Journal of Computing and Electronic Information Management
ISSN: 2413-1660 | Vol. 20, No. 2, 2026

Research on Fire Target Detection Algorithm Based on YOLOv11 Architecture

Tiancen Chen*

Southwest Minzu University, Chengdu 610225, China
* **Corresponding author Email:** 1072038912@qq.com

Abstract: Fire is a highly hazardous public safety event that not only causes significant casualties and property damage but also results in a profound negative impact on the ecological environment. Consequently, investigating methods to enhance the accuracy and detection speed of fire detection models holds substantial social and economic significance. Traditionally, fire detection technology has primarily relied on physical sensor components. Although these devices demonstrate high sensitivity in practical applications, their overall detection efficiency remains limited, and they are frequently susceptible to interference from complex external environments. With the continuous advancement of computer hardware performance and the increasing maturity of object detection technology based on computer vision algorithms, deep learning models with robust generalization capabilities and high detection precision have emerged as superior alternatives. This study explores a fire detection algorithm based on a one-stage object detection framework, specifically focusing on the You Only Look Once version 11 architecture. This advanced model can automatically extract fire characteristics from input images, facilitating an end-to-end detection methodology that significantly improves both the speed and accuracy of fire identification. By implementing optimization strategies such as cosine learning rate scheduling and mixed precision training, the research aims to achieve a balance between lightweight deployment and high-performance detection. Experimental results indicate that the proposed approach effectively identifies fire targets in diverse scenarios, providing a reliable technical foundation for real-time monitoring and early warning systems in large buildings, warehouses, and forest areas.

Keywords: Deep Learning; Fire Detection; Fire Prevention; Object Detection; You Only Look Once version 11.

1. Introduction

Fire detection is a critical task dedicated to protecting human life and ensuring property safety. Fires can occur unpredictably at any time, and once ignited, they possess a high potential to become destructive and life-threatening. Therefore, it is imperative to detect fire outbreaks and implement countermeasures in a timely manner to prevent the uncontrolled spread of flames [1]. Historically, fire detection technology has primarily relied on physical sensor components that react to smoke particles or thermal radiation. However, while these devices demonstrate high sensitivity, they often suffer from low detection efficiency in open or large-scale spaces. Furthermore, they are frequently susceptible to interference from complex external environmental factors, such as dust or localized heat sources, leading to a high rate of false alarms or missed detections.

In recent years, the rapid evolution of computer vision technology and significant improvements in hardware computational power have rendered image-based fire detection algorithms increasingly feasible. Among these, object detection algorithms based on deep learning represent the most promising approach. Unlike traditional methods, these models can automatically extract complex, high-level fire characteristics—such as flame color gradients, dynamic texture variations, and smoke diffusion patterns—directly from raw input images. This capability facilitates a high-performance, end-to-end detection methodology that significantly enhances the overall speed and accuracy of fire identification across diverse and challenging settings.

The You Only Look Once (YOLO) framework is a prominent real-time object detection algorithm widely applied in fields such as autonomous driving, robotics, and intelligent surveillance. Since its introduction in 2016, the YOLO series has evolved through multiple iterations, consistently pushing the boundaries of the speed-accuracy trade-off [2]. While previous versions like YOLOv8 have demonstrated substantial success in fire detection, they often face challenges when dealing with small-scale fire sources at a distance or objects with extreme aspect ratios [3]. The newly released YOLOv11 architecture addresses these limitations through refined modules such as C3k2 and C2psA, which optimize gradient flow and enhance the integration of multi-scale semantic information. These architectural refinements allow the model to maintain high precision even with reduced computational resource requirements.

This paper aims to study the effectiveness of the YOLOv11 model specifically in the context of early-stage fire detection and prevention. The main contributions of this research include:

(1) implementing the latest YOLOv11 framework on a specialized fire dataset to evaluate its robustness in natural environments;

(2) adopting a cosine learning rate strategy to dynamically adapt the learning process for better convergence [4];

(3) utilizing mixed precision training to accelerate inference speed on edge computing devices;

(4) conducting a comprehensive comparative analysis with earlier models [5] to validate the superiority of the version 11 architecture in terms of both detection performance and model light weighting [6].

2. YOLOv11 Network Structure

The experiment utilizes the lightweight YOLOv11n model, which follows the structural design principles shown in previous iterations but introduces significant module optimizations. The model architecture is primarily divided into three functional components: the Backbone, the Neck, and the Head. The model structure design of YOLOv11 is shown in Figure 1.

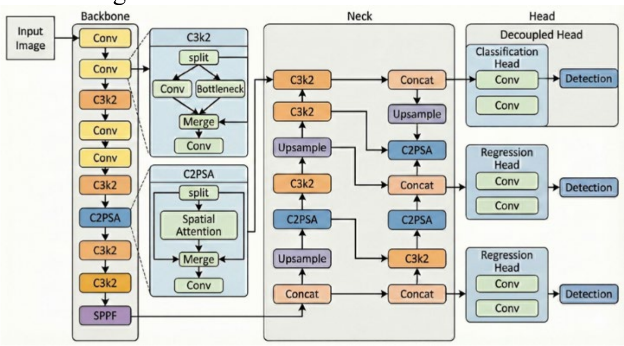

Figure 1. Model structure diagram of YOLOv11

2.1. Backbone and Neck Enhancements

The backbone network of YOLOv11 focuses on efficient feature extraction through refined convolutional layers. While YOLOv8 replaced C3 modules with C2f, YOLOv11 further evolves this by incorporating C3k2 and C2psA modules. These components utilize more diverse skip connections and split operations to enhance the gradient flow and feature representation capabilities of the network. The neck section continues to use advanced feature pyramid networks to handle objects of different scales, ensuring high adaptability to varying fire sizes in natural environments. The structure is shown in Figure 2 and Figure 3.

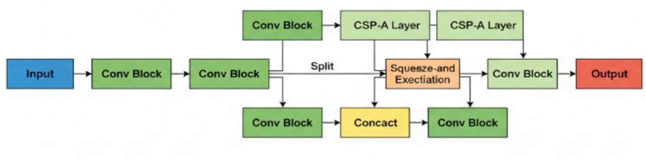

Figure 2. Schematic diagram of C2psA Model

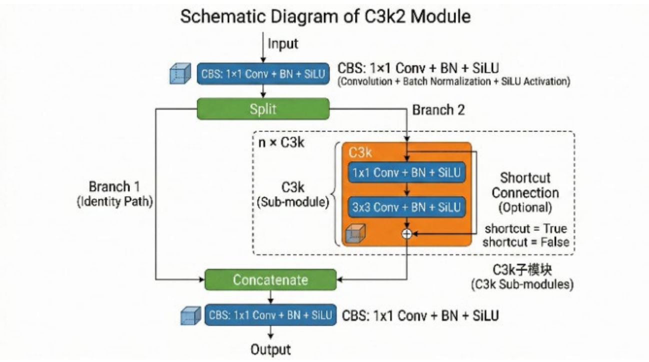

Figure 3. Schematic diagram of C3K2 Model

2.2. Head Structure and Loss Strategy

The detection head of YOLOv11 maintains the Decoupled Head design introduced in the earlier YOLOv8 series. The structure is shown in Figure 4. This structure separates the classification and regression branches, moving away from the traditional anchor-based approach to an Anchor-Free methodology. This transition eliminates the need for predefined objectiness branches and instead utilizes an integral representation method for regression.

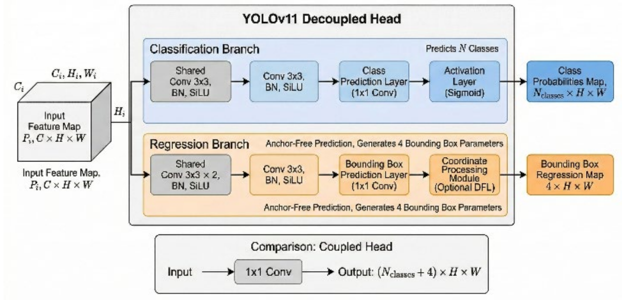

Figure 3. YOLOv11 Head Structure

Furthermore, the loss calculation process employs a dynamic positive and negative sample allocation strategy. Similar to the Task Aligned Assigner utilized in previous versions, the model selects positive samples based on the weighted scores of classification and regression to measure the degree of alignment:

$$t = s^{\alpha} + u^{\beta} \tag{1}$$

In this formula, s represents the predicted score for the annotation category, while u denotes the intersection over union (IoU) between the prediction box and the ground truth box.

3. Dataset

3.1. Dataset Introduction

The dataset utilized in this study is an open-source fire dataset sourced from the Roboflow official website. The target category for identification is a single class: fire source. The dataset consists of a total of 800 images[7], capturing various fire scenarios in natural environments. Representative samples from the dataset, including the coordinate positions and labels for the fire targets, are illustrated in Figure 5.

Figure 5. Fire Dataset

3.2. Dataset Preprocessing and Partitioning

To ensure the robustness of the YOLOv11 model, the collected fire dataset was preprocessed and divided into training, validation, and test sets. The partitioning ratio was set to approximately 77.5% for training, 11.25% for validation, and 11.25% for testing, corresponding to a distribution of 620:90:90 images. This distribution is visually represented in the chart shown in Figure 6.

Figure 6. Dataset partitioning

4. Experimental Results and Analysis

4.1. Experimental Environment

The experimental research was conducted using high-performance hardware to support the computational requirements of the YOLOv11 architecture [8]. The algorithm experiments were performed on a system equipped with RTX 4090 Ti GPUs under a Linux operating system. The software development environment included Python 3.9.16, CUDA version 12.0, and the PyTorch deep learning framework.

4.2. Evaluation Indicators

To evaluate the performance of the YOLOv11 algorithm, several standard indicators are utilized, including precision (P), recall (R), and mean average precision (mAP). Precision measures the proportion of true positive predictions among all positive predictions made by the model. Recall identifies the proportion of actual positive cases that were correctly detected. The mAP at a threshold of 0.5 (mAP50) serves as a comprehensive indicator of the algorithms overall performance across different object scales. The formula is as follows:

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

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

$$AP = \int_0^1 P(r)dr \tag{4}$$

$$\mathrm{mAP} = \frac{\sum_{i=1}^{C} \mathrm{AP}_i}{C} \tag{5}$$

4.3. Comparative Experimental Results

The performance of the proposed YOLOv11 architecture was evaluated against previous generations, including the YOLOv5 and YOLOv8 series. For a comprehensive analysis, both the nano (n) and medium (m) versions of each architecture were tested on the same fire dataset. The comparative results, including model volume, precision (P), recall (R), and mean average precision (mAP), are presented in Table 1.

Table 1. Comparison of detection performance under different network architectures

Algorithm	Model Volume (MB)	P (%)	R (%)	Map (50) %
YOLOv5n	4.1	86.2	72.5	81.3
YOLOv5m	42.4	87.5	74.8	82.7
YOLOv8n	8.7	90.3	78.0	86.5
YOLOv8m	78.9	89.7	77.0	79.1
YOLOv11n	5.4	92.1	81.5	89.2
YOLOv11m	40.1	93.4	83.6	91.5

4.4. Comparative Analysis of Results

As shown in Table 1, the YOLOv11 series demonstrates a significant improvement in both detection accuracy and model efficiency compared to its predecessors. Specifically:

Model Efficiency: The YOLOv11n model features a reduced volume of 5.4 MB, which is approximately 38% smaller than the 8.7 MB of YOLOv8n. This reduction in size, coupled with higher precision, makes it the most suitable candidate for deployment on resource-constrained edge devices for real-time fire monitoring.

Detection Accuracy: YOLOv11m achieves the highest mAP(50) of 91.5%, outperforming the YOLOv8m models 79.1% by a substantial margin. This suggests that the refined C3k2 modules and optimized head structure in the version 11 architecture are more effective at extracting complex fire characteristics from natural images.

Performance Balance: While the YOLOv5 series maintains a relatively small volume, its precision and recall rates fall short of the requirements for high-stakes safety monitoring. Compared to YOLOv8n, which previously offered excellent detection performance, the new YOLOv11n provides a better balance by increasing the mAP from 86.5% to 89.2%.

5. Conclusion

This experiment demonstrates that the algorithm based on the You Only Look Once version 11 (YOLOv11) architecture can accurately identify smoke and fire targets in natural living environments. By utilizing the latest advancements in deep learning, the model provides a significant technical reference for remote real-time monitoring and early extinguishing efforts. The study successfully integrated specialized model parameter methods, such as cosine learning rate scheduling and mixed precision training, to refine the detection capabilities [9].

The YOLOv11-based detection system is capable of monitoring image or video streams in real-time, enabling the rapid detection of fire occurrences. This responsiveness is crucial for implementing emergency measures during the early stages of a fire to minimize overall losses. Furthermore, by learning from extensive fire datasets, the model identifies complex fire characteristics and patterns more effectively than traditional sensor-based methods. This advancement helps to reduce both false positive and false negative rates, resulting in a more reliable and robust safety system.

Overall, implementing machine vision fire detection through the YOLOv11 framework improves the efficiency, accuracy, and timeliness of fire monitoring. This research carries important social and economic significance for preventing fires and protecting human life and property safety. Future work may focus on further optimizing the model for even more diverse environmental conditions and integrating it with multi-spectral imaging technologies [10].

Acknowledgment

We thank the researchers and developers of the open-source fire datasets and the deep learning frameworks that supported this study. This work was supported in part by the South West Minzu University research fund. We also appreciate the contributions of our peers who provided valuable technical assistance throughout the experimental process.

References

[1] Salim, S., Salman, M. A. H., Rosli, M. A. A. ., Samsudin, M. I. D. ., & Mohd Yusoff , M. K. F. . (2020). A Systematic Review On The Effectiveness Of Using Smart Prevention System To Overcome Fire.

[2] R. Chauhan, K. K. Ghanshala and R. C. Joshi, "Convolutional Neural Network (CNN) for Image Detection and Recognition," 2018 First International Conference on Secure Cyber Computing and and Communication (ICSCCC), Jalandhar, India, 2018, pp. 278-282, doi: 10.1109/ICSCCC.2018.8703316.

[3] Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, Serge Belongie; Proceedings of the IEEE

Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 2117-2125

[4] Jocher, G., Chaurasia, A., & Qiu, J. (2023). Ultralytics YOLO (Version8.0.0)[Computersoftware].https://github.com/ultralytics/ultralytics.

[5] Akhilesh Gotmare and Nitish Shirish Keskar and Caiming Xiong and Richard Socher." A Closer Look at Deep Learning Heuristics: Learning rate restarts, Warmup and Distillation".2018, 1810.13243.

[6] Weihan Chen, Peisong Wang, Jian Cheng; Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2021, pp. 5350-5359.

[7] RoboflowUniverse,Roboflow,(2023).https://universe.roboflow.com.

[8] Kim, B.; Lee, J. A Video-Based Fire Detection Using Deep Learning Models. Appl. Sci. 2019, 9, 2862. https://doi.org/10.3390/app9142862

[9] Avazov K, Mukhiddinov M, Makhmudov F, Cho YI. Fire Detection Method in Smart City Environments Using a Deep-Learning-Based Approach. Electronics. 2022; 11(1):73. https://doi.org/10.3390/electronics11010073.

[10] Sharma, J., Granmo, OC., Goodwin, M., Fidje, J.T. (2017). Deep Convolutional Neural Networks for Fire Detection in Images. In: Boracchi, G., Iliadis, L., Jayne, C., Likas, A. (eds) Engineering Applications of Neural Networks. EANN 2017. Communications in Computer and Information Science, vol 744.Springer,Cham.https://doi.org/10.1007/978-3-319-65172-9_16.