Proceedings of the 3rd International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/2025.30446

A Federated Learning Fire Detection Method Integrating YOLO11n-P2 and Fire-Yolopa

Xiaoyan Lu[1], Xinru Li[1], Jian Hu[1*]
[1]North China University of Technolog, Beijing, China
**Corresponding Author. Email: erxing41@gmail.com*

Abstract. The existing multi-source heterogeneous fire detection faced with issues of high privacy leak risk for data, low small-flame detection accuracy, and high communication overhead, so this paper proposed a federated fire detection model based on integrating YOLO11n-P2 detection model and Fire-Yolopa detection with partial aggregation. This framework is also the first to use YOLO11 series and federated learning together for fire detection tasks. Using the federated learning will allow for each client to participate in cooperation and share their knowledge with each other without having to share their real data, solving the issue of closed data silo in fire detection. Through experimental results, it can be seen that the optimal client model has an mAP50 value of 80.2%， only different from the centralized baseline model by 2.68%, and the small flame recall rate improves from 50.00% to 54.64% This result shows the improvements for Small flame detection in the terms of accuracy as well as the overall detected. This work will give a reusable technical framework which can help the community to collaboratively improve on small objects and privacy as well as deployment edge.

Keywords: Fire detection, federated learning, YOLO11n, small target detection, partial aggregation strategy

1. Introduction

Fire is a kind of sudden and destructive hazard and has caused great harm to human lives and properties. with the rapid increase of urbanization,industry,the application scene of fire detection has become more and more extensive. including indoor,outdoor,forest,and industrial. The videos and photos produced from these situations have useful fire-related information but may also include private person information or sensitive site information like where houses are laid out, how industrial buildings' pieces go together, what's on camera for public to watch over people do, or really important places that need super secret security like army bases or sick folks getting care at hospitals. A few recent works added federated learning on top of fire detection for cross device distributed training without sacrificing users' data privacy. Take for example Panneerselvam et al. [1], they put forward an IOFireNet framework based on MobileNet, the framework proved its effectiveness by showing great resilience in different types of complex environment like the indoor, outdoor, forest, and night-time scene with greatly reduced amount of communicated data and risk of privacy leak. Hu et al. [2] proposed the FedVIS algorithm, in which the federated dropout method is

DOI: 10.54254/2755-2721/2025.30446

combined with gradient selection methods to greatly enhance the communication efficiency and model convergence rate in multi-source data environments Lee et al. [3] Combined image clustering methods to form a distributed fire classification and localization model, achieving nearly 99% detection precision under non-Iid situations; taking the traits of edge environments into account, several effective ways of cooperation – like Partial Variable Training (PVT) [4] Partial Model Aggregation Chen et al. [5], and CluHFed [6], the hierarchical aggregation algorithm – were put forward to ease communication burdens and improve model adaptability. Although these studies show the great potential of federated learning on privacy protecting and distributed training, object detection algorithms using FL still have problems such as low detection accuracy for small-scale flames, poor adaptability in complex situations, and great communication burden.

As for fire detection, YOLO object detection technology is used as a new research point. Vasconcelos et al. [7] systematically reviewed recent developments in deep learning – Based fire detection and pointed out that the combination of remote sensing data and CNNs made the automation and accuracy of fire monitoring systems more automated and accurate. Megalingam et al. [8] continue to verify the effectiveness of the YOLO model and other popular object recognition algorithms, such as faster R-CNN, in detecting fires early from video images. Elhanashi et al. [9] did a thorough review with a view to model structure, development of dataset for training, and real-world deployment, pointing out that the current fire-detection systems are still encountering obstacles like being too sensitive, not having enough data marked, and being hard to set up on devices lacking resources. Among many detection methods, the YOLO series has become one of the main technologies used in fire-detection research because the model is complete and can make quick calculations. Since Redmon et al. [10] first gave birth to the YOLO framework, its framework has continuously evolved, with lightweight and highly accurate detection process. For example, Wang et al. [11] proposed YOLO-LFD and DSS-YOLO, with improved detection results for forest fire and small objects scenarios respectively. Pan et al. [12] proposed the YOLO-FireAD model, which can better adapt to the complex environment by introducing attention and dual pooling mechanism. Most of these models focus more on improving the detection accuracy and the speed inference rather than paying attention to the data safety and privacy during sensitive environment.

Federated learning deeply combined by the YOLO framework is of great academic value, there has been a lot of material which shows it working when it comes together. take for example Quéméneur and Cherkaoui [13], who combined YOLOv7 and a federated object detection framework under vehicle network scenario,showing that the approach could not only improve the detection accuracy but also protect the sensitive data well. Similar to this, Kim et al. [14] proposed a semi supervised Federated object detection framework, this framework made great progress in solving the problem of data heterogeneity and label scarcity. Although the combination of the YOLO framework and federated learning shows very effective results in some object detection fields, there is no research on the integration of the combination in the fire detection field. Therefore, this research plans to use the YOLO algorithm together with federated learning to detect fires, so as to establish a high-precision, low communication cost, and privacy preserving distributed fire detection system.

2. Method

2.1. Framework of federated fire detection

For the challenges of data privacy protection, detection accuracy, and communication overhead in fire detection scenarios under multi-source heterogeneous environments, this study proposes a federated fire monitoring framework based on the YOLO object detection model and a federated learning mechanism (as Figure 1 shows), each client locally trains independently, the central server employs the partial aggregation method Fire-YOLOPA, which allows multiple clients to achieve collaborative training and knowledge sharing. Training: Each edge client node trains the local YOLOv111P2submodel and only uploads the updated model parameters, not the raw images, to the central server, which periodically receives model updates from clients and performs partial aggregation on the Backbone and Neck modules, assigning aggregation weights for theP2andP3layersofthe feature perception layers that make up the global model; Meanwhile, the parameter of the Head module will be updated by each client in local to better adapt to the different distribution of data.

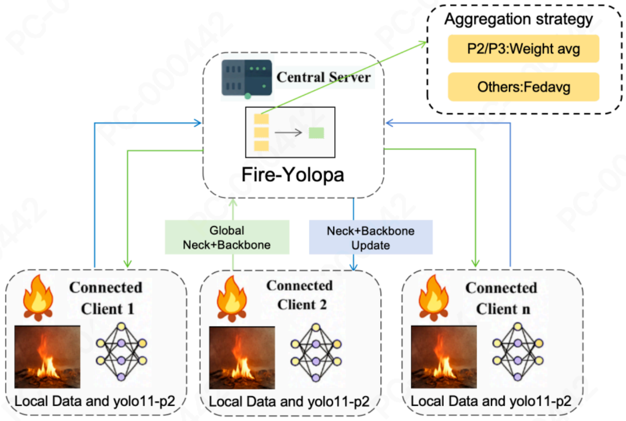

Figure 1. Framework of the federated fire detection system

2.2. YOLO11-P2 object detection model

YOLO11-P2, which is an upgraded version of the ultra-lightweight YOLO11n to overcome some of the technical difficulties found in situations related to very small objects. Its main architecture is a classic "Backbone–Neck–Head", with most major improvements coming from the multi-scale feature fusion mechanism at the Neck. To alleviate the feature undersampling problem for ultra-small targets with 10-25 pixels, it adds another P2 high-resolution feature perception layer on top of the previous P3 - P5 three-scale structure from YOLO11n and forms a P2 - P3 - P4 - P5 four-scale collaborative fusion system. This makes the model better at noticing very small fires and picking up on their tiny details, this helps when dealing with tricky cases where a lot is going on and you can get your hands messy trying to tell how everything should be reacted to by our little helpers up there in space.Considering the main core modular functionality, the Backbone of YOLO11-P2 is a collaborative module formed by Conv, C3k2, SPPF, C2PSA This design can efficiently extract cross-scale foundational and higher-level semantic features, and at the same time retain important details of ultra-small objects. In the Neck layer, the new P2 module adopts the nearest-neighbour up-sampling and channels concatenation to merge the Backbone’s high resolution with P3’s lower resolutions. fusion to solve the problem of attenuated deep semantic information and provide complete fine-grained information. On an input size of 640x640, it can still perfectly represent very minute fire sources. Therefore, this model can still be accurately detected even on complex and constantly changing surroundings.

DOI: 10.54254/2755-2721/2025.30446

2.3. Partial aggregation strategy (Fire-YOLOPA)

2.3.1. Bottleneck analysis of traditional federated aggregation

In the case of fire detection, in federated learning, full aggregation methods such as Federated Averaging adopt a strategy of equal weighting based on size when aggregating data. But there's a problem with this method when it comes to meeting the multi-scale small fire detection needs of the YOLO11n-P2 model. In the YOLO11n-P2 model, the P2 layer (which is aimed at very tiny fires) mainly has to depend on the details coming from each pixel, so the YOLO11n-P2 has to use more detailed contour information. On the other hand, the P3 layer (which tackles smaller fires) mainly relies on the structure around the fire's edge (structured contour). A traditional full-layer uniform weighting does a poor job of telling these different features apart, and so it results in small fire features not being properly represented in the global model, also, in terms of the actual data, fire datasets usually suffer from imbalance, having too many medium and big fire samples compared to small-fire ones. The above problem can be made worse by this imbalanced state. In the local feature extraction phase, the YOLO11n-P2 model was trained to acquire biased representations towards medium and large fires because of their predominant occurrence in the data. When these biased local models are put together via full-layer averaging, this imbalance carries through the feature transfer chain, making the smaller fire perception layer P2 weaker. Consequently, the small fire's minor visual characteristics get covered up by the big fire's obvious ones, so there is an extremely big drop in how well the small fire can be spotted. In reality, it means that very hard to detect small fire sources which is obviously a huge weakness for federated fire detection systems.

2.3.2. Construction of a hierarchical and multi-scale aggregation framework

Based on the principle of "hierarchical function-sampled attribute matching", which differentiates the aggregation of weighted samples for different core perception layers, while selectively aggregating non-core layers in the network to improve its capability of detecting small-fire targets and to achieve stable and efficient global aggregations of the whole model. In detail, for the ultra-small and small fire perception layer we apply the sampling proportion-oriented sample weighting strategy. Within the layers, the features of ultra-small and small fire targets mainly consist of pixel-level fine-grained details and structured contour information, and their relative sample proportions indicate each client's contributions to these feature representations. By utilizing sample-proportion based weighting, it increases the effect of client contribution to small-fire features to better detect small and ultra-small fire targets. At the same time, base feature extraction layers and medium-size to large-size target layers still use standard Fed Avg aggregation to satisfy the requirement for cross-client general feature learning and maintain generalization for medium size to large size fire targets. The second is using the "non-participation in federated aggregation" for the multi-scale detection head at layer 29. A way to preserve the independence of local client parameters that reduces the amount of redundant parameter aggregation and communication costs. It improves communication efficiency well and avoid redundant computations, it could support robust multi-scale fire detection with federation learning setting.

DOI: 10.54254/2755-2721/2025.30446

Table 1. Hierarchical division of YOLO11n-P2 module

Module category	Corresponding laver number/structure	Resolution (640 input)
The basic feature extraction layer	0-10(Conv,C3k2,SPPF,C2PSA)	320×320→20×20
ultra-small fire perception layer	17-19(Upsample,Concat,C3k2)	160×160
Slightly lower temperature sensing layer	14-16(Upsample,Concat,C3k2)	80×80
Medium and large-sized target layer	20-28(P4/16,P5/32)	40×40→20×20
Multi-scale detection head	29(Detect (P2,P3,P4,P5))	-

3. Experiments and analysis

3.1. Data and experimental preparation

The fire detection datasets adopted in this study were taken from the publicly available Kaggle datasets – "Fire / Smoke Detection YOLO v9." which is dedicated exclusively to conducting fire and smoke object detection tasks. In terms of coverage, The dataset encompasses a large quantity of scenes and interference, offering fire and smoke samples which can adequately back up the performance evaluation of models in complex real world scenarios. It includes extensive imagery of fire and smoke under indoor scenes and outdoor scenes, as well as commonly encountered interference conditions such as high illumination levels, smoke occlusions, and complex backgrounds. To focus on the core task of flame detection, we select only the data with the "fire" class from the collected sample set, and finally get an experimental dataset of 18607 fire images. We performed a global random shuffle after which the data were equally split among three clients such that each received roughly 6200 images. For each client's dataset, it must preserve all type of scenes and interference situations to avoid performances fluctuation from federated training brought on by biased data distribution. This kind of allocation creates a decent and credible information cornerstone for later model training and strategy verification.

3.2. Analysis of local and federated experiments

To compare the detection performance with non-federated training, standard local training experiments were conducted on each client's data partition. These models were trained exclusively on the client-specific dataset and did not use weights from models trained on other clients' data. During training, no data exchange occurred between clients. The experimental results are summarized as follows:

DOI: 10.54254/2755-2721/2025.30446

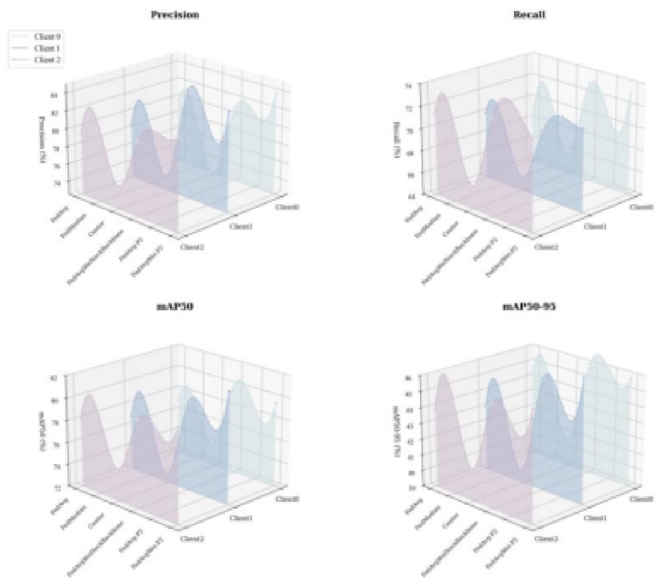

Figure 2. Comparison of different framework strategies on client-side datasets

From the validation result shown in figure 2, we can see that even if it is federated learning strategies such as Federated learning strategies such as Fed Avg, Fed Median and so on, which all outperform non federated local training models such as Center series in terms of every evaluation metric for both YOLO11n and YOLO-1n-P2 model.Fed Avg strategy under the YOLO11n model gets an accuracy of 80.82% and recall of 66.80% and mAP50 of 76.05% on Client 0, which has increased by 5.14%, 2.94%and 3.85%than Center series local training model's 75.68%, 63.86%, 72.20%. Fed Median achieves an mAP50-95 of 42.07% for Client 2, 3.11% more than the local model's 38.96%. Even with the simpler Fed Avg strategy under the YOLO11n-P2 framework, Client 2 reaches a precision and recall of 77.83% and 72.30% respectively, outperforming the best local YOLO11n training result which has a precision and recall of 77.59% and 63.86% respectively. the optimal Fire-Yolopa strategy has improved in detection, achieving mAP50 = 79.91, and mAP50-95 = 45.92 on Client 1, an improvement from the best local result of 7.71% and 6.96%, respectively. The improvement is consistent across different frameworks, showing that federated learning does not depend on a specific model framework, but rather a cross-client fusion of features, resulting in better results, by compensating for small, non-distributed client sets.

3.3. Ablation study design and analysis

To verify the effectiveness of the model architecture improvements and aggregation strategy optimizations introduced in this study, an ablation experiment was designed using the controlled variable method.The experiment uses centralized training baselines and YOLO11n with full aggregation as reference points. The proposed improvements were then incrementally introduced, and their impact on small-fire detection performance was quantified using the metrics: precision, recall, mAP50, and mAP50-95. The aggregated results are summarized in Table 2, while detailed performance comparisons for each client are provided in Table 3.

DOI: 10.54254/2755-2721/2025.30446

Table 2. Comparison of key results from ablation experiments (server side/client side best)

Model/Training Strategy	Precision (%)	Recall (%)	mAP50 (%)	mAP50-95 (%)	Core improvement direction
Center_YOLO11n	84.55	75.51	82.88	48.74	-
Center_YOLO8n	81.68	70.53	78.24	44.62	-
FedMedian_Sever_YOLO11n	79.69	69.31	76.50	42.63	YOLO11n+Fully Aggregated FL
FedAvg_Sever_YOLO11n	82.46	67.30	76.47	42.57	YOLO11n+Fully Aggregated FL
FedAvg_Sever_YOLO11-P2	81.60	72.10	78.10	45.33	Introduce the P2 high-resolution feature layer
FedAvgWei_Sever_YOLO11-P2	84.02	73.25	81.53	47.81	P2 layer + neck layer - weighted P2/P3 sections
FedAvgWeiNeckBackbone_YOLO11-P2(best)	84.74	72.01	79.9	45.98	Partial weighting + Only aggregation of neck/backbone

3.3.1. Performance comparison between baseline models and YOLO11n full aggregation

First, centralized training models (YOLO8n, YOLO11n) and the YOLO11n full aggregation (FedAvg) framework were used as baselines to evaluate the performance differences between federated training and different base architectures. The results show that centralized YOLO11n (mAP50: 82.88%, recall: 75.51%) outperforms YOLO8n (mAP50: 78.24%, recall: 70.53%), demonstrating the superior adaptability of the YOLO11 series for fire detection tasks.Meanwhile, the server-side FedAvg aggregation of YOLO11n achieved a precision of 82.46%, slightly lower than centralized YOLO11n, but still 0.78% higher than centralized YOLO8n, while effectively eliminating data privacy risks inherent in centralized training. This result validates the balance between model performance and privacy protection achieved by federated training and establishes a baseline reference for subsequent methodological improvements.

3.3.2. Validation of YOLO11n-P2 model structure

Figure 3 Comparison between YOLO11n Server-side Full Aggregation and YOLO11n-P2 server-side full aggregation (both using FedAvg). YOLO11n-P2 reaches a considerable increase in recall by 4.8% from 67.3% to 72.1% and a higher mAP50 by 1.63% from 76.47% to 78.1%, however the precision does not drop greatly from 82.46% to 81.6%. These findings prove the core value of the newly defined P2 high-res feature layer in YOLO11n-P2 by making use of 160*160 feature map to seize local details about ultra small fire, and reducing the undersampling issue of small target. Recall, there is significant upping, there is an obvious drop with what was missed to detect, making it certain that a baseline in detecting, the YOLO 11n-P2, is where this took its start and ended up being.

DOI: 10.54254/2755-2721/2025.30446

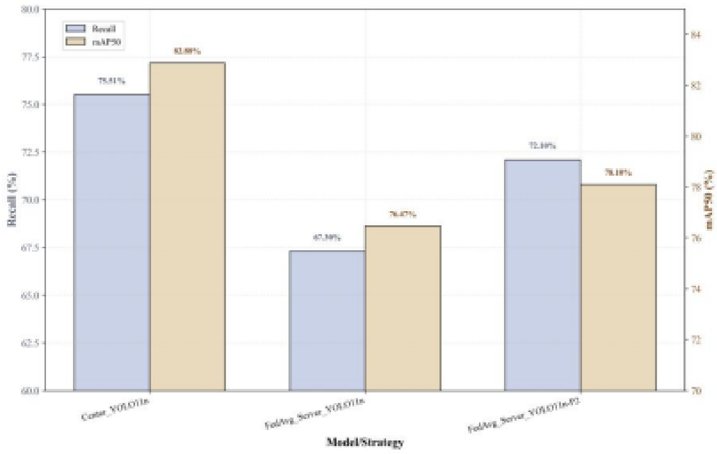

Figure 3. Comparison of model structure before and after optimization

And then the detailed metrics for each of the clients also confirm this. As you can see in Table 3: And for all client’s recall rate of the FedAvg_YOLO11 - P2 model is more than 72%, compared with the FedAvg_YOLO11n, it has increased by 5-6 percentage points. This shows that the P2 layer always improves small-target recall no matter which client has different data distribution, which means it’s very tough and can do well regardless of the local data bias.

Table 3. Comparison of key results from ablation experiments (client side)

Model / Strategy	Client	Precision(%)	Recall (%)	mAP50 (%)	mAP50-95 (%)
	Client0	80.82	66.80	76.05	42.05
FedAvg_YOLO11n	Client1	79.50	66.47	75.72	41.75
	Client2	80.24	67.03	76.28	42.20
	Client0	80.60	73.35	78.15	44.79
FedAvg_YOLO11-P2	Client1	79.96	72.82	77.55	43.94
	Client2	77.83	72.30	76.73	43.47
	Client0	82.13	73.52	79.34	45.62
FedAvgWei_YOLO11-P2	Client1	82.71	72.83	78.85	46.13
	Client2	81.54	72.21	79.62	44.91
	Client0	81.07	72.83	80.26	46.03
FedAvgWeiNeckBackbone_YOLO11-P2	Client1	84.74	72.01	79.91	45.92
	Client2	80.73	72.68	79.63	45.46

3.3.3. Validation of aggregating only Neck/Backbone

As can be seen from Table 2, after the introduction of the Neck/Backbone-only aggregation scheme along with the partial weighting method, it yields an optimal client model which reports precision @ 0.5 of 84. 74%, recall @ 0.5 of 72. 01%, and MAP@ 0. 50 of 79. 90%. Compared to the server-side partial weighting + full Head aggregation model, all metrics saw decreases under 1.5%, even showing slight improvements in precision, while also being able to reduce communication data

DOI: 10.54254/2755-2721/2025.30446

volume by about 18%. This indicates that the Head layer, as a mapping mechanism from features to detection outputs, has great generalization for fire detection tasks. not excluded from the aggregation process, and will not cause severe decrease in performance, but reduce the federated communication cost, which verifies the engineering advantage of this strategy. As shown in table 3, Under this strategy, Clients maintained a stable mAP50 between 79.63% and 80.26%, and the recall rate was between 72% and 73%, with no obvious volatility. These results further suggest that the Head layer's ability to generalize its mapping capacity enables the system to keep the stable object detection performance even when communication efficiency is improved which makes this model achieve a trade-off between accuracy and practical deployment efficiency, thus can be implemented on edge-device quite well.

To sum up, the ablation experiments were performed incrementally from good, with improved modules. The performance improvement of different parts is clear. YOLO11n-P2 is the main framework for increasing small object recall, and partial weighting improves feature aggregation precision. The "Neck/Backbone-only aggregation" approach improves engineering efficiency, but maintains performance. Considering performance comparison, the Federated Learning – enabled algorithm shows strong competitive performance in terms of performance. Take FedAvgWei Neck Backbone(Fire-Yolopa)_YOLO-P2 as an example; client1 reaches 84.74% precision that beats Central_YOLO11n precision of the centralized training baseline. Though the recall, mAP50, and mAP50-95 for each client are slightly below the centralized baseline, the differences are still acceptable. And all federated schemes outperformed the YOLO8n centralized training and the YOLO11n full-aggregation strategy. The results show that our proposed federated improvement scheme can approximate the performance of centralized training, while eliminating the risk of raw data privacy leakage and allowing for distributed training. Consequently, it achieves a balance among "privacy protection – detection performance – engineering efficiency," providing a solution that ensures both technical effectiveness and practical value in multi-source heterogeneous fire detection scenarios.

3.4. Validation of small-fire target recall

To explore the effectiveness of the proposed improvements towards resolving the core problem of missed small-fire targets even further, this section studies the small-fire recall rate (Table 4), and contrasts the different model architecture and aggregation strategies, to point out the specific value of the improvements presented. As per Table 4, the proposed strategies have major advantages in the small-fire recall rates on 3 dimensions: Model architecture optimization, which is the foundational performance upgrade for small-fire detection. The YOLO11n-P2 series keeps achieving higher small-fire recall than the YOLO11n and YOLO8n baselines. FedAvg_Server_YOLO11-P2 gets a small fire recall of 54.53%, 2.23% higher than FedAvg_Server_YOLO11n with the same aggregation strategy. This difference highlights the core value of the extra P2 high-resolution feature layer in YOLO11n-P2.And to improve the aggregation strategies allows for even more potential to be detected by the small flame. FedAvgWei_Server_YOLO11n-P2 recognized as the best strategy in this study, reaches a small-flame recall of 55.12%,which is higher than the FedAvg method under the same architecture and the baseline model YOLOv11n_Center. This is due to the specific design of the "partial weighting of the Neck layer P2 / P3," strategy, which dynamically assigns an aggregation weight to the P2 layer according to the proportion of ultra-small flame samples; the contribution to features from clients with many ultra-small flame images is intensified. This can avoid the dilution problem of small flame which is caused by uniform full aggregation. This way we achieve better recall performance of small-flame detection.

DOI: 10.54254/2755-2721/2025.30446

Finally, in terms of the aggregation scope optimization, stable small fire detection with engineering efficiency optimization. Under the FedAvg Wei Neck Backbone (Fire-YOLOPA) scheme, the client level small-fire recall range is between 52.98% - 54.64%. The highest deviation from the FedAvgWei strategy across all clients was only 0.66%. Small fire recall is almost stable on both sides without a prominent fluctuation trend. It shows that except for the aggregation of the Head layer, there is almost no reduction in the detection performance of small fires. Compared with the original strategy, about 18% of the communication data can be saved, which also shows that the strategy has engineering significance.

Table 4. Comparison of small fire target recall rates for different models / training Strategies

Model / Strategy	Small target flame recall (%)
YOLOv8n_Center	40.31
YOLOv11n_Center	50.00
FedAvg_Sever_YOLO11n	52.30
FedAvg_Client0_YOLO11n	48.72
FedAvg_Client1_YOLO11n	47.70
FedAvg_Client2_YOLO11n	50.00
FedMedian_Sever_YOLO11n	51.79
FedMedian_Client0_YOLO11n	47.19
FedMedian_Client1_YOLO11n	48.15
FedMedian_Client2_YOLO11n	49.52
FedAvg_Sever_YOLO11-P2	54.53
FedAvg_Client0_YOLO11-P2	52.59
FedAvg_Client1_YOLO11-P2	51.79
FedAvg_Client2_YOLO11-P2	51.21
FedAvgWei_Sever_YOLO11n-P2	55.12
FedAvgWei_Client0_YOLO11n-P2	54.35
FedAvgWei_Client1_YOLO11n-P2	53.84
FedAvgWei_Client2_YOLO11n-P2	53.07
FedAvgWeiNeckBackbone_Client0_YOLO11n-P2	53.68
FedAvgWeiNeckBackbone_Client1_YOLO11n-P2	54.64
FedAvgWeiNeckBackbone_Client2_YOLO11n-P2	52.98

4. Conclusion

In response to the suppression of small-flame features in traditional federated full-aggregation strategies, this paper proposes a new partial aggregation approach called “Fire-Yolopa” strategy, which establishes a two-level structure with “differential weighting in core perception layer + selective aggregation in non-core layer”. The ultra-small and small-flame perception layers adopt a dynamic weighting scheme based on the proportion of small-flame samples, magnifying the weight contribution of clients with higher contribution. The basic feature extraction layers and medium-to-large object layers (layers 20-28) adopt standard FedAvg aggregation for ordinary performance and the multi-scale detection heads keep their local updates to avoid redundant costs, supporting edge

DOI: 10.54254/2755-2721/2025.30446

deployment and diverse situations such as residential areas and industrial sites. Experimental results show that using this method can make the global model reach an mAP50 of 79.9 %, reducing Communication data by about 18 %. Full-head aggregation leads to mAP50 81.53%, almost as good as the centralized Center-YOLO11n baseline at 82.88%. Even under the “Neck/Backbone-only aggregation” scheme, there is an improvement in precision, recall and mAP50, all within 1.5%, achieving a threefold increase in detection performance and privacy and communication efficiency.

It is the first research paper that integrates YOLO11 series with Federated Learning to detect fire; it fills the research gap of small object federated detection research in this field. The proposed Fire-Yolopa strategy represents a reusable technical recipe for the orchestration of "small-object perception, privacy safety, and edge deployments," it can be put into use with different kinds of small-object detection efforts too, like observing forest fires, industrial blazes. Future study may also be done on federated fusion of multi-modal fire data and investigating adaptive aggregation schemes w.r.t dynamic clients, so as to cater for more complex scenarios with heterogeneous devices/data distributions.

References

[1] Panneerselvam S, Thangavel S K, Ponnam V S, et al. Federated learning based fire detection method using local MobileNet [J]. Scientific Reports, 2024, 14(1): 30388.

[2] Hu Y, Fu X, Zeng W. Distributed fire detection and localization model using federated learning [J]. Mathematics, 2023, 11(7): 1647.

[3] Lee J, Kang J, Park C S, et al. Distributed Fire Classification and Localization Model Based on Federated Learning with Image Clustering [J]. Applied Sciences, 2024, 14(20): 9162.

[4] Yang T J, Guliani D, Beaufays F, et al. Partial variable training for efficient on-device federated learning [C]//ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2022: 4348-4352.

[5] Chen Z, Yi W, Shin H, et al. Efficient wireless federated learning with partial model aggregation [J]. IEEE Transactions on Communications, 2024, 72(10): 6271-6286.

[6] Chen J, Li W, Yang G, et al. Federated learning meets edge computing: A hierarchical aggregation mechanism for mobile devices [C]//International Conference on Wireless Algorithms, Systems, and Applications. Cham: Springer Nature Switzerland, 2022: 456-467.

[7] Vasconcelos R N, Franca Rocha W J S, Costa D P, et al. Fire detection with deep learning: A comprehensive review [J]. Land, 2024, 13(10): 1696.

[8] Megalingam R K, Kota A H, Kota A H. Fire Detection: Harnessing the Power of Object Detection Algorithms for Advanced Visual Analysis [C]//International Conference on Information and Communication Technology for Intelligent Systems. Singapore: Springer Nature Singapore, 2024: 317-327.

[9] Elhanashi A, Essahraui S, Dini P, et al. Early Fire and Smoke Detection Using Deep Learning: A Comprehensive Review of Models, Datasets, and Challenges [J]. Applied Sciences, 2025, 15(18): 10255.

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

[11] Wang H, Zhang Y, Zhu C. YOLO-LFD: A Lightweight and Fast Model for Forest Fire Detection [J]. Computers, Materials & Continua, 2025, 82(2).

[12] Pan W, Xu B, Wang X, et al. YOLO-FireAD: Efficient Fire Detection via Attention-Guided Inverted Residual Learning and Dual-Pooling Feature Preservation [J]. arXiv preprint arXiv: 2505.20884, 2025.

[13] Quéméneur C, Cherkaoui S. Fedpylot: Navigating federated learning for real-time object detection in internet of vehicles [J]. arXiv preprint arXiv: 2406.03611, 2024.

[14] Kim T, Lin E, Lee J, et al. Navigating data heterogeneity in federated learning: a semi-supervised federated object detection [J]. Advances in Neural Information Processing Systems, 2023, 36: 2074-2096.