International Journal of Engineering Research And Management (IJERM)
ISSN: 2349- 2058, Volume-10, Issue-08, August 2023

Fire Detection Algorithm Based on Improved YOLOv5

Yuchen Xie, Zhou Yu, Mengyuan Liu

***Abstract*— Aiming at the existing deep learning-based flame detection methods, which have problems such as slow detection speed, large model difficult to deploy to edge computing devices and poor flame detection effect for small targets, an improved model based on YOLOv5 is proposed and applied to flame detection. Firstly, to address the problem of the model being large and difficult to deploy, this paper proposes a lightweight evolution method based on GhostNet; in order to further improve the accuracy and the effect of detecting small target flames, Coordinate Attention is added to the backbone network, which makes the network pay more attention to important information and reduces the influence of useless information, thus improving the performance of the network; finally, the EIoU Loss instead of CIoU Loss as the loss function of the algorithm, which improves the localization accuracy while increasing the rate of bounding box regression. The experimental results show that the new model obtained based on the above improvement method has obvious advantages, the model size is only 7.6MB, which is 45.3% smaller than the original model; the inference speed reaches 99.4FPS on RTX3080 device, which is 12.8% higher than the original model; and the mAP reaches 0. 881, which is 2.3% higher than the original model.**

***Index Terms*—Fire detection, smoke detection, YOLOv5, attention mechanism.**

I. Introduction

Fire is an extremely hazardous event, posing a great threat to the economy, the environment and people's lives. Negligence of living fire, industrial fire operation is not standardized, and even man-made or natural factors lead to fire are a serious threat to the safety of people's lives and property. The occurrence of fire is characterized by complex causes, many origins, high incidence rate, difficult to find, and the location of the fire is difficult to predict. In factories and forest areas that are difficult to inspect manually, due to the dense distribution of materials in these areas, the fire hazards are large and the potential risk is high, and any hidden and tiny fire starter may become the source of a large fire.

Traditional fire detection based on physical sensors is realized by detecting chemical information such as smoke, light, sound and infrared of the flame. It is easy to realize and effective in simple scenarios, but it also has disadvantages such as small monitoring range, susceptibility to external interference, and slow detection speed. Fire detection based on machine learning algorithms, compared with the sensor detection of fire accuracy is higher, more suitable for complex scenes, but often need to be based on the characteristics of the flame for complex artificial feature extraction, and generalization is general, the accuracy and detection speed still need to be further improved.

Target detection algorithms based on deep learning gradually show their advantages, and the classical target detection algorithm YOLOv5 excels in the model in all aspects. Therefore, an improved flame detection algorithm based on YOLOv5s is proposed to address the existing problems. First, on the basis of the existing small fire and smoke dataset, the dataset is supplemented by searching fire and smoke related images on the Internet through crawler technology to expand the original small dataset in order to achieve better training effect. Then the network structure of YOLOv5s is improved, and the GhostNet lightweight network is used to replace YOLOv5s CSPDarkNet-53 in YOLOv5 to complete the feature extraction, and add Coordinate Attention after the backbone network, and finally improve the loss function to improve the comprehensive performance of the algorithm.

II. YOLOv5 Algorithm

A. Introduction to YOLOv5

YOLOv5 is one of the widely used target detection networks, which has achieved good results in various industrial problems by virtue of its high detection accuracy and fast inference speed. Meanwhile, YOLOv5 has been updating and iterating, and the current version is 6.0, which is also the best overall performance version [1]. The YOLOv5s network is a lighter version of YOLOv5, which is more in line with the requirements of real-time fire detection.The network structure of YOLOv5s is shown in Figure 2.1. The network structure of YOLOv5s is shown in Fig. Its backbone network consists of Focus, Conv, C3 and SPP [2]. The definitions and roles of each of these structures are as follows. The Focus structure reduces the computation of the convolutional layer by slicing the input image, so as to improve the speed of convolutional processing. Assuming the input image is 640x640x3, Focus slices it into 320x320x12 feature maps, and then generates the feature maps by 64 convolution kernels.

Manuscript received August 23, 2023.

Yuchen Xie, School of Computer Science and Technology, Tiangong University, China
Zhou Yu, School of Computer Science and Technology, Tiangong University, China
Mengyuan Liu, School of Computer Science and Technology, Tiangong University, China

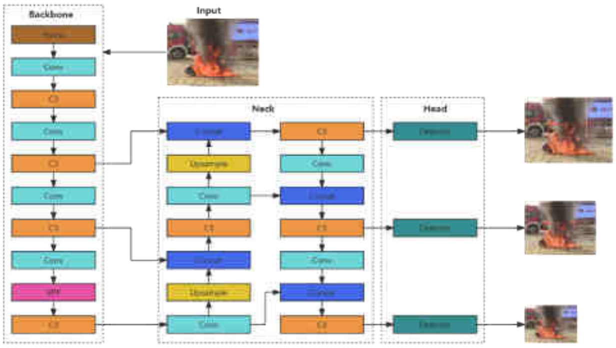

Figure 2.1 YOLOv5 structure

B. Backbone

C3 is used for feature extraction by a combination of BottleneckCSP and Conv of the CSPDarkNet53 network, which splits the input into two parts. One part is processed by the dense block and the other part is sent directly to the next stage without any processing, which allows the different dense layers to iteratively learn the replicated gradient information.

The SPP (Spatial Pyramid Pooling) structure is called Spatial Pyramid Pooling, which maximizes the pooling by three convolution kernels of sizes 5, 9, and 13, respectively, to improve the sensory field of the network.

C. Feature Fusion

After the input image is processed by the backbone network, the features of the image will be transformed into semantic features. Starting from the lower layer, the deeper the input image passes through the network, the higher the complexity of its semantic features, while the spatial resolution of the feature map is decreasing due to the downsampling process, which will lead to the loss of spatial information and fine-grained features. In order to preserve the fine-grained features, the idea of FPN (Feature Pyramid Network) [3] is applied to the feature fusion network of YOLOv5, and PANet (Path Aggregation Network) [4] is used as the feature fusion network.

D. Prediction Network

The YOLOv5 prediction network consists of three detectors, each of which is assigned three Anchors of different scales to predict the feature maps of different scales in the P3~P5 detection layer based on the bounding box, which ensures that the feature maps used for detection contain both the visual features of the lower layers of the input image and the semantic information of the higher-level features.

III. Improved Method

A. Model Lightweight

GhostNet [5] is a classical lightweight convolutional neural network. The core idea of the model is to use the Ghost module to generate more Ghost feature maps from a set of intrinsic feature maps with a low resource linear transformation, so as to tap more intrinsic feature information, thus reducing the redundancy of feature maps in traditional convolutional neural networks and realizing the lightweight of the network. The Ghost module is a plug-and-play base block, and the Ghost bottleneck is composed of it. The Conv module and the C3 module of YOLOv5s have redundant feature maps in the convolutional layer, so this paper uses the Ghost bottleneck to reconstruct the feature layer. to reconstruct the feature layer.

B. Add Coordinate Attention

Attention Machanism in Convolutional Neural Networks (CNNs) is a method of allocating network resources to improve the performance of the network [6], which makes the network pay more attention to important information and reduces the impact of useless information, thus improving the performance of the network. Typically Channel Attention brings more significant performance improvement to the model, but they usually ignore the position information [7]. In this paper, we choose Coordinate attention [8], which decomposes the channel attention into two one-dimensional features for encoding, along different spatial directions, capturing the long-range dependence and precise location information, respectively, and encoding to form direction-aware and location-sensitive attention feature maps to be applied to the input feature maps in a complementary way, so as to enhance the representation of the target of interest, and its structure is shown in Figure 3.1.

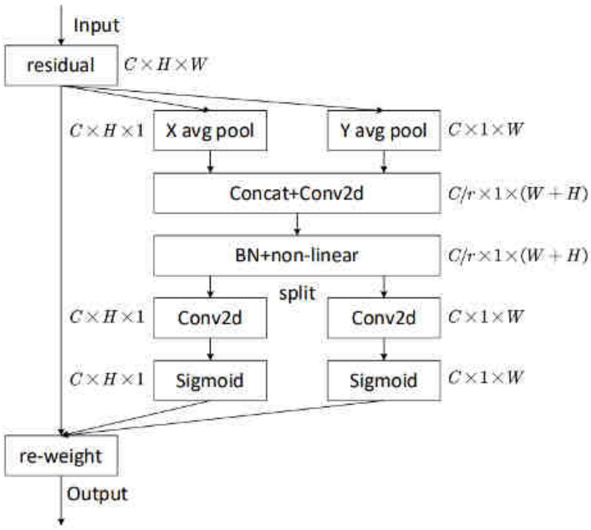

Figure 3.1 Coordinate attention structure

Considering that replacing the Ghost bottleneck bottleneck block may lead to a decrease in the characterization ability of the feature maps in the feature layers at each scale, thus affecting the detection accuracy of the new model, Coordinate Attention is introduced at the end of the feature layers from P3 to P5 to obtain the flame feature information with stronger correlation [9].

After reconstructing the backbone network using Ghost Bottleneck according to the above and adding Coordinate Attention later, the network structure of the model is shown in Figure 3.2.

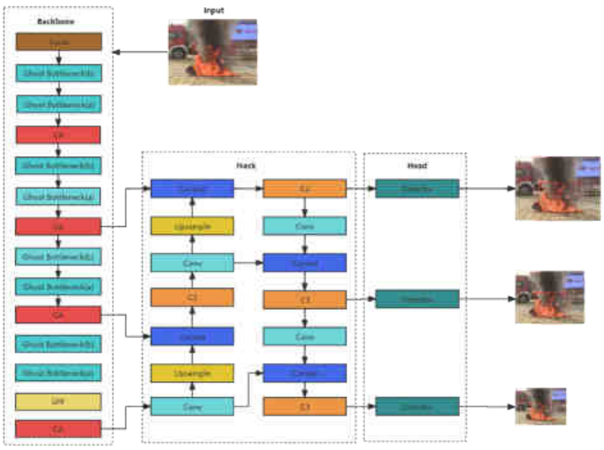

Figure 3.2 Improved model structure

C. Improvement of the loss function

CIoU Loss considers the coverage area, centroid distance and aspect ratio, and is able to measure its relative position well, and at the same time, it can solve the problem of optimizing the horizontal and vertical directions of the prediction frames, but the method does not consider the matching of the directions between the prediction frames and the target frames, which is insufficient and leads to a slow convergence speed. In the edge regression loss, the width and height loss in the EIOU Loss [10] makes the convergence speed faster and the accuracy higher, which is better than the CIOU Loss in the original network, so this paper adopts the EIOU Loss edge regression loss function with better performance.

IV. EXPERIMENTS

A. Comparative Experiment on Lightweight Improvement

In order to verify the effectiveness of the lightweight improvement, this paper uses Ghost bottleneck to reconstruct the new backbone network for feature extraction based on YOLOv5s, and the improved model is named YOLOv5s-G, which is compared with YOLOv3 [11], YOLOv4 [1], and YOLOv5 series. Through experimental comparisons, it is found that the addition of the Coordinate Attention module can effectively improve the comprehensive performance of the model by sacrificing a small amount of model volume and detection speed in exchange for the improvement of the checking accuracy P and the mean average precision mAP.

B. Comparative Experiments on Loss Function Improvement

In order to verify the impact of the improved loss function on the performance of the algorithm, the loss function in the algorithm is replaced individually in the experiments, and a side-by-side comparison experiment is carried out using the CIOU Loss and the EIOU Loss in the original algorithm, respectively, and the results are shown in Table 4.1.

Table4.1 Comparative Experimental Results of Loss Function Improvement

Function	mAP	FPS(frame/s)
CIOU	0.861	88.1
EIOU	0.872	88.7

From the results, it can be clearly seen that EIoU Loss has an advantage over CIoU Loss in terms of detection speed and measurement precision, and this experiment proves the superiority of the EIoU loss function, which improves the model performance more.

C. Model Detection Effect

By comparison, we can find that using Ghost bottleneck to reconstruct the new backbone network can greatly reduce the number of parameters in the model and make the model size smaller; while using Coordinate Attention can improve the detection accuracy while the detection speed of the model size is almost unchanged; and the performance of EIoU Loss in the present dataset has advantages in terms of both accuracy and speed. Figure 4.1 shows the comparison of the detection results before and after the improvement.

(a)original (b)improved

Figure 4.1 Comparison of detection results before and after improvement

V. CONCLUSION

In this paper, an improved YOLOv5s algorithm is proposed based on YOLOv5s, which is mainly designed to solve the problems of slow detection speed, low accuracy, and large number of model parameters in fire and smoke detection. Based on YOLOv5s, the algorithm uses GhostNet to reconstruct the backbone network to reduce the complexity of the model and lighten the model; secondly, Coordinate Attention is added to the backbone network to improve the detection accuracy of the algorithm, and finally, the original loss function is replaced by EIOU Loss to improve the speed of the bounding box regression while improving the localization accuracy. Based on the existing small dataset, the fire and smoke dataset is self-developed, and experiments are carried out on this dataset to verify the algorithm. The experimental results show that the improved model has obvious advantages, the model size is only 7.6MB, which is 45.3% smaller than the original model; the inference speed reaches 99.4FPS on the RTX3080 device, which is 12.8% higher than the original model; and the mAP reaches 0. 881, which is 2.3% higher than the original model. The real-time requirement is achieved while ensuring the accuracy. In the future, the algorithm will be further improved and deployed to edge computing devices to realize real-time fire detection on mobile.

REFERENCES

[1] Baek Jaeseung, et al."A wavelet-based real-time fire detection algorithm with multi-modeling framework." Expert Systems With Applications 233.(2023).

[2] Bochkovskiy, Alexey , C. Y. Wang , and H. Y. M. Liao . "YOLOv4: Optimal Speed and Accuracy of Object Detection." (2020).

[3] Lin, Tsung Yi , et al. "Feature Pyramid Networks for Object Detection." IEEE Computer Society (2017).

[4] Wang, Wenhai , et al. "Efficient and Accurate Arbitrary-Shaped Text Detection with Pixel Aggregation Network." IEEE (2019).

[5] Han, Kai , et al. "GhostNet: More Features from Cheap Operations." (2019).

[6] Yuchen Zhao, et al."Fire Detection Algorithm Based on an Improved Strategy of YOLOv5 and Flame Threshold Segmentation." 75.3(2023).

[7] Zheng Hongtao, et al."Real-time fire detection algorithms running on small embedded devices based on MobileNetV3 and YOLOv4." Fire Ecology 19.1(2023).

[8] Hou, Qibin , D. Zhou , and J. Feng . "Coordinate Attention for Efficient Mobile Network Design." (2021).

[9] Xu Hao, Li Bo,and Zhong Fei."Light-YOLOv5: A Lightweight Algorithm for Improved YOLOv5 in Complex Fire Scenarios." Applied Sciences 12.23(2022).

[10] Zheng, Zhaohui , et al. "Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression." arXiv (2019).

[11] Redmon, Joseph , and A. Farhadi . "YOLOv3: An Incremental Improvement." arXiv e-prints (2018).

Yuchen Xie, School of Computer Science and Technology, Tiangong University, China.

Zhou Yu, School of Computer Science and Technology, Tiangong University, China.

Mengyuan Liu, School of Computer Science and Technology, Tiangong University, China.