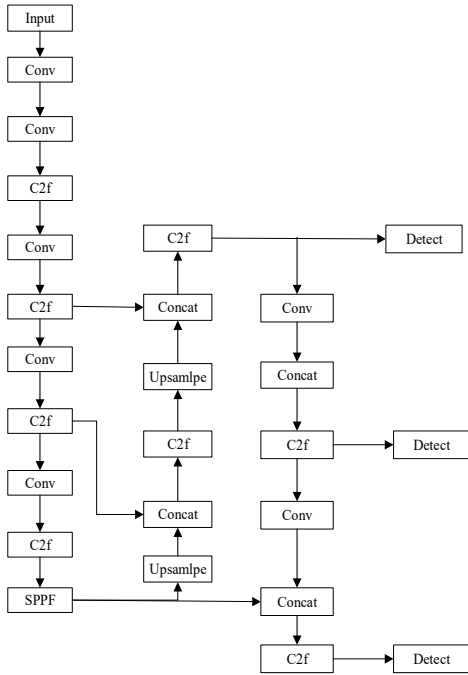


Figure 1. The network structure of YOLOv8

### 3. Network Improvement of YOLOv8

#### 3.1. SEAttention Attention Mechanism

The innovation of SEAttention lies in its focus on the relationships between channels, hoping that the model can automatically learn the importance of different channel features, and focus more on the channel features with the highest amount of information, while suppressing those unimportant channel features.

SEAttention mainly consists of two parts: Squeeze and Excitation [5]. Firstly, the SE attention module aggregates feature maps across spatial dimensions through global pooling, i.e.  $H * W * C \rightarrow 1 * 1 * C$ , and converts them into feature maps of size  $1 * 1 * C$ , through the Squeeze compression operation; The incentive process involves non-linear changes to the compression results, combining the learned channel attention information with feature maps to reduce the complexity of the model and improve its generalization ability. The structural diagram of SEAttention is shown in Figure 2, Place the SEAttention attention mechanism module in the backbone, as shown in Figure 3:

#### 3.2. Loss Function SIOU:

Object detection is one of the core issues in computer vision tasks, and its effectiveness largely depends on the definition of the loss function. The traditional object detection loss function relies on the aggregation of bounding box regression indicators, such as the distance, overlapping area, and aspect ratio between the predicted box and the true box (i.e. GIoU, CIoU, ICIoU, etc.). In this article, a new loss function SIOU is proposed, which redefines the penalty metric taking into account the vector angles between the required regressions. SIOU consists of four parts, namely angle loss, distance loss, shape loss, and IOU loss. SIOU can effectively reduce the impact of errors on the loss function and improve network robustness. When detecting small targets, the estimation method used by SIOU is weighted. the SIOU loss function is used to improve the convergence speed and detection accuracy of the model [6]. Applying the SIOU loss function to the algorithm model for detecting controlled tools can also improve the recognition efficiency of the model. The SIOU loss function formula is shown in (1), (2), (3), and (4):

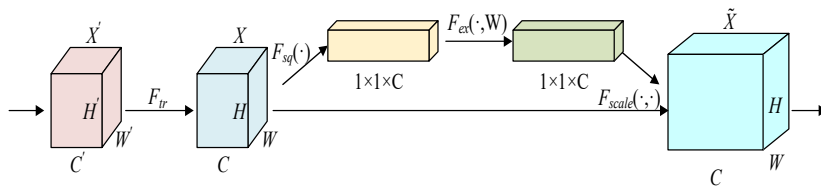


Figure 2. The structural diagram of SEAttention

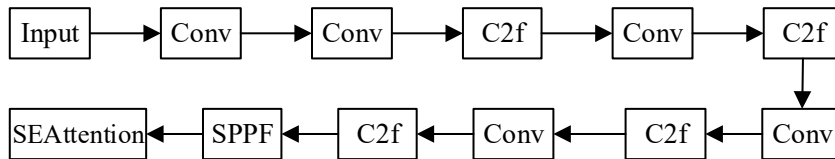


Figure 3. Place the SEAttention attention mechanism module in the backbone

The calculation formula for angle cost is:

$$A = 1 - 2\sin^2\left(\arcsin(x) - \frac{\pi}{4}\right) \quad (1)$$

The definition of distance cost is:

$$\Delta = \sum_{t=x,y} (1 - e^{-\gamma \rho_t}) \quad (2)$$

The definition of shape cost is:

$$\Omega = \sum_{t=w,h} (1 - e^{-w_t})^\theta \quad (3)$$

The definition of the SIOU loss function is:

$$SIOU\_LOSS = 1 - IOU + \frac{\Delta + \Omega}{2} \quad (4)$$

$\Delta$ : Distance loss,  $\Omega$ : Shape loss,  $\theta$ : The degree of attention paid to shape loss

## 4. Experimental Results and Analysis

### 4.1. Dataset Processing:

The dataset used in this experiment is partly from video screenshots of controlled cutting tools, and partly from controlled cutting tool datasets sold on the internet. Clear or complex scene controlled cutting tool images are selected from them and expanded to 600 datasets through data augmentation methods (such as increasing brightness, flipping, etc.). Randomly divide the images into 480 training set images and 120 validation set images in an 8:2 ratio. Use labeling to frame and label the controlled tools in the images with the keyword "knife".

### 4.2. Experimental Environment:

After preparing the dataset, a remote server with RTX 4090, PyTorch 1.11.0, Cuda 11.3 graphics card, programming language in Python, and RTX 4090 (24GB) \* 1 GPU was created on Autodl, and a YOLOv8 GPU runtime environment was created on the server.

### 4.3. Evaluation Criteria:

The indicator used to evaluate the improvement effect of the algorithm in this article is the average accuracy mAP, As shown in (5):

$$mAP = \sum AP / N \quad (5)$$

mAP stands for mean average accuracy, which generally refers to the average value of AP for all categories within all images.

The P-R curve is a curve enclosed by Precision and Recall coordinates, respectively. By comparing the PR curve and map obtained from the experiment, the efficiency of the improved algorithm model can be determined.

### 4.4. Result Analysis:

After 200 rounds, the mAP of the original model is shown in Figure 4, and the mAP of the improved model is shown in Figure 5:

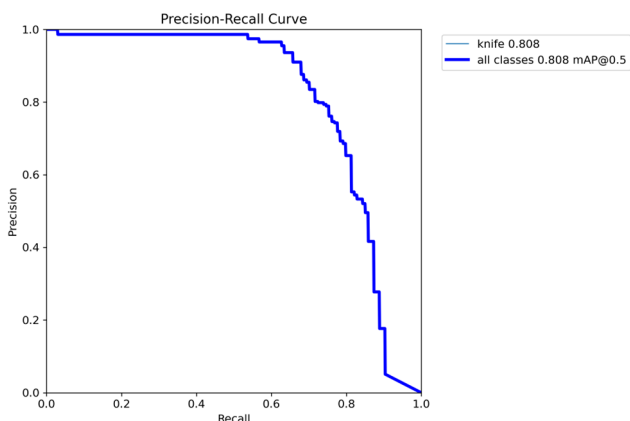


Figure 4. The mAP of the original model

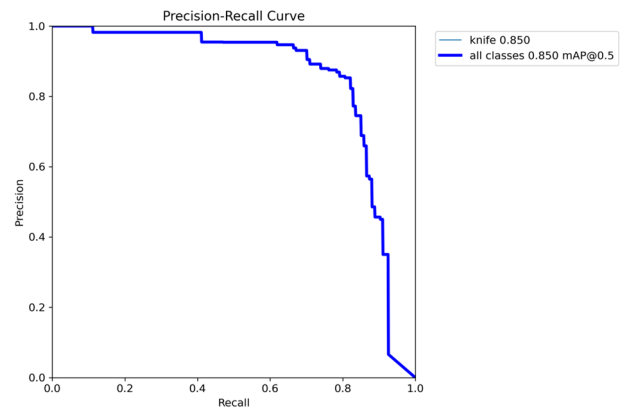


Figure 5. The mAP of the improved model

The improved YOLOv8 model has been trained a total of 200 rounds, and the mAP has increased from the initial 80.8% to the improved 85%, an increase of 4.2%. Therefore, experimental results have shown that the improved algorithm performs well in detecting controlled cutting tools.

## 5. Summary

This article proposes a controlled tool detection algorithm based on the improved YOLOv8 model. After multiple experiments, SE attention mechanism was added to the original YOLOv8 algorithm model, and the loss function SIOU was introduced into the YOLOv8 algorithm. The improved mAP increased from the initial 80.8% to the improved 85%, an increase of 4.2%. This provides a certain reference value for future algorithm models for identifying controlled tools.

## References

- [1] Sun Y, Sun Z, Chen W . The evolution of object detection methods [J]. Engineering Applications of Artificial Intelligence, 2024, 133(PE):108458-.
- [2] Liu Y, Chen H . A Lightweight Container Detection Network Model Based on YOLOv5s[C]// (Technical Committee on Control Theory, Chinese Association of Automation), (Chinese Association of Automation), (Systems Engineering Society of China). Proceedings of the 42nd Chinese Control Conference (8). School of Artificial Intelligence, Hebei University of Technology; Control Engineering Technology Innovation Center of Hebei Province, Hebei University of Technology, 2023:6. DOI: 10. 26914/c. cnkihy. 2023.036061.
- [3] LIANG F, ZHOU Y, CHEN X, et al. Review of Target Detection Technology based on Deep Learning[C]// (APISE). Proceedings of 5th International Conference on Control Engineering and Artificial Intelligence (CCEAI 2021).ACM, 2021:4.DOI: 10.26914/ c.cnkihy. 2021. 001576.
- [4] Liu, Qingtang, et al. "YOLOv8n\_BT: Research on Classroom Learning Behavior Recognition Algorithm Based on Improved YOLOv8n." IEEE Access (2024).
- [5] Zhigang L, Baoshan S ,Kaiyu B .Optimization of YOLOv7 Based on PConv, SE Attention and Wise-IoU[J].International Journal of Computational Intelligence and Applications, 2024, 23(01).
- [6] Feng L, Kang X ,Zhengpeng H , et al. Fabric defect detection algorithm based on improved YOLOv5[J]. The Visual Computer, 2023,40(4):2309-2324.