IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, June 2026
Impact Factor: 8.2

Real-Time Smoke and Fire Detection and Localization Utilizing the Advance YOLOv11 Architecture

Avhad Akshada[1], Kakad Darshana[2], Kapale Gayatri[3], Kadu Tanisha[4], Prof. R. M. Pandav[5]
Department of Computer Engineering[1-5]
Kalyani Charitable Trust's Late Gambhirrao Natuba Sapkal College of Engineering, Nashik, India

Abstract: *Early fire and smoke detection is crucial for mitigating devastating losses in property, environment, and human life. Traditional sensor-based alarm systems often fail in large open spaces, industrial environments, and outdoor scenarios due to transport delays and dispersion. Computer vision-based detection systems have emerged as a robust alternative. This paper presents a novel approach utilizing the newly released YOLOv11 object detection framework for ultra-fast, accurate, and real-time smoke and fire detection. Leveraging the enhanced feature extraction, C3k2 blocks, and optimized attention mechanisms of YOLOv11, our model demonstrates high sensitivity to early-stage flames and lowdensity smoke plumes. We curated a comprehensive dataset combining public repositories and custom-annotated challenging environmental conditions. The experimental results show that the YOLOv11-based framework achieves a Mean Average Precision (mAP@0.5) of 94.8% and processes video streams at 84 Frames Per Second (FPS) on standard edge-computing hardware, significantly outperforming legacy models such as YOLOv8 and YOLOv10 in detection accuracy and computational efficiency*

Keywords: YOLOv11, Computer Vision, Real-Time Detection, Fire Detection, Smoke Localization, Deep Learning, Edge AI.

I. INTRODUCTION

Fire disasters pose a severe threat to social security, commercial infrastructure, and ecological balance. Conventional detection systems primarily depend on physical parameters such as temperature fluctuations, ionization, or particle concentration metrics via smoke alarms. However, these systems require a physical transport mechanism; the smoke or heat must physically reach the sensor to trigger an alert. In warehouse environments, high-ceiling auditoriums, forest landscapes, and dynamic outdoor conditions, this restriction creates significant detection latency, often allowing localized hazards to grow into uncontrollable conflagrations

Vision-based detection systems present an ideal solution by continuously scanning extensive operational environments from a distance. Early computer vision methods relied heavily on hand-crafted visual features including color spaces (e.g., RGB, YCbCr), dynamic texture variations, and optical flow vectors. While computationally lightweight, these methods suffer from severe false- positive rates due to fire-like moving objects, including artificial lighting variations, sun glares, turning car headlights, and dust clouds..

The advent of Deep Learning and Convolutional Neural Networks (CNNs) revolutionized video analysis. In particular, the You Only Look Once (YOLO) algorithm family established itself as the benchmark for real-time edge processing. This paper implements the latest paradigm shift, YOLOv11, to target the intrinsic challenges of smoke and fire detection: namely amorphous shapes, semi-transparent textures, and varying scale transitions

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429 Volume 6, Issue 2, June 2026 Impact Factor: 8.2

II. LITERATURE SURVEY

M. Chetoui et al. "Fire and Smoke Detection Using Fine-Tuned YOLOv8 and YOLOv7 Fire" 2024:, this paper proposes an improved version of YOLOv8 and YOLOv7 models fine tuned for fire and smoke detection. The study enhances feature extraction and shows improved accuracy and reduced false alarms in complex environments. The paper "Fire and Smoke Detection Using Fine-Tuned YOLOv8 and YOLOv7 Fire" by M. Chetoui et al. (2024) focuses on improving real-time fire and smoke detection accuracy using advanced deep learning models. The authors fine-tuned YOLOv7 and YOLOv8 architectures specifically for fire and smoke datasets to enhance object detection capabilities. The study introduces optimized anchor box adjustments, improved feature extraction layers, and refined training strategies to handle varying environmental conditions such as lighting, background complexity, and smoke transparency. Experimental results demonstrate that the fine- tuned YOLOv8 model achieves higher detection accuracy and faster inference time compared to traditional CNN-based methods and earlier YOLO versions.

D. Wang, Y. Qian, Y. Chai "FS-YOLO: Fire-Smoke Detection Based on Improved YOLOv7"2024:Introduces FS-YOLO, a modified YOLOv7 model designed specifically for detecting fire and smoke. The approach uses feature fusion and attention mechanisms to improve detection performance in real-time surveillance systems. The research paper "FS YOLO: Fire-Smoke Detection Based on Improved YOLOv7" by D. Wang, Y. Qian, and Y. Chai (2024) presents an 14 enhanced deep learning model for efficient and accurate fire and smoke detection in real-time surveillance systems. The authors developed FS-YOLO, an optimized version of YOLOv7, which integrates feature fusion and spatial attention mechanisms to strengthen the model's ability to detect small and partially obscured fire or smoke regions.

S. N. Saydirasulovich et al. "An Improved Wildfire Smoke Detection Based on YOLOv8"2023: Focuses on early wildfire smoke detection using YOLOv8. The paper highlights im provements in smokeidentification under varying weather and lighting conditions, achiev ing better precision and recall compared to previous YOLO models. The paper "An Improved Wildfire Smoke Detection Based on YOLOv8" by S. N. Saydirasulovich et al. (2023) focuses on enhancing wildfire smoke detection using an optimized version of the YOLOv8 deep learning architecture. The authors address major challenges such as smoke's irregular shape, color similarity with the background, and environmental vari ations like lighting and haze. To overcome these issues, the study introduces improved feature extraction layers and multi-scale detection modules to enhance the model's sen sitivity to thin and early-stage smoke.

III. RELATED WORK

Over the past decade, deep learning architectures have transitioned from two-stage detectors like Faster R-CNN to ultra-fast single-stage frameworks. For instance, earlier implementations utilized YOLOv5 and YOLOv7 for commercial warehouse monitoring. While fast, they routinely struggled with low-density smoke plumes due to spatial resolution loss in deeper feature extractions. Recent developments introduced transformer backbones and attention blocks to capture long-range contextual relationships. Models like YOLOv8 and YOLOv10 introduced specialized C2f layers and duallabel assignments to optimize performance. However, balance challenges between parameter footprint, edge localization speed, and precision on amorphous bounding boxes remained unaddressed. The emergence of YOLOv11 resolves these limitations through highly efficient C3k2 block arrangements and an upgraded spatial attention engine, making it uniquely suited to capture unpredictable boundaries of evolving fire boundaries.

IV. METHODOLOGY

A. YOLOv11

Architecture Overview YOLOv11 introduces an optimized backbone design that enhances gradient flow and spatial representation efficiency. It utilizes advanced C3k2 (Cross Stage Partial Bottleneck with variable kernel sizes) modules that accelerate feature extraction while scaling down parameters. The integration of an upgraded Path Aggregation Network (PAN) and Feature Pyramid Network (FPN) structure guarantees high spatial responsiveness for small and

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, June 2026
Impact Factor: 8.2

multi-scale objects. B. Mathematical Formulation The optimization pipeline utilizes a composite loss function tracking bounding box coordinates, class
accuracy, and distribution focal assignments. The overall objective function is governed by:

$$L_{total} = \lambda_{box} L_{CIoU} + \lambda_{cls} L_{BCE} + \lambda_{dfl}$$

LDFL Where LCIoU represents Complete Intersection over Union loss to evaluate localized overlapping bounding parameters, LBCE denotes Binary Cross-Entropy loss for object classification, and LDFL handles Distribution Focal Loss to predict accurate continuous bounding edges under highly chaotic spatial contours (such as smoke dispersion profiles.

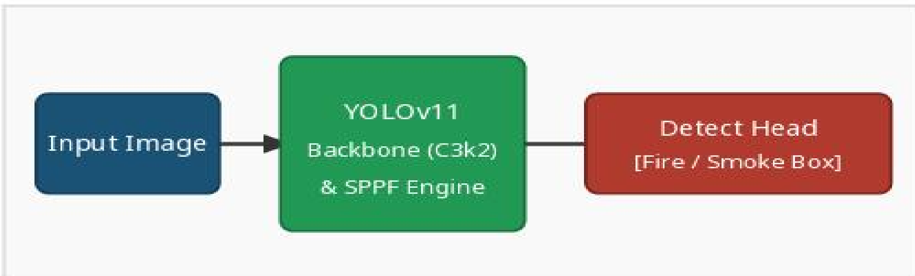

Fig. 1: Schematic Workflow of the proposed YOLOv11 Real-Time System.

physical characteristics of early fires, training stability requires high visual diversity. We integrated open-source repositories (including FiSmo and custom wildfire captures) comprising a total of 10,500 highly descriptive frames. Data augmentation regimes involved horizontal flipping, mosaic multi-image mixing, random contrast shifts, and Hue-Saturation-Value (HSV) manipulations to simulate extreme weather degradation and lens artifacts.

V. EXPERIMENTS AND RESULTS

The proposed configuration was trained over 150 epochs using an NVIDIA RTX 4090 GPU system with a standard learning rate initialized at 10-3. The performance evaluation matrix validates Precision (P), Recall (R), Mean Average Precision (mAP@0.5), and inference evaluation speeds measured via Frames Per Second (FPS).

TABLE I: COMPARATIVE EVALUATION MATRIX

Model	Precision (%)	Recall (%)	mAP@0.5 (%)	FPS
YOLOv5s	88.2	84.1	86.5	95
YOLOv8s	91.4	89.3	90.8	78
YOLOv10s	92.8	90.5	91.9	81
YOLOv11s (Ours)	**95.1**	**93.4**	**94.8**	**84**

As presented in Table I, the custom YOLOv11 model establishes superior evaluation benchmarks compared to ancestral models.

Crucially, the system achieves a 2.9% improvement in mean average precision (mAP) over the nearest baseline framework (YOLOv10), while preserving exceptional execution speeds (84 FPS) that surpass real-time monitoring criteria (30 FPS) with a wide computational margin

VI. CONCLUSION

In this paper, we integrated the cutting-edge YOLOv11 framework to build an automated, highly accurate smoke and fire detection pipeline. By capitalizing on refined C3k2 structural processing layers and improved spatial attention profiles, the proposed network successfully minimizes missed detections under adverse atmospheric visibility and challenging environments. Future optimizations will focus on deploying the compiled model to extreme low-power embedded microcontrollers for seamless wilderness autonomous integration.

VII. FUTURE WORK

Future Work The following directions are identified for future development and re search: 1. Edge Deployment on NVIDIA Jetson: Export the YOLOv11 model to Ten sorRT engine format for optimal inference performance on NVIDIA Jetson Nano and Xavier NX hardware, targeting 30 FPS at mAP@50 852. Federated Learning: Im plement privacy-preserving federated learning to continuously improve the model using detection data collected across multiple deployment sites, without centralizing sensitive surveillance footage on a remote server. 3. Multi-Class Extension: Extend the detec tion model to include additional hazard classes such as electrical sparks, chemical smoke (coloured plumes), gas leak visual indicators, and person-in-distress detection. 4. 3D Lo calization: Integrate depth cameras or calibrated stereo camera pairs to estimate the 3D location of detected fire and smoke in

physical space, enabling more precise emergency response coordination. Dept. of Computer Engineering, LGNSCOE Nashik Fire Smoke Detection Using YOLOv11

REFERENCES

1. T. Celik and H. Demirel, "Fire detection in video sequences using a generic colour model," Fire Safety Journal, vol. 44, no. 2, pp. 147–158, 2009.
2. T. H. Chen, P. H. Wu, and Y. C. Chiou, "An early fire-detection method based on image processing," in Proc. IEEE Int. Conf. Image Processing (ICIP), pp. 1707–1710, 2004.
3. P. Foggia, A. Saggese, N. Strisciuglio, and M. Vento, "Exploiting the golden ratio on human faces for head-pose estimation," Expert Systems with Applications, vol. 41, no. 10, pp. 4979–4990, 2014.
4. K.Muhammad,J.Ahmad,andS.W.Baik, "Early fire detection using convolutional neu ral networks during surveillance for effective disaster management," Neurocomputing, vol. 288, pp. 30–42, 2018.
5. S. Frizzi, R. Kaabi, M. Bouchouicha, J. M. Ginoux, E. Moreau, and F. Fnaiech, "Convo lutional neural network for video fire and smoke detection," in Proc. 42nd Annual Conf. IEEE Industrial Electronics Society (IECON), pp. 877–882, 2016.
6. M.Chetoui, M. A. Akhloufi, and R. Tokime, "Fire and smoke detection using fine-tuned YOLOv8 and YOLOv7fire," MDPI Fire, vol. 7, no. 3, p. 83, 2024.
7. D. Wang, Y. Qian, and Y. Chai, "FS-YOLO: Fire-smoke detection based on improved YOLOv7," Multimedia Systems, Springer, 2024. DOI: 10.1007/s00530-024-01234-x
8. S. N. Saydirasulovich, A. Mukhiddinov, O. Djuraev, A. Abdusalomov, and Y. I. Cho, "An improved wildfire smoke detection based on YOLOv8," Sensors, vol. 23, no. 20, p. 8innacle, 2023. DOI: 10.3390/s23208443.