Journal of Artificial Intelligence Practice (2022)
Clausius Scientific Press, Canada
DOI: 10.23977/jaip.2022.050206
ISSN 2371-8412 Vol. 5 Num. 2

A Medical Waste Classification Method Based on Yolov4

Yupeng Mou[1,a], Liyan Zhang[1,b,*]

[1]*School of Computer & Telecommunication Engineering, Dalian Jiaotong Unniversity, Huanghe Road, Dalian, China*
[a]*631440689@qq.com*, [b]*zhangliyan@126.com*
**Corresponding author*

Keywords: YoloV4, CBAM, Classification

Abstract: In recent years, with the outbreak of corona virus, medical waste will also become a way of virus transmission. Therefore, this paper presents a medical waste classification method based on improved YoloV4. This method uses hard mish activation function to replace mish activation function on the basis of YoloV4 algorithm, which reduces the amount of calculation and improves the classification speed. Two CBAM Block attention mechanism module are added, and the experimental results verify the effectiveness of this algorithm.

1. Introduction

In 2019, the corona virus broke out, and the hospital cleaning staff will be affected by the virus when they handle the medical waste, which will cause the spread of the virus. Therefore, this paper presents a medical waste treatment method based on neural network. According to the classification principles of medical waste treatment, the waste classification is completed through the target detection algorithm [1], which can be more conducive to the waste disinfection and prevent the spread of the virus. It can also realize full intelligent garbage classification by controlling the manipulator. By training convolutional neural network, this paper presents an improved algorithm based on YoloV4 algorithm, which can better realize garbage classification, and the classification speed is faster than manual speed.

2. Improved YoloV4 algorithm

Yolo algorithm was proposed by Redmon et al. in 2016. [2] This method can improve the speed of the target detection. The algorithm converts the detection process into a regression problem and discards the candidate region generation stage in the two-stage detection algorithm represented by mask- RCNN, which greatly improves the detection speed. The backbone feature extraction network of YoloV4 uses cspparknet53, which integrates CSP structure into darknet53 of Yolov3[3]. Its advantage is that it can enhance the learning ability of CNN, eliminate the computing bottleneck and reduce the memory occupation. That is to reduce the amount of the network computing and the occupation of video memory, while ensuring that the capacity of the network remains unchanged or slightly improved. In order to improve the classification speed, this paper replaces the hard mish activation function with the mish activation function in YoloV4, and introduces the attention module.

2.1. Activate function

In 2017 [4], Google tried to find a new activation function through automatic search technology. It found many innovative activation functions through violent search and reinforcement learning based search, and directly evaluated these activation functions through experiments rather than theoretical analysis, thus obtaining the swish activation function as shown in equation (1).

$$Swish(x) = x * sigmoid(\beta * x) \tag{1}$$

However, in 2019, it was found that the swish activation function only works in the deep network and has a certain amount of computation. In order to solve these problems, HardSwish activation function [5] as shown in equation (2), that is, the hard coded version of swish activation function, was proposed. Since no sigmoid function is used by swish, HardSwish requires significantly less computation than the former and is more friendly to model lightweight.

$$\text{HardSwish}(x) = \begin{cases} 0 & \text{if } x \le -3 \\ x & \text{if } x \ge +3 \\ x*(x+3)/6 & \text{otherwise} \end{cases} \tag{2}$$

2.2. CBAM

CBAM (Convolutional Block Attention Module) is a plug and play module, which can be placed behind any layer in theory. However, if the attention mechanism module is inserted into the backbone network, the pre training weight of the network will not be used. Therefore, this paper chooses to place the CBAM Block behind the backbone network. Too many attention mechanism modules will seriously affect the training speed. Considering the small data set in this paper, we choose to insert two CBAMs, which are located between the SPP after 3 times of splicing and convolution and the down sampling and 5 times of splicing and convolution.

2.3. Recommended algorithm network structure

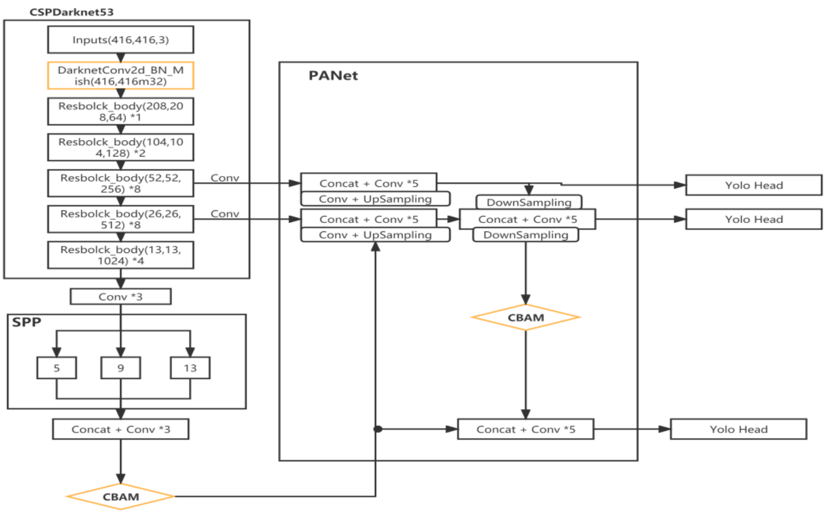


Figure 1: Improved YoloV4 network structure diagram

Taking the input as (416 * 416) as an example, the network structure of the improved YoloV4 algorithm is shown in Figure 1.

3. Algorithm simulation and result analysis

3.1. Data set establishment

In this paper, we try to collect and screen out the pictures that are suitable for the data set from the network. Each category retains at least 300 pictures. In this paper, we use data enhancement technology to expand the data set. By rotating and changing the color of the original image, we can not only retain the characteristics of the original object, but also expand the data set to improve the training effect. After data enhancement, a total of 12099 medical waste images were used in this training. The details of the new dataset are shown in Table 1.

Table 1: Detailed Information of the Datasets

Object name	Number of pictures (PCs.)
m_bottle	1936
m_box	1672
syringe	1012
thermometer	1242
rabbits	1474
scalpel	704
rats	1551
glaves	1474
caps	475
capsule	559
totle	12099

3.2. Algorithm parameter setting

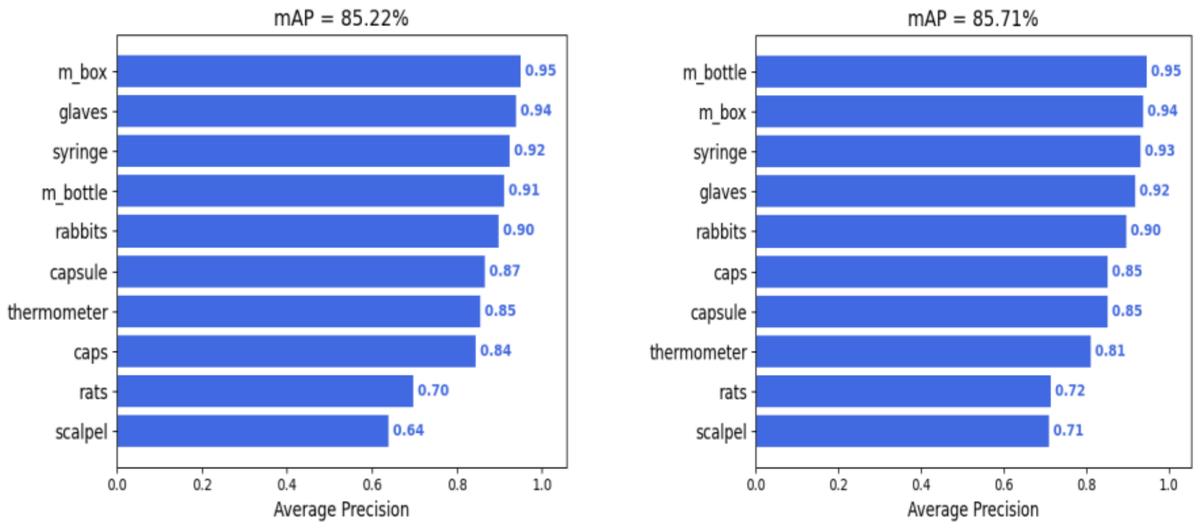
Network hyper-parameter settings are shown in Table 2.

Table 2: Network hyper-parameter settings

Batch Size	8
initial learning rate	0.0001
smooth label	0.005
optimizer	Adam
weight decay	—
input shape	608*608
epoch	300
NMS	0.45

3.3. Simulation results and analysis

(1) The precision ratio refers to the proportion of "true" samples that are actually true among the "true" samples, mAP is the average precision rate, the average of the precision rates of all categories. As shown in Figure 2.

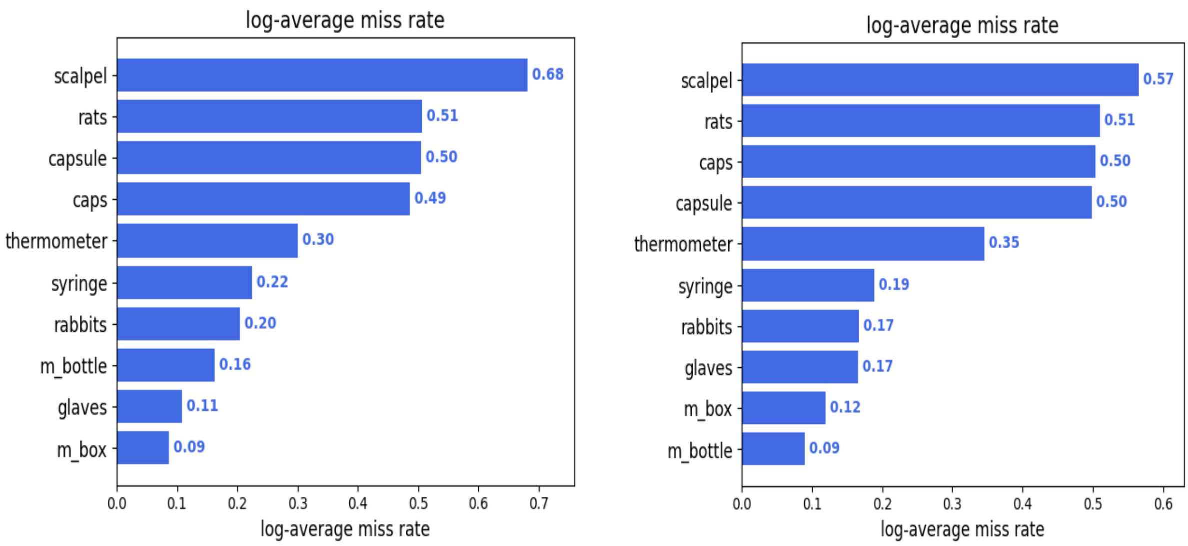


(a) YoloV4 (b) Improved YoloV4

Figure 2: Result of mAP

It can be seen from the experimental results, the map of improved YoloV4 is the highest, 85.71%; It is 0.49% higher than 85.22% of YoloV4, with significant improvement. The precision ratio YoloV4 and YoloV4 for each subclass. The distribution of improved YoloV4 is roughly the same, which is the difference caused by different network structures.

(2) Lamr (log average miss rate): it is the missed detection rate, which indicates the proportion of undetected targets in the total number of targets. The lower the value, the better the network performance, as shown in Figure 3.



(a) YoloV4 (b) Improved YoloV4

Figure 3: Result of lamr

From the results of average false detection rate, it can be seen that the improved YoloV4 has the best overall performance and the lowest average false detection rate.

(3) FPS (frames per second) is the number of detected images per unit time. It is used to evaluate whether the speed of the model designed in the computer hardware device can be detected in real time through reasoning and detecting the number of frames or not, as shown in Table 3. The

improved YoloV4 is faster than YoloV4.

Table 3: Detecting Speed of Different Network

Network	Network parameter	Model size (MB)	Detection speed (FPS)
YoloV4	64,171,171	244.79	36.99
Improved YoloV4	64,433,511	248.79	34.61

4. Conclusions

This paper presents an improved YoloV4 garbage classification method. By improving the activation function and introducing CBAM, the overall performance of the network is improved. The simulation results show that the algorithm in this paper is effective.

References

[1] L.Y. Zhang, Y.X. Zhao, Y.P. Mou, et al. Automatic classification algorithm of train common garbage based on resnet50, Journal of Dalian Jiaotong University, 2021, 42 (04): 101-105

[2] J Redmon, S Divvala, R Girshick, et al. You only look once: Unified, real-time object detection, IEEE conference on computer vision and pattern recognition. 2016: 779-788.

[3] X.T. Du, L.Y. Zhang, Y.X. Zha, et al. Vehicle distance measurement algorithm based on Yolov3, Computer programming skills and maintenance, 2020, (04): 15-16.

[4] P.Ramachandran, B. Zoph, Q.V.Le. Searching for activation functions. arXiv preprint arXiv:1710.05941, 2017.

[5] A.Howard, M. Sandler, G. Chu, et al. Searching for mobilenetv3. the IEEE/CVF international conference on computer vision. 2019: 1314-1324.