International Journal of Science and Engineering Applications
Volume 13-Issue 10, 124 – 127, 2024, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1310.1025

# Research on a Lightweight Fire Detection Algorithm Based on Improved YOLOv8

Yijian Xu
School of Electronic Information and Electrical Engineering
Yangtze University
Jingzhou, China

Yin Liu
School of Electronic Information and Electrical Engineering
Yangtze University
Jingzhou, China

**Abstract**: In order to enhance the accuracy of fire detection, particularly the ability to detect small fire sources, and to increase the speed of fire detection, this paper proposes an improved fire detection algorithm based on YOLOv8. By leveraging the DualConv to improve the C2f and construct a lightweight structure, and introducing the Slou loss function for small fire sources, the system's accuracy and speed are improved. Experiments were conducted on a custom-built fire dataset, and the results show that compared to the original YOLOv8, the improved model increases the mean Average Precision (mAP@50) by 1.5% and reduces model parameters by 10.3%. This effectively lowers the false alarm rate and enhances the response speed to fires.

**Keywords**: YOLOv8;Fire Detection; Lightweight; DualConv; Slou;

## 1. INTRODUCTION

We ask that authors follow some simple guidelines. This document is a template. An electronic copy can be downloaded from the journal website. For questions on paper guidelines, please contact the conference publications committee as indicated on the conference website. Information about final paper submission is available from the conference website.

According to statistics from China's National Fire and Rescue Department, there were 825,000 fire incidents nationwide in 2022, an increase of 7.8% compared to the previous year, resulting in direct property losses of 7.16 billion yuan. The number of fire-related deaths rose by 1.2%[1]. In the first half of 2023, a total of 550,000 fire incidents were reported across the country, averaging over 3,000 fires per day[2]. Therefore, establishing a timely and accurate fire detection system is crucial. Such a system needs to quickly identify fire sources in the early stages of a fire and provide critical information to fire departments, thereby reducing emergency response times, allowing more time for evacuation and firefighting efforts, and minimizing potential losses.

However, traditional fire detection methods currently rely mainly on sensor technology. This sensor-based fire detection approach has obvious drawbacks. For instance, sensors are easily affected by environmental factors such as smoke, dust, and humidity, which can lead to false alarms or missed detections. The installation and maintenance costs of sensor equipment are relatively high, and for fire detection in large areas or complex environments, the difficulty and cost of deploying sensors increase significantly.

In the rapidly evolving field of artificial intelligence, deep neural networks have become a crucial research direction due to their exceptional performance and wide range of applications. Particularly in fire detection, AI can achieve accurate identification and rapid response to fires by analyzing vast amounts of data and images. This technology not only overcomes limitations of traditional sensor methods, such as susceptibility to environmental interference, but also provides more accurate alerts. However, as the depth and complexity of deep neural networks increase, their computational costs and parameter quantities also rise, posing significant challenges for resource-constrained environments. Therefore, reducing computational overhead and parameter quantities while maintaining accuracy is an urgent issue that researchers need to address.

This paper proposes a fire detection system based on an improved YOLOv8 and introduces an enhancement mechanism that utilizes DualConv[3]to improve C2f, constructing C2f in a lightweight manner. This innovation not only enhances the model's efficiency but also effectively improves its performance. To address the challenge of small and inconspicuous fire sources in the early stages of a fire, the Slou loss function[4] is introduced to enhance small object detection. The model analyzes images through deep learning, enabling rapid identification of fire sources in the early stages of a fire and promptly issuing alerts. This effectively reduces false alarm rates and improves response times to large-scale fires.

## 2. YOLOv8 ALGORITHM

The overall structure of the YOLOv8 algorithm largely continues the design philosophy of YOLOv5, maintaining its efficient and flexible characteristics. However, the most notable change in YOLOv8 is its shift from the traditional Anchor-Based method to an Anchor-Free approach. This method fundamentally alters the operation of the YOLO series, adopting an anchorless approach. It no longer relies on pre-set anchors, instead directly inferring object positions and sizes within the feature map. This transition significantly improves inference speed. Compared to Anchor-Based methods, the Anchor-Free approach simplifies model design, eliminating the need for complex configurations of anchor-related parameters such as scale sizes, aspect ratios, and IoU thresholds. Consequently, YOLOv8 achieves higher computational efficiency while maintaining detection accuracy, opening up new possibilities for real-time object detection applications.

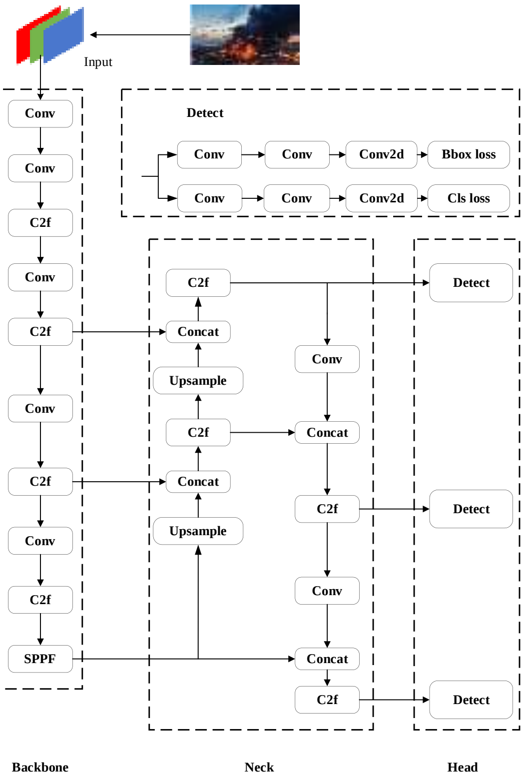


Figure. 1 YOLOv8 Network Architecture

As shown in Figure 1, the design of the YOLOv8 model is primarily divided into three key components: the Backbone network, the Neck network, and the Head network. Among these, CSPDarknet serves as the backbone network, while a decoupled head functions as the output head.

The Backbone plays a crucial role in the construction of YOLOv8, with its structural design determining the overall network's performance in feature extraction. The combination of backbone blocks lays a solid foundation for the network. In each convolutional module, a stride-2 kernel is used for downsampling. This design strategy not only effectively reduces the size of the feature map but also enhances the network's expressive power in feature capture by increasing the number of channels. Each convolutional operation is followed by batch normalization to stabilize the training process and employs the SiLU activation function to enhance the model's nonlinear expression capability. This continuous design strengthens the extraction of detailed information. The C2f module introduces another innovative design. The input data first undergoes processing through a convolutional module with specified parameters, followed by a residual connection module. This residual connection not only facilitates gradient flow but also aims to capture more feature information. The processed result is then concatenated with the output of the main backbone module, producing a richer feature map. The SPPF module merges multiple pooling results and unpooled data by serially using three pooling operation layers, extracting multi-scale features. The core task of the backbone is to gradually reduce the size of the feature map through these different modules and extract deep semantic and feature information from the image, enhancing the model's expressive capability.

The Neck section begins further processing of the rich feature maps extracted by the Backbone, utilizing methods such as feature fusion, upsampling, downsampling, and feature transformation to enhance the precision of target localization and the representation capability of semantic information. This module design significantly strengthens the performance and robustness of the object detection system.

The Head takes on the crucial responsibility of carrying out the actual object detection task. At this stage, the network needs to predict the object's bounding box, class label, and confidence score. An outstanding Head module design means being able to accurately locate and classify objects, providing a strong foundation for the accuracy of the entire detection model.

## 3. IMPROVEMENT METHODS

To enhance detection accuracy and reduce model complexity, this paper proposes improvements to YOLOv8. Figure 2 shows the network model used for fire detection, with the blue and green boxes indicating the improved sections.

### 3.1 C2f-Dual

DualConv is an innovative convolutional network architecture whose core idea is to combine two different sizes of convolutional kernels—3×3 and 1×1—to process the same input feature map channels. Traditional convolutional networks often require a trade-off between accuracy and efficiency, but DualConv cleverly alleviates this issue by optimizing information processing and feature extraction.

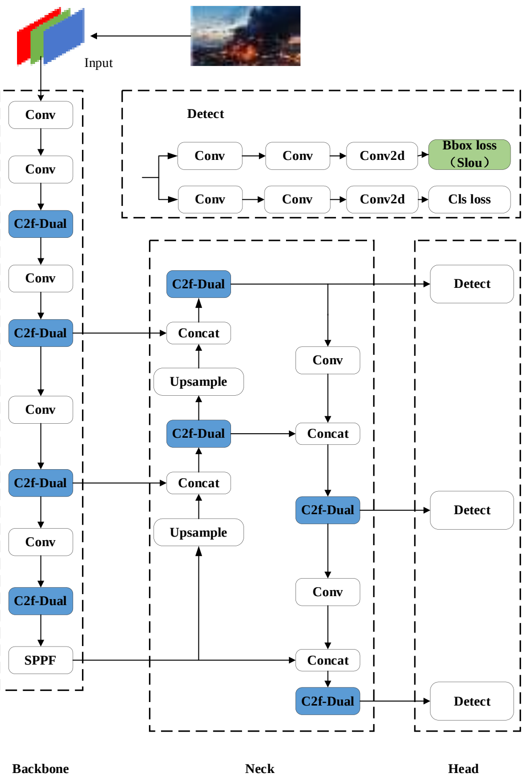


Figure. 2 Improved YOLOv8 Model Overall Structure Diagram

In practical implementation, the 3×3 convolutional kernel is responsible for extracting local spatial features, while the 1×1 convolutional kernel is used to enhance nonlinear feature representation capabilities and reduce the channel dimensions of the feature map. This combination not only allows for more precise feature extraction but also retains more input information, effectively reducing information loss.

DualConv utilizes group convolution techniques to efficiently arrange convolutional filters. Group convolution divides the

International Journal of Science and Engineering Applications
Volume 13-Issue 10, 124 – 127, 2024, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1310.1025

channels of the convolution operation into multiple groups, performing convolution operations separately for each group. This approach not only significantly reduces computational complexity but also decreases the number of parameters. Compared to the full-channel convolution method of standard convolution, this strategy is more flexible and efficient.

This article presents an improved lightweight architecture based on YOLOv8, aimed at enhancing the model's efficiency and accuracy. The improvement leverages an innovative DualConv structure, which optimizes the feature extraction process by combining 3×3 group convolutions and 1×1 pointwise convolutions.

DualConv is introduced to perform convolution operations. This module combines group convolution and pointwise convolution, significantly enhancing information processing capabilities while reducing computational cost and the number of parameters. By integrating these convolution operations, DualConv achieves more efficient feature extraction.

In the Bottleneck module, the standard bottleneck structure is improved by incorporating DualConv to replace traditional convolutional layers. This modification ensures efficient information flow while reducing unnecessary computations by selectively retaining or skipping certain connections.

Furthermore, the C2f-Dual module is designed with a bottleneck structure of a CSP (Cross Stage Partial) network, applying DualConv to the core part of the model. This module provides a more lightweight and efficient path for feature extraction through a series of dual convolutional operation layers, integrating initialization and expansion operations. A comparison of the structures of C2f and C2f-Dual is shown in Figure 3.

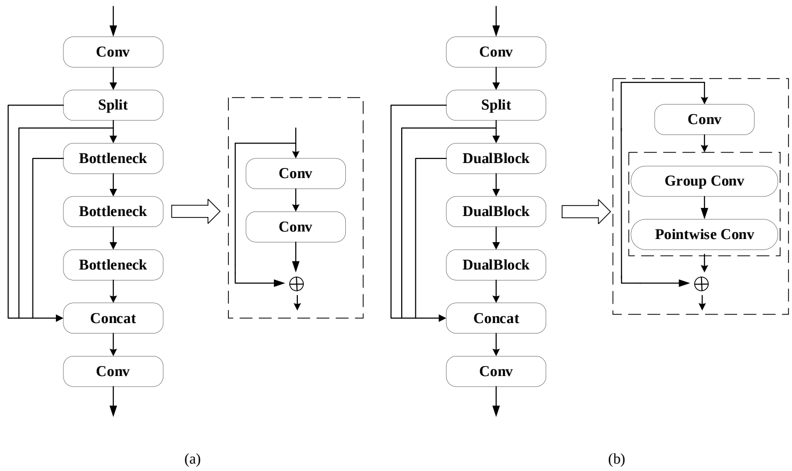


Figure. 3 Comparison of C2f and C2f-Dual: (a) C2f (b) C2f-Dual

## 3.2 Loss Function

The loss functions in the YOLOv8 algorithm primarily include category classification loss and bounding box regression loss. The bounding box regression loss utilizes a combination of DFL (Distribution Focal Loss) and CIoU Loss.

DFL aims to enhance the accuracy of predicted box regression, particularly for more precise predictions of the bounding box coordinates.

CIoU Loss is an improved version of IoU loss that takes into account factors such as position, size, aspect ratio, and distance between bounding boxes to optimize detection boxes more efficiently. The formula is as follows:

$$CIoU = IoU - \frac{\rho^2(b, b_{gt})}{c^2} - \alpha v \quad (1)$$

$$v = \frac{4}{\pi^2}(arctan\frac{w^{gt}}{h^{gt}} - arctan\frac{w}{h})^2 \quad (2)$$

$$\alpha = \frac{v}{(1 - IoU) + v} \quad (3)$$

**IoU** is the ratio that describes the overlap between the target box and the predicted box. $\rho$ is the Euclidean distance between the center point of the predicted box $b$ and the center point of the ground truth box $b_{gt}$. $c$ is the diagonal length of the smallest enclosing area that surrounds both the predicted box and the ground truth box. $\alpha$ is the weight balancing parameter. $v$ is used to measure the aspect ratio consistency between the predicted box and the ground truth box.

To further enhance the detection of small objects, the SIoU loss function is introduced. By focusing on the shape alignment and geometric fitting of bounding boxes, SIoU significantly improves the localization accuracy of small objects. The formula for the SIoU loss function is as follows :

$$SIoU = 1 - IoU + \frac{\Delta + \Omega}{2} \quad (4)$$

# 4. EXPERIMENT AND RESULT ANALYSIS

## 4.1 Experimental Environment and Parameter Settings

The experimental setup utilized a computer system equipped with an Intel Core i5-12400F 2.50 GHz CPU, an NVIDIA GeForce RTX 3060 12 GB GPU, and 32 GB of RAM, running on the Windows 11 Professional operating system. The neural network model was built and trained using the PyTorch 2.3.1 framework, in conjunction with CUDA version 12.1 and Python version 3.11.5. The training parameters for the experiment are as follows: a batch size of 16, 200 epochs, a learning rate of 0.01, and the Adam optimizer, to ensure the reproducibility of the experiment and the accuracy of the results.

## 4.2 Fire Dataset

In this text, Python web scraping technology is utilized to collect fire images and public datasets, creating a new fire dataset. The dataset includes two labels: flames and smoke, and incorporates some unlabeled samples to enhance the model's generalization capability. LabelImg is used to annotate the images, and the results are stored in .txt files.

The fire dataset contains a total of 10,461 images. They are divided in a ratio of 7:2:1, resulting in 7,510 images for the training set, 1,878 images for the validation set, and 1,073 images for the test set.

## 4.3 Evaluation Metrics

This experiment evaluates the fire detection performance of the model using recall (R), precision (P), and mean Average Precision (mAP@50). Higher values indicate better overall performance. Additionally, the model's complexity and detection speed are assessed through the number of parameters and floating-point operations per second (FLOPS).

a. Precision (P) represents the proportion of true positive samples among all samples predicted as positive.

$$P = \frac{TP}{TP + FP} \tag{5}$$

b. Recall (R) represents the proportion of actual positive samples that are correctly identified as positive samples.

$$R = \frac{TP}{TP + FN} \tag{6}$$

c. Mean Average Precision (mAP@50) measures the accuracy of a model by calculating the area under the precision-recall curve for each category.

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \tag{7}$$

In this experiment, TP represents the number of samples correctly predicted as fire; FP indicates the number of non-fire samples incorrectly judged as fire, reflecting the number of false detections; FN signifies the number of fire samples incorrectly judged as non-fire. mAP@50 refers to the mean average precision calculated for all categories when the Intersection over Union (IoU) reaches 50%. It is used to evaluate the overall performance of the model's detection and recognition accuracy for various classification targets at this threshold.

## 4.4 Experimental Results Analysis

Compared to YOLOv8, this model improves precision by 2%, recall by 2.3%, and mAP@50 by 1.5%. The model also features reduced complexity and faster detection speed, with a reduction of 0.31M parameters and 0.7 GFLOPS in computational load. Detailed comparison results are shown in Table 1.

Table 1. Comparison Results

| Models | Precision | Recall | mAP@50 | FLOPS(G) | Parameter (M) |
|---|---|---|---|---|---|
| YOLOv8 | 0.803 | 0.755 | 0.821 | 8.2 | 3.02 |
| Ours | 0.823 | 0.778 | 0.836 | 7.5 | 2.71 |

# 5. CONCLUSION

This paper presents a fire detection algorithm based on an improved YOLOv8, achieving enhancements in both performance and efficiency of the model[5][6].

1) Enhancement of the C2f structure: By introducing DualConv, the C2f module is optimized. This enhancement not only improves the model's performance but also increases efficiency through lightweight design. The improvement in the C2f module allows the network to maintain high detection accuracy while offering better real-time performance.

2) Introduction of the Slou loss function: Addressing the challenge of detecting small targets in fire detection, this paper proposes and introduces the Slou loss function. This function boosts the detection accuracy of small targets and enhances the model's adaptability in complex environments.

3) Performance boost: Compared to the original YOLOv8 model, the method proposed in this paper improves accuracy by 2%, increases recall by 2.3%, and raises mAP@50 by 1.5%. These improvements indicate that the model can achieve better detection outcomes while maintaining a lightweight design.

4) Computational efficiency: By reducing the parameter count by 0.31M and the computational load by 0.7 GFLOPS, the improved model significantly decreases computational complexity. This not only enhances speed but also reduces hardware resource requirements, making the model more suitable for applications in resource-constrained environments.

# 6. REFERENCES

[1] China National Fire and Rescue Administration. National Police and Fire Incident Statistics for 2022[Online].(2023-03-24)[2024-09-06]. https://www.119.gov.cn/qmxfxw/xfyw/2023/36210.shtml.

[2] China National Fire and Rescue Administration. Over 3,000 Daily Fire Incidents Reported Nationwide in First Half of 2023 [Online]. (2023-07-17)[2024-09-06]. https://www.119.gov.cn/qmxfgk/sjtj/2023/38420.shtml,2023-07-17/2024-6-7.

[3] Zhong J, Chen J, Mian A. DualConv: Dual convolutional kernels for lightweight deep neural networks[J].IEEE Transactions on Neural Networks and Learning Systems, 2022, 34(11): 9528-9535.

[4] Gevorgyan Z. SIoU loss: More powerful learning for bounding box regression[J]. arXiv preprint arXiv:2205.12740, 2022.

[5] Wang G, Chen Y, An P, et al. UAV-YOLOv8: A small-object-detection model based on improved YOLOv8 for UAV aerial photography scenarios[J]. Sensors, 2023, 23(16): 7190.

[6] Chen T H, Wu P H, Chiou Y C. An early fire-detection method based on image processing[C]//2004 International Conference on Image Processing, 2004. ICIP'04. IEEE, 2004, 3: 1707-1710.