IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 9 August 2024, accepted 19 August 2024, date of publication 23 August 2024, date of current version 2 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3448608

RESEARCH ARTICLE

Research on Fire Smoke Detection Algorithm Based on Improved YOLOv8

TIANXIN ZHANG, FUWEI WANG, WEIMIN WANG, QIHAO ZHAO, WEIJUN NING, AND HAODONG WU
School of Artificial Intelligence and Software, Liaoning Petrochemical University, Fushun 113005, China

Corresponding author: Fuwei Wang (wangfw4048@lnpu.edu.cn)

This work was supported by Liaoning Provincial Department of Education Youth Project for Research on Record Linking Technology Based on Privacy Protection (LJKQZ20222440).)

ABSTRACT Fire has consistently posed a significant disaster risk worldwide. Current fire detection methods primarily rely on traditional physical sensors such as light, smoke, and temperature detectors, which often struggle in complex environments. The susceptibility of existing fire detection technologies to background interference frequently results in false alarms, missed detections, and low detection accuracy. To address these issues, this paper proposes a fire detection algorithm based on an improved YOLOv8 model. First, to enhance the detection capabilities for large-scale fire and smoke targets, a large target detection head is added to the backbone of the YOLOv8 model. This modification enhances the network's receptive field, allowing it to capture a broader range of contextual information and identify fires over extensive areas. Secondly, an efficient multi-scale attention mechanism, EMA (Efficient Multi-Scale Attention Module), based on cross-space learning is integrated into the FPN (Feature Pyramid Network) part of the model. This mechanism highlights target features while suppressing background interference. Additionally, a PAN-Bag (Path Aggregation Network Bag) structure is proposed to help the model more accurately detect objects such as fire and smoke, which have uneven feature distributions and variable morphologies. With these improvements, we introduce the YOLOv8-FEP algorithm, which offers higher detection accuracy. Experimental results demonstrate that the YOLOv8-FEP algorithm improves the mAP by 3.1% and the accuracy by 5.8% compared to the original YOLOv8 algorithm, proving the effectiveness of the enhanced algorithm.

INDEX TERMS Fire detection, YOLOv8, EMA, PAN-Bag.

I. INTRODUCTION

Fire is a natural and man-made disaster worldwide that presents a serious risk to human civilization. The escalation of urbanization and the severity of climate change have resulted in an increase in the frequency of fire incidents and the resulting losses, profoundly influencing social development. First of all, the most obvious hazards are those that cause death or property damage. Fires have the ability to spread quickly, endangering lives and destroying structures. Second, toxic smoke and copious volumes of greenhouse gases released by fires exacerbate air pollution, lower air quality, and harm the environment over time. Because they upset the natural equilibrium and have an impact Earth's carbon cycle, forest fires are also a major threat to biodiversity. The majority of sensor elements used in fire detection technologies today, such as smoke and thermistors, have poor detection efficiency and are highly susceptible to complicated external conditions, which eventually reduces detection accuracy. Consequently, one of the main areas of current research is the application of powerful computer vision technologies to automatic fire detection.

Most of the traditional fire detection methods are based on rules or feature engineering, and they are often difficult to achieve satisfactory recognition results for complex and changing fire scenes. In contrast, by automatically identifying intricate details in images, deep learning models greatly increase the accuracy and robustness of fire detection.

The associate editor coordinating the review of this manuscript and approving it for publication was Kashif Sharif.

Two-stage detection and single-stage detection are the two primary method types in the field of deep learning target identification. Two-stage detection techniques, like the R-CNN [1] series, Fast R-CNN [2], Faster R-CNN [3], etc., create a set of candidate regions first and then perform bounding-box regression and classification on these regions. This approach produces higher accuracy results, but it is slower and requires more computing power. Regarding single-stage detection, which is represented by SSD [4], YOLO [5] series, which includes YOLOv1 through YOLOv8 [6], [7], they do away with the laborious process of generating candidate regions and proceed directly to the completion of object classification and location prediction simultaneously in a network, greatly increasing the speed at which a detection is made. While the accuracy of the early versions in certain scenarios is marginally lower than that of the two-stage methods, their performance has improved greatly with continuous algorithmic iteration.

The most recent algorithm model in the YOLO series, YOLOv8, is used as the foundation model in this work. Thanks to its quick and effective performance, it can effectively meet the need for real-time detection. When it comes to general object detection, YOLOv8 has no issues and is clearly superior to competing methods. Nonetheless, the algorithm's shortcomings in identifying intricate scenes persist, such as its limited detection range compared to the original model, susceptibility to background interference, and low detection accuracy. This study proposes a fire detection algorithm based on YOLOv8-FEP, with the following primary modifications to overcome the aforementioned issues:

- A large target detection header is added to the backbone network part to increase the sensory field of the network, which facilitates the model to detect fires at different scales, e.g., to identify smoke, fire light, or fire spread in the distance.
- The EMA (Efficient Multi-Scale Attention Module) attention mechanism is introduced in the FPN (Feature Pyramid Network) part to make the model focus more on the target area location information to improve the detection accuracy of the target area, which helps the model to accurately identify the fire signs, such as fire light, smoke, and other features in the complex background, and reduces the false alarm rate and missed alarm rate.
- Meanwhile, the PAN-Bag (Path Aggregation Network Bag) module is incorporated into the FPN part. It enhances the model's detection performance in complex scenarios and is able to integrate local details and global contextual information more effectively.

The rest of the paper is structured as follows: Section II discusses the current state of research and application of fire detection algorithms in the field of deep learning; Section III details an improved approach based on the YOLOv8 fire detection model; Section IV elaborates on the experimental design and the analysis of the results; and section V summarizes the paper and outlines future research directions.

II. RELATED WORK

In recent years, thanks to advances in various artificial intelligence fields, certain results have been achieved in vision-based research areas such as image processing and computer vision. These results are widely used in fire smoke detection. The first proposed algorithm is a smoke and fire detection algorithm based on a target classification model. Shen et al. [8] were among the pioneers in proposing a deep learning-based approach for fire detection using Convolutional Neural Networks (CNN) in smoke detection. Ba et al. [9] introduced 'SmokeNet,' a novel Convolutional Neural Network (CNN) model that enhances visual image classification by integrating spatial and temporal attention mechanisms. Yuan et al. [10] developed deep multi-scale neural networks with a feature extraction layer comprising multiple parallel convolutional layers to achieve high accuracy. Myeongho et al. [11] proposed a multi-scale fire prediction framework based on convolutional neural networks, where the feature maps at each scale predict fires by feature squeezing blocks and softmax functions to determine the final result. Jia et al. [12] put forth a research that advocated for a combined strategy utilizing both Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) network in the task of fire detection.

Since the target classification algorithm can only determine whether there is a fire but cannot locate the fire, a smoke fire detection algorithm based on target detection has been proposed. The two-stage detection methods include Fast R-CNN, Faster R-CNN, etc. Chaoxia et al. [13] proposed an improved Faster R-CNN fire detection method, including a color-guided anchoring strategy and embedded global information-guided fire detection; Zhang et al. [14] used Faster R-CNN for recognition and detection of smoke images, but the use of FPN for feature fusion process is prone to information attenuation, and there is also the effect of aliasing in cross-scale fusion, the model is more complex, and the detection speed is slow. Therefore, smoke fire detection algorithms based on target detection models are beginning to be oriented towards efficient end-to-end, one-stage target detection algorithms. Single-stage detection methods include the SSD and YOLO series. Qin et al. [15] proposed a fire detection method combining the classification model and target detection model, using separable convolution to classify the fire image first, and then using YOLOv3 to judge the location of the fire in the image, which not only improves the detection efficiency but also avoids the problem of degradation of detection accuracy caused by directly using YOLOv3 for classification and location judgment. Zhao et al. [16] proposed Fire-YOLO, an improved fire detection algorithm based on YOLOv3, which uses EficientNet to extract the features of the input image, which facilitates the feature learning of the model, improves the performance of the network, and optimizes the detection process of the YOLOV3 model for very small targets. Liu et al. [17] used the improved YOLOv5n model for forest fire detection, which can be deployed on low-power devices. Li et al. [18] improved the YOLOv5 algorithm by replacing the SPPF module with RFB, which enabled the improved model to more accurately extract information about fire and smoke in the forest and reduced the model's misdetection and omission rates. Wu et al. [19] improved the YOLOv4 algorithm and proposed a YOLOv4-minor based fire detection algorithm for ships by considering the characteristics of fires on ships as well as the unique characteristics of the ocean environment. Al-Smadi et al. [20] introduced an innovative framework designed to mitigate sensitivity issues in various YOLO object detection models, thereby enhancing the accuracy of smoke detection. Xue et al. [21] proposed a small-object forest fire detection model based on improved YOLOv5. The researchers added the CBAM attention mechanism to the YOLOv5 model, which effectively solved the problem of information loss caused by the small number of pixels of small objects in forest fires. Chen et al. [22] proposes a multi-modal dataset combining RGB and thermal dual-feed video and a deep learning-based fire detection method to improve the accuracy of early wildland fire detection and assessment.

III. METHODS

The addition of EMA and PAN-Bag modules to the FPN as well as a big target detection head to the backbone network are the suggested improvement points in this paper. Figure 1 illustrates the structure of the updated algorithm model.

A. LARGE TARGET DETECTION HEADS

Although the original backbone network of YOLOv8 already has the ability of multi-scale feature extraction, which can effectively deal with the detection of objects of different sizes, YOLOv8 still suffers from issues such as an excessively small detection range and a tendency toward duplicate recognition, for the special needs of fire scenarios, especially for the recognition of smoke at long distances and the recognition of a wide fire field. Because of this, we specifically improve the YOLOv8 backbone network topology by including a big target detection head, a 10×10 feature layer (P6 layer). An additional downsampling layer can greatly increase the network's perceptual range and improve the multi-scale feature pyramid. This multi-scale feature fusion not only strengthens the model's target capture capability at large scales but also promotes deeper multi-scale feature fusion, ensuring that the model strikes a balance between microscopic details (e.g., localized fire sources) and macroscopic scene comprehension (e.g., fire spread, environmental interactions). Of particular importance, the lower resolution large target detection head provides the necessary efficiency gains for real-time fire monitoring systems while ensuring that the model captures large-scale fire features. This optimization is particularly critical in a wide range of monitoring scenarios, such as forests, industrial areas, or urban surveillance, where immediate response is required to accelerate the early identification of fires and the triggering of alarms, which buys valuable time for emergency response to mitigate the damage caused by fires.

B. ATTENTION MECHANISM

There are many interfering factors in fire scene detection. For example, lighting changes that may lead to uneven brightness distribution of the image, affecting the contrast between the smoke and fire target and the surrounding background. Secondly, the surrounding environment may present an intricate scene, with many buildings, roads, vehicles, crowds, and other various background elements, which will introduce a lot of redundant information and interference in the image. Due to the excessive repetition and interference introduced by these backdrop features, it will be challenging to distinguish the pyrotechnic target from the complicated background. The basic principle of the attention mechanism is to suppress useless feature information while strengthening useful feature information so that the model can focus on the important regions in the image more adaptively. The traditional CBAM (Convolutional Block Attention Module) [23] emphasizes the important channels and spatial locations in the feature map to improve the expressive ability of convolutional neural networks, but can only effectively capture local information, making it difficult to establish long-distance channel dependence; CA (Channel Attention) [24] focuses on the channel dimension of the feature map to achieve the reinforcement of important channels, but it ignores the importance of the interactions between the entire spatial location.

To address the YOLOv8 model's susceptibility to misdetection and omission in complex fire scenarios, we introduce the state-of-the-art EMA (Efficient Multi-Scale Attention Module) [25] attention mechanism. The EMA attention mechanism proposes a new cross-spatial learning method and is designed to establish short-term and long-term dependency relationships in a multi-scale parallel sub-network. It also considers a general approach to reshaping some of the channel dimensions into batch dimensions and grouping them into multiple sub-features to ensure an even distribution of spatial semantic features within each feature group. As a result, the EMA can weight the feature channels to highlight important features relevant to fire detection and suppress other irrelevant and distracting information in the image. This selective enhancement helps the model accurately identify fire signs, such as fire, smoke, and other features, in complex environmental contexts. Meanwhile, compared with CBAM, SA, and CA, EMA not only has higher performance but also is more efficient in terms of required parameters, so we chose to introduce the EMA module in this paper.

By introducing the EMA module into the FPN part of YOLOv8, the model focuses more on the location information of the target area, in order to improve the detection accuracy of the target area and reduce the false

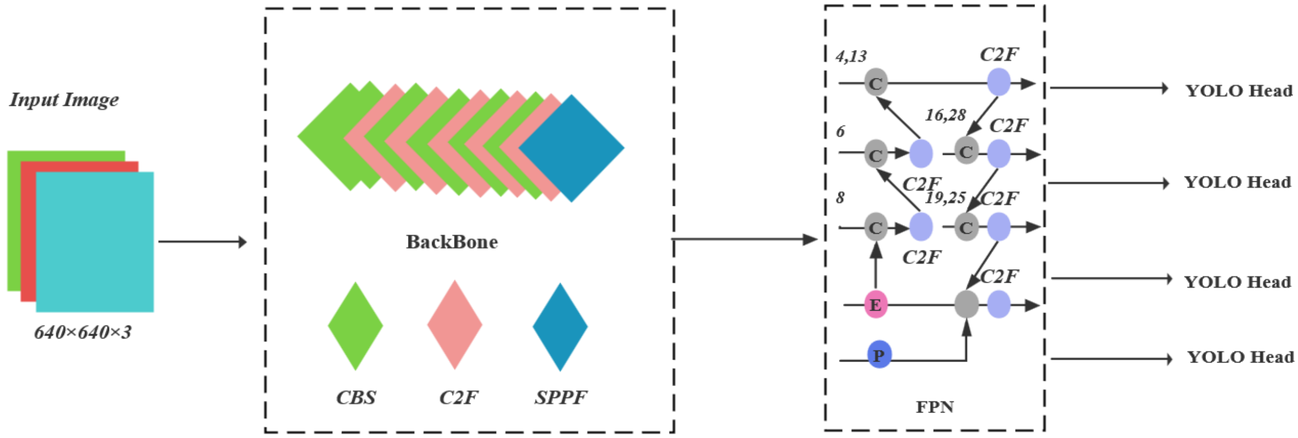

FIGURE 1. Structure of the YOLOv8-FEP network.

alarm and missed alarm rates. The model diagram of EMA is shown in Figure 2.

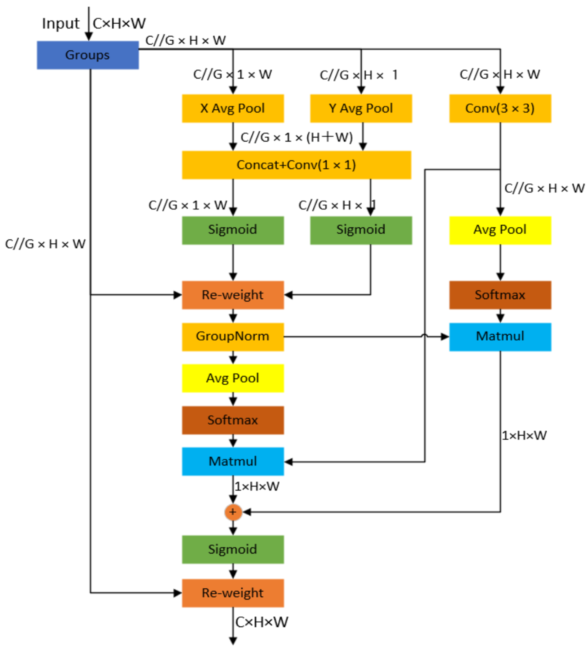

FIGURE 2. The architecture of the EMA.

C. PAN-BAG

The issue with the popular PAN (Path Aggregation Network) [26] structure is that target properties for smoke and fire cannot be effectively extracted. For this reason, we designed the PAN-Bag module. The PAN-Bag is an improvement to the classical PAN structure, which improves the model's capability of extracting complex target features, especially for challenging detection tasks such as fire and smoke, by introducing a feature fusion strategy at multiple levels. The PAN-Bag module upgrades the PAN structure by means of a well-designed ProgressiveConvUnit, which is embedded with a sequence of highly optimized operations, including a 1×1 convolution for channel compression and feature recalibration, This is followed by two 3×3 convolutions for spatial feature extraction and enhancement, and finally the initial features are fused with the transformed features by the element summing operation, which preserves the original information and incorporates rich contextual knowledge. This kind of design guarantees the effective fusion and conveyance of information in addition to enhancing the expressiveness of the features.

In the PAN-Bag module, we first apply progressive convolutional units to the feature maps of the P3 layer (80×80) to enhance their feature representations, while maintaining the feature map size unchanged. This lays a solid foundation for subsequent fusion. Subsequently, each feature map (sizes: 80×80, 40×40, 20×20, 10×10) is halved by level-by-level average pooling operations, aligning the size with the next smaller scale feature map in preparation for splicing. The splicing operation not only fuses features at different scales but also enriches the lower-level features by introducing semantic information from the higher-level features. After each splicing, the progressive convolution unit is applied again to further refine the fused features. This process is executed layer-by-layer from P4 to P6, ensuring that each level of features is meticulously processed and deeply fused with cross-scale information.

In this way, the PAN-Bag not only optimizes cross-scale communication between feature maps but also strengthens the recognition of features with significant scale variations, such as fire and smoke, through progressive convolutional units before each fusion. This design improves the accuracy of the model in recognizing fire signs in complex backgrounds, reduces false detections and omissions through an efficient information integration strategy, and enhances the robustness and practicality of the model. The structure of the PAN-Bag and the progressive convolutional units are shown in Figures 3 and 4.

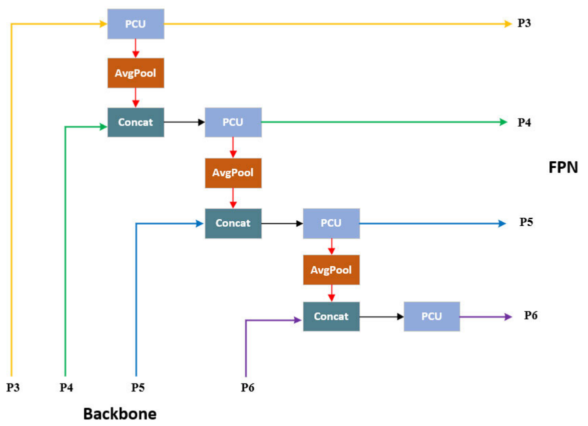

FIGURE 3. The architecture of the PAN-Bag.

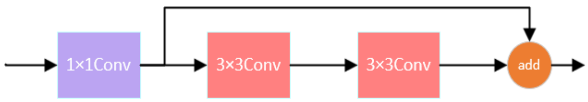

FIGURE 4. The architecture of the ProgressiveConvUnit.

IV. DATASET AND EXPERIMENTAL SETTING

A. FIRE DATASET

Due to the limited availability of opensource datasets for fire and the lack of annotation in most of the existing datasets, we constructed a fire dataset of our own. The dataset for this experiment consists of 4402 images obtained through online searching as well as a selection of fire and smoke images from a number of public datasets. We use the LabelImg tool to manually label, including drawing bounding boxes and classification categories. Due to the lack of fixed shapes of fire and smoke, which are susceptible to human factors during the labeling process, the maximum bounding rectangle labeling method is used in this experiment to minimize the background information of the target. Taking into account the correspondence between labels and data, to ensure that the dataset is evenly distributed, the dataset is randomly divided into the training set, validation set, and test set according to the ratio of 80%, 10%, and 10%, including 3521 training sets, 440 validation sets, and 441 test sets. The final dataset is stored in the PASCAL VOC dataset format. The dataset includes scenes of forest fires, factory fires, vehicle fires, building fires, outdoor fires, and torch flames. Part of the dataset is shown in Figure 5.

B. EXPERIMENTAL ENVIRONMENT AND PARAMETER SETTING

The experimental environment of this paper: operating system: Windows 11; GPU: NVIDIA RTX 2080; the deep learning framework is PyTorch; the programming language is python3.8; and the experimental environment is built in PyCharm.

By monitoring the model training process, it can be found that the loss curve tends to flatten when the model has an epoch of about 130-160, indicating that the model is close to convergence at this time, so the Eopchs are set to 150. The model training parameters are shown in Table 1.

FIGURE 5. Examples of the experimental data.

TABLE 1. Training parameters of the model.

Training Parameters	Details
Epochs	150
Batch-size	16
Image-size	640×640
Initial learning rate	0.01

C. MODEL EVALUATION

The main metrics used in this paper are Precision, Recall, and mAP.

Precision refers to the proportion of samples that are actually positive in all the samples predicted by the model as positive categories, which measures the accuracy of the model in the prediction of positive categories; Recall refers to the proportion of samples that are actually positive in all the samples that are correctly predicted by the model as positive categories, which measures whether the model is able to efficiently find all the positive categories of samples;

The mAP is a commonly used evaluation metric in target detection tasks that combines the precision and recall curves of the model on different categories and calculates the average value, which measures the detection performance of the model on multiple categories, and is often used to evaluate the overall effectiveness of target detection algorithms. Where mAP50 denotes the mAP value at the 50% loU threshold. mAP50-95 is a more stringent evaluation metric that calculates the mAP values in the range of 50-95% loU thresholds and then averages them.

The calculation formula is:

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$AP = \int_0^1 P(R)\, dR \quad (3)$$

$$mAP = \frac{1}{C} \sum_{i=1}^{C} AP_i \quad (4)$$

where: TP denotes the number of samples correctly predicted by the model to be in the positive category; FP denotes the number of samples incorrectly predicted by the model to be in the positive category; FN denotes the number of samples incorrectly predicted by the model to be in the negative category; AP_i denotes the average precision of a single category, and C denotes the number of all categories.

V. RESULT AND DISCUSSION

A. EXPERIMENTAL RESULTS

The experimental result data of the improved YOLOv8-FEP algorithm is shown in Figures 6-8, Figures 6 and 7 demonstrate the Precision-Confidence Curve and Re-call-Confidence Curve, and Figure 8 demonstrates other related result plots such as mAP. Figure 9 illustrates the detection results of some validation sets of the model.

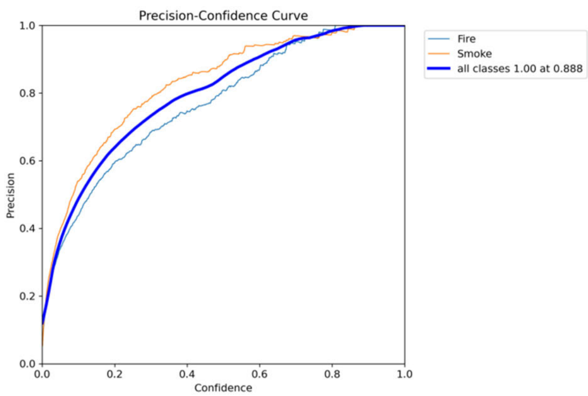

FIGURE 6. YOLOv8-FEP precision-confidence curve.

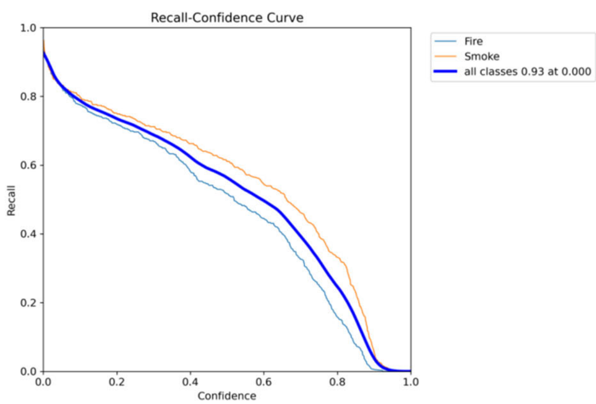

FIGURE 7. YOLOv8-FEP recall-confidence curve.

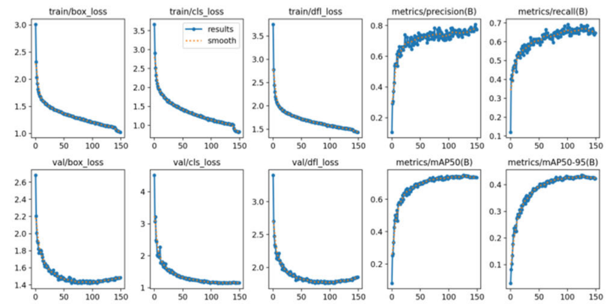

FIGURE 8. YOLOv8-FEP other relevant result charts.

FIGURE 9. Partial detection result graph of YOLOv8-FEP.

B. COMPARATIVE EXPERIMENTS ON DIFFERENT ATTENTION MECHANISMS

Table 2 displays the outcomes of the incorporating different attention mechanisms, encompassing four distinct types: CA, SA, CBAM, and EMA. Yolov8 incorporating EMA yields the most favorable outcomes. Specifically, it exhibits a notable improvement of 2% in mAP50 compared to the original YOLOv8.

TABLE 2. Comparative experiments on attention mechanisms.

Model	P(%)	mAP50(%)
YOLOv8	70.1	71.5
+CA	73.2	72.7
+SA	72.9	72.9
+CBAM	72.3	72.6
+EMA	75.1	73.5

C. EXPERIMENT

Ablation tests are conducted by merging many modules to confirm the efficacy of the modified method in this study. The experimental results on the same test set are shown in Table 3. From the table, we know that taking YOLOv8 as the baseline model: Combination 1 is the original YOLOv8 model without improvement; Combination 2 adds the large target detection head into the backbone module of YOLOv8, and the precision rate is improved by 4.5% and the mAP is improved by 2%; Combination 3 is the combination of Combination 2 and then introduces the fusion EMA, and the precision rate reaches 75.1%. Comparing with Combination 2, the mAP does not improve, but the precision rate and recall rate are both improved by 0.5%. Combination 4 introduces the PAN-Bag module, and the accuracy reaches the maximum value of 76.1%, but the recall and mAP decrease compared to combination 3. Combination 5 (YOLOv8-FEP) fuses combinations 3 and 4, and improves the mAP by 3.1%, the accuracy by 5.8%, and the recall by 2.1% compared to the original model. Meanwhile, the FPS value of Combination 5 is the lowest among the five models, and this parameter is an important indicator of the model's detection speed, which fully demonstrates that this model is more capable of meeting the demand for real-time fire detection. It is the model with the strongest performance in this experiment, proving the effectiveness of the improved algorithm in this paper.

TABLE 3. Results of ablation experiment.

	P6	EMA	PAN -Bag	P	R	mAP50	mAP50 -95
1	—	—	—	70.1	64.6	71.5	41.5
2	√	—	—	74.6	65.2	73.5	43.1
3	√	√	—	75.1	65.7	73.5	43.2
4	—	—	√	76.1	62.4	73.2	42.8
5	√	√	√	75.9	66.7	74.6	43.6

Data visualization work was done to illustrate the performance comparison graph between YOLOv8-FEP and YOLOv8, with the goal of more intuitively reflecting the effectiveness of the new algorithm. The comparison results are displayed in Figure 10. In Figure 10, the horizontal coordinates indicate the epoch number, and the vertical coordinates from Figure (a) to Figure (d) indicate the changes in Precision, Recall, mAP50, and mAP50-95, respectively.

D. COMPARISON WITH OTHER MODELS

In order to investigate more deeply the performance of the improved model proposed in this paper in the fire detection task, we conducted comparison experiments with other mainstream target detection models, including YOLOv5s, YOLOv3-Tiny, YOLOv7-Tiny [27], YOLOv8n, YOLOv8n-World, and the latest YOLOv9-Tiny, under the same dataset and training parameters. The results of the comparison experiments are shown in Table 4 and in Figure 11. From the table and figure, we know that the detection performance of the YOLOv8-FEP model in this paper is better than that of the YOLOv5s, YOLOv3-Tiny, YOLOv7-Tiny, YOLOv8n, YOLOv8n-World, and YOLOv9-Tiny models.

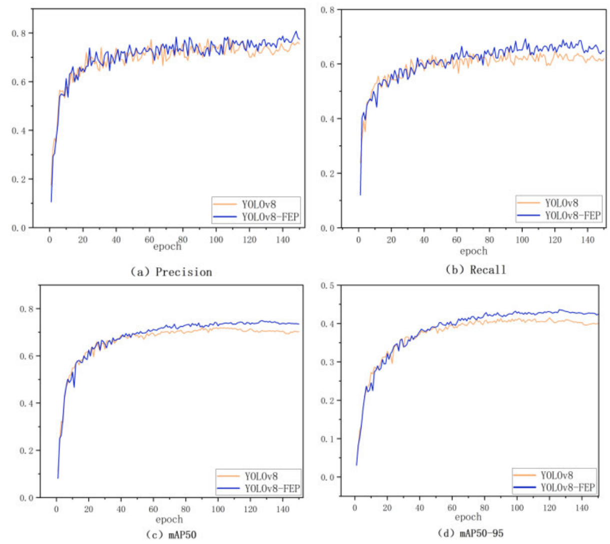

FIGURE 10. Comparison of results between YOLOv8-FEP and YOLOv8.

TABLE 4. Results of comparative experiments.

Models	P(%)	R(%)	mAP50(%)	FPS
YOLOv3-Tiny	60.3	59.6	60.3	132
YOLOv5s	71.1	62.0	69.8	125
YOLOv7-Tiny	68.6	63.7	68.3	181
YOLOv8n	70.1	64.6	71.5	426
YOLOv8n-World	73.4	65.2	72.4	413
YOLOv9-Tiny	74.3	64.3	74.1	172
YOLOv8-FEP	75.9	66.7	74.6	395

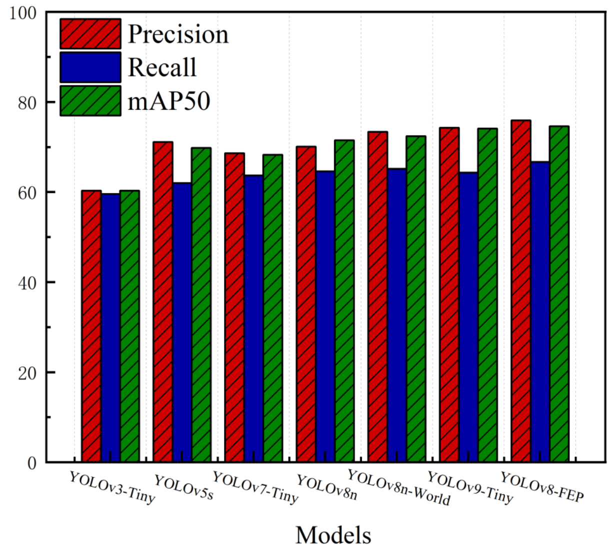

FIGURE 11. Visualization of the results of comparative experiments.

E. EXPERIMENTAL EFFECT VERIFICATION

Images from various scenarios in the test set are chosen in order to compare the model detection results before and after

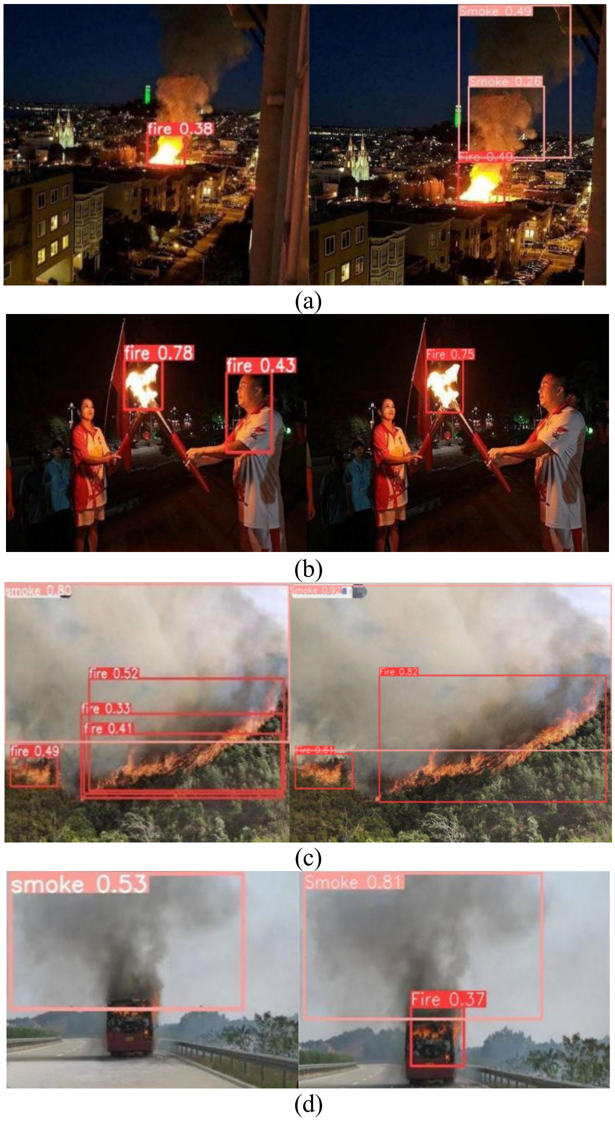

FIGURE 12. Comparison of YOLOv8-FEP and YOLOv8 detection results.

the improvement in order to clearly display the difference brought about by the improvement. Figure 12 displays the results of the comparison. The upgraded YOLOv8-FEP model's detection results are presented on the right side, while the original YOLOv8 model's detection results are displayed on the left. (a) The left figure does not identify smoke due to the background interference, and the right figure identifies light-colored smoke, which improves the model's detection performance of smoke; (b) The left figure incorrectly identifies the person as fire due to light interference from the flame, whereas the right figure accurately identifies the fire without misidentifying the person as fire. This reduces the problems of omissions and false alarms. (c) In the left figure, most of the fire are identified repeatedly. In the right figure, after increasing the detection range with the new model, there is no repeated identification, and the detection of fire is more comprehensive and accurate. Additionally, the confidence level for smoke identification is higher. (d) In the left figure, the fire on the red car body was not identified due to the influence of the car body, and only the smoke above the car was identified. In the right figure, the fire on the car body is accurately identified, and the confidence level for smoke identification is higher.

VI. CONCLUSION

In this study, we improve the YOLOv8 model and propose the YOLOv8-FEP model. The innovations of this model are the addition of a large target detection head in the backbone network part and the creative integration of an efficient multi-scale attention EMA and PAN-Bag module in the FPN. This design not only increases the sensory field of the network but also suppresses the interference of the background environment while highlighting the target features, so that the model pays more attention to the information in the target area, thus achieving the ability to extract the features of fire and smoke in complex scenes. In comparison to the original YOLOv8 and other mainstream detection models, the experimental results demonstrate that the YOLOv8-FEP model exhibits higher accuracy and robustness on the fire and smoke detection tasks. This results in a significant improvement in detection accuracy as well as a reduction in the likelihood of false alarms and missed detections. In addition to being useful for fire prevention and management in a range of settings, including public safety, industrial safety, and environmental monitoring, the algorithmic model offers a potent tool for fire monitoring that effectively lowers the amount of property loss and environmental damage caused by fire, raises society's standard of safety and disaster preparedness, and offers technical support for fire prevention and emergency response in the future. In the future, the optimization potential of the YOLOv8-FEP model remains to be deeply explored. By expanding the training dataset to include more diverse environmental conditions and complex scenarios, the model's detection accuracy in extreme or special environments can be further improved. Furthermore, ongoing research and development into the algorithm's capacity to recognize smoke and fire characteristics is strategically valuable in the long run for advancing the field of fire prevention science and fortifying the emergency response system.

REFERENCES

[1] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 23–28.

[2] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Santiago, Chile, Dec. 2015, pp. 1440–1448.

[3] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process. Systems*, vol. 28, Montreal, QC, Canada, Dec. 2015, pp. 7–12.

[4] W. Liu, D. Anguelov, and D. Erhan, "SSD: Single Shot MultiBox Detector," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 21–37, doi: 10.1007/978-3-319-46448-0_2.

[5] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 779–788.

[6] M. J. A. Daasan M. H. I. B. Ishak, "Enhancing face recognition accuracy through integration of YOLO V8 and deep learning: A custom recognition model approach," in *Proc. Asia Simulation Conf.* Cham, Switzerland: Springer, 2023, pp. 242–253.

[7] J. Terven and D. Cordova-Esparza, "A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS," 2023, *arXiv:2304.00501*.

[8] Y. Shen, J. Liu, G. Zhao, and X. Li, "A deep learning approach for fire detection using convolutional neural networks," *IEEE Access*, vol. 5, pp. 13251–13258, 2017.

[9] R. Ba, C. Chen, J. Yuan, W. Song, and S. Lo, "SmokeNet: Satellite smoke scene detection using convolutional neural network with spatial and channel-wise attention," *Remote Sens.*, vol. 11, no. 14, p. 1702, Jul. 2019.

[10] F. Yuan, L. Zhang, B. Wan, X. Xia, and J. Shi, "Convolutional neural networks based on multi-scale additive merging layers for visual smoke recognition," *Mach. Vis. Appl.*, vol. 30, no. 2, pp. 345–358, Mar. 2019, doi: 10.1007/s00138-018-0990-3.

[11] M. Jeon, H.-S. Choi, J. Lee, and M. Kang, "Multi-scale prediction for fire detection using convolutional neural network," *Fire Technol.*, vol. 57, no. 5, pp. 2533–2551, Sep. 2021.

[12] J. Jia, Y. Cao, X. Wang, and J. Huang, "Fire detection based on multi-model fusion," in *Proc. 4th Int. Conf. Image, Vis. Comput. (ICIVC)*, Sep. 2019, pp. 329–332.

[13] C. Chaoxia, W. Shang, and F. Zhang, "Information-guided flame detection based on faster R-CNN," *IEEE Access*, vol. 8, pp. 58923–58932, 2020.

[14] Q.-X. Zhang, G.-H. Lin, Y.-M. Zhang, G. Xu, and J.-J. Wang, "Wildland forest fire smoke detection based on faster R-CNN using synthetic smoke images," *Proc. Eng.*, vol. 211, pp. 441–446, Jan. 2018.

[15] Y.-Y. Qin, J.-T. Cao, and X.-F. Ji, "Fire detection method based on depthwise separable convolution and YOLOv3," *Int. J. Autom. Comput.*, vol. 18, no. 2, pp. 300–310, Apr. 2021.

[16] L. Zhao, L. Zhi, C. Zhao, and W. Zheng, "Fire-YOLO: A small target object detection method for fire inspection," *Sustainability*, vol. 14, no. 9, p. 4930, Apr. 2022, doi: 10.3390/su14094930.

[17] H. Liu, H. Hu, F. Zhou, and H. Yuan, "Forest flame detection in unmanned aerial vehicle imagery based on YOLOv5," *Fire*, vol. 6, no. 7, p. 279, Jul. 2023.

[18] J. Li, R. Xu, and Y. Liu, "An improved forest fire and smoke detection model based on YOLOv5," *Forests*, vol. 14, no. 4, p. 833, Apr. 2023.

[19] H. Wu, Y. Hu, W. Wang, X. Mei, and J. Xian, "Ship fire detection based on an improved YOLO algorithm with a lightweight convolutional neural network model," *Sensors*, vol. 22, no. 19, p. 7420, Sep. 2022.

[20] Y. Al-Smadi, M. Alauthman, A. Al-Qerem, A. Aldweesh, R. Quaddoura, F. Aburub, K. Mansour, and T. Alhmiedat, "Early wildfire smoke detection using different YOLO models," *Machines*, vol. 11, no. 2, p. 246, Feb. 2023.

[21] Z. Xue, H. Lin, and F. Wang, "A small target forest fire detection model based on YOLOv5 improvement," *Forests*, vol. 13, no. 8, p. 1332, Aug. 2022.

[22] X. Chen, B. Hopkins, H. Wang, L. O'Neill, F. Afghah, A. Razi, P. Fulé, J. Coen, E. Rowell, and A. Watts, "Wildland fire detection and monitoring using a drone-collected RGB/IR image dataset," *IEEE Access*, vol. 10, pp. 121301–121317, 2022, doi: 10.1109/ACCESS.2022.3222805.

[23] S. Woo, J. Park, and J. Y. Lee, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2018, pp. 3–19.

[24] Q. Hou, D. Zhou, and J. Feng, "Coordinate attention for efficient mobile network design," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Nashville, TN, USA, Jun. 2021, pp. 13708–13717.

[25] D. Ouyang, S. He, and G. Zhang, "Efficient multi-scale attention module with cross-spatial learning," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2023, pp. 1–5.

[26] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path aggregation network for instance segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 8759–8768.

[27] J. Zhu, J. Zhang, Y. Wang, Y. Ge, Z. Zhang, and S. Zhang, "Fire detection in ship engine rooms based on deep learning," *Sensors*, vol. 23, no. 14, p. 6552, Jul. 2023, doi: 10.3390/s23146552.

TIANXIN ZHANG was born in 1999. He is currently pursuing the master's degree with the School of Artificial Intelligence and Software, Liaoning Petrochemical University. His research interests include fire detection, YOLO algorithm, and image recognition.

FUWEI WANG was born in 1976. He is currently teaching with the School of Artificial Intelligence and Components, Liaoning Petrochemical University. He is also working as the Master's Tutor. His research interests include computer network and security and artificial intelligence.

WEIMIN WANG was born in 1975. He is currently teaching with the School of Artificial Intelligence and Components, Liaoning Petrochemical University. He is also working as an Associate Tutor for the master's degree students. His research interest includes artificial intelligence application technology.

QIHAO ZHAO was born in 2000. She received the degree in software engineering from the North University of China. She is currently pursuing the degree in computer science and technology with the School of Artificial Intelligence and Software, Liaoning Petrochemical University. Her research interests include network security, intrusion detection, and neural networks.

WEIJUN NING was born in 1999. He is currently pursuing the master's degree with the School of Artificial Intelligence and Software, Liaoning Petrochemical University. His research interests include the direction of text classification and sentiment classification for natural language processing.

HAODONG WU was born in 1998. He is currently pursuing the degree with the School of Information Control and Engineering, Liaoning Petrochemical University. His research interest includes text classification for natural language processing.