Enhanced Real-Time Fire Detection Systems Using Deep Learning and Differentiating Between Dangerous and Non-Dangerous Fires

Mohamed Youssef[1], Mohamed Marie[2], Sarah Naiem[3]
Department of Computer Science, Faculty of Computer Science and Artificial Intelligence, Capital University, Cairo, Egypt

***Abstract*—Fire is a major hazard in many disasters, creating risks to public safety and the surrounding environment. This study aims to improve the accuracy and reliability of fire detection using deep learning, addressing the key limitations of traditional sensor systems, such as latency and poor adaptability. The proposed model presents a customized Fire-Smoke-YOLOv8x Model trained from scratch on a dataset of 100,000 images representing diverse fire and smoke conditions paired with adaptive algorithms to differentiate dangerous from non-dangerous fire scenarios and achieved a performance with 98.2 % precision, 97.3% recall, and 94% mAP@50. The model processes RGB video under varied lighting and environments and achieves strong detection results across a wide range of fire and smoke scenarios. Incorporating thermal or multispectral data, as noted in future work, could further improve performance under extremely low visibility conditions. This framework supports real-time surveillance in smart cities, transport hubs, and industrial safety, where fast and accurate detection is critical. It combines fine-grained detection with risk-aware post-processing, providing a high-performance and scalable solution for real-world fire detection.**

***Keywords*—Deep learning; fire detection; smoke detection; computer vision; real-time detection; attention mechanism; fire risk classification**

I. INTRODUCTION

Human history has been shaped by a range of natural and human-caused disasters that threaten lives and disrupt communities, often resulting from environmental shifts that have led to widespread loss and destruction. Among these events, fire is especially dangerous due to its rapid spread and unpredictable behavior [1],[2],[3].

Fires are one of the most critical threats to public safety and community growth. They present risks to both cities and rural areas, pose a danger to human life and property, causing fear among people, damaging homes and buildings, destroying farms and forests, spreading quickly in dry conditions, harming animals, reducing air quality, and making recovery slow and difficult for affected communities [4], [5], [6].

Exacerbating this crisis, rising temperatures and long periods without rain make it easier for wildfires to start. Additional things, such as rising sea surface temperatures and extreme weather events, have also been linked to increased fire activity across urban and natural ecosystems [7]. Because of all that, strengthening early fire detection capabilities is becoming critical to emergency response strategies. Efforts have focused on deploying advanced, real-time fire detection systems that provide faster alerts, support early detection, and improve overall emergency response efficiency [8].

Factors such as global warming and land-use changes have contributed to the increased frequency and intensity of fire-related events. Recent years have shown a rise in extreme weather conditions, including heatwaves, which create favorable environments for fire ignition and spread. Many regions have reported unusual and large-scale fires, highlighting the growing need for reliable early detection systems. Early fire identification is critical because rapid detection enables timely intervention, reduces potential damage, and helps protect both lives and property. Effective early warning systems can support firefighters by providing prompt alerts in high-risk areas, by this way limiting the spread of fires before they escalate [9], [10]. To reduce the risk of loss of life and damage to property, it is important to detect fires as early as possible, deploy warning systems in at-risk locations, and intervene quickly to limit fire spread [11],[12]. Conventional fire detection technologies such as heat sensors, smoke sensors, and threshold-based alarm systems have been widely used for decades. However, their performance can be limited under varying environmental conditions, resulting in delayed or unreliable detection [13].

Recent advancements in computer vision provide promising alternatives. Deep Learning Models can analyze complex visual patterns, allowing better detection of flames and smoke under diverse lighting and environmental conditions [14],[15],[16].

II. RELATED WORKS

A wide range of approaches has been proposed for fire and smoke. Traditional fire detection methods typically rely on sensors that are activated only after the fire reaches a certain stage, meaning that detection often occurs after the fire has already spread significantly within the environment [17],[18]. This limitation has motivated researchers to shift toward intelligent vision-based systems, where models are trained to visually detect fire and smoke at earlier stages before widespread propagation occurs [19],[20]. Early detection is critical in reducing damage and improving response time, especially in scenarios where fires may be difficult to recognize in their initial phases [21]. Also, while sensor-based systems are commonly deployed in enclosed environments, their applicability in open or outdoor settings is limited.

This creates a need for intelligent detection systems capable of monitoring such environments using visual data, enabling timely identification and response [22],[23].

To address these limitations, several advanced fire and smoke detection methods have been introduced, as summarized in Table I

TABLE I. SUMMARY OF REPRESENTATIVE PRIOR FIRE AND SMOKE DETECTION STUDIES

	Authors names	Approach Category	Core Technique(s) and Features Used	Reported Strengths	Limitations
1	Chettoui, M., & Akhloufi, M. A. [24]	Deep learning–based fire/smoke detection using YOLOv7/YOLOv8 with transfer learning.	on a >11k image fire–smoke dataset; evaluation at IoU=0.5; trained using transfer learning with optimized hyperparameters	Best model is (YOLOv8x): mAP@0.5 = 92.6, Precision = 95.4, Recall = 84.8.	Evaluation is performed on a still image dataset; no video-based evaluation.
2	Prakash M., Neelakandan, S., Tamilselvi, M., Velmurugan, S., Priya, S. B., & Martinson, E. O. [25]	Deep learning (transfer learning)	Inception-v3 transfer learning to classify satellite/aerial images into fire vs. non-fire	Accuracy = 97.55%, Precision = 94.19%, Recall = 96.44% and F1 = 93.33% on their dataset.	No localization capability Trained and evaluated exclusively on satellite/aerial imagery, limiting generalization to indoor or ground-level scenes.
3	Vignesh R. P., Darshan Kumar P., Rajasekar S., & Prakash K. [26]	Fire and smoke detection using YOLO-based object detection	Real-time fire and smoke detection using YOLOv8 on annotated image datasets	Highlights YOLOv8's suitability for practical deployments (industrial/urban/forest). Demonstrates good accuracy on their custom dataset.	tested only on static images (no video/temporal robustness); limited dataset diversity.
5	S. Rahman, S. M. H. Jamee, J. K. Rafi, J. S. Juthi, S. A. Aziz and J. Uddin [27].	Deep learning (object detection; YOLOv8)	Fine-tune YOLOv8 on a 5,700-image custom dataset (fire/smoke / fire+smoke)	Real-time detection with 68.3% precision, 54.6% recall, and 60.7% F1; Outperformed several conventional CNN baselines.	Small dataset (~5.7k images); limited generalization; modest recall; no segmentation capability; performance drops in complex smoke conditions.
6	Mukhiddinov, M., Abdusalomov, A. B., & Cho, J. [28].	Deep learning (object detection; optimized YOLOv5m)	YOLOv5m enhanced with anchor-box clustering (K-means++), SPPF-plus, feature pyramid enhancement, pruning, and transfer learning;	Achieves AP@50 = 81.5%, AP = 73.6% on UAV-based wildfire smoke	UAV-only wildfire dataset (~6000 frames); limited generalization to indoor/ground fires; no FPS reported; higher model complexity.
7	N. Borges, L. Fonseca, P. S. Barreto, E. A. P. Alchieri, M. F. Caetano, P. A. Resende, L. Brandão and L. Vieira. [29].	Two-phase deep-learning smoke detection integrated in a distributed PTZ-camera network (SEMFOGO-DF)	Phase-1 analyzes panoramic image sequences to flag regions with high smoke probability. Phase-2 automatically zooms the PTZ camera and runs a smoke classifier on the zoomed view to confirm/reject the alert.	Two-phase workflow reduces false alerts from Phase-1 with a small reduction in detection rate; the system is designed for distributed, real-world operation in Brazil's Federal District.	No quantitative precision metrics provided The sequential two-phase process may increase computational overhead and response latency.
8	Y. Xu, J. Li, L. Zhang, H. Liu, and F. Zhang [30]	Deep learning (object detection; YOLOv7 enhanced)	YOLOv7 enhanced with ConvNeXtV2 backbone and CBAM attention (CNTCB-YOLOv7)	Precision 86.18%, Recall 81.85%, AP 88.36%	Detection was evaluated on static images only; no real-time video inference or latency analysis was reported.

III. RESEARCH METHODOLOGY

Deep learning for fire detection relies on large, diverse datasets. Flames vary in color (blue, red, yellow, or sometimes even green, it's different depending on the material/temperature), and smoke is gray, white, and black. Because fires occur in agriculture, infrastructure, homes, and forests, datasets must span various conditions for robust detection. Key public datasets:

BoWFires: 226 images (119 fire and 107 non-fire) at various resolutions [31].

Corsican Fire Dataset: 500 RGB fire images, 100 RGB+NIR pairs, and five RGB–NIR sequences. Each RGB image has a pixel fire mask, and NIR/sequence data mainly for multispectral analysis [32].

Smoke100k: 100,000 images with smoke/no-smoke backgrounds [33].

D-Fire: 21,000+ images in fire, smoke, fire + smoke and background classes, with YOLO-format bounding boxes [34].

VisiFire Annotations: 2,684 frames from 12 public videos, with pixel-level fire segmentation for evaluation [35].

A. Fire-Smoke 100k (FS-100k)

The FS-100k dataset offers a large-scale and diverse benchmark for combined fire–smoke detection. It contains 100,000 images sourced from both real and synthetically generated material. It covers a wide range of flame and smoke characteristics and environmental contexts.

Data Sources and Diversity: Frames were extracted at different intervals from public clips, capturing both indoors (kitchens and garages) and outdoors (wildfires and industrial sites). The dataset includes transparent smoke, flickering flames, night scenes, and off-angle viewpoints, providing a range of visual conditions that commonly occur in real situations. To represent rare or hard-to-capture scenarios, real images were supplemented with synthetic cases created using compositing techniques and generative models.

These cases include:

- large wildfire edges
- Multi-colored flames
- heavy smoke in low light or fog

The synthetic overlays introduce additional variety in smoke density and flame appearance, helping cover situations that may be underrepresented in standard footage.

B. High-Resolution Close-UP Fire Imagery

- To capture fine-grained flame traits, the dataset included high-resolution video frames from a modern smartphone with optical and telephoto sensors. These samples show close-range fire behavior from controlled sources such as:
- Small ignition devices (matchsticks, lighters)
- Steady indoor flames under varied lighting
- Low velocity ignition sources resembling real triggers.
- This enables detailed flame representation, including:
- The Edge texture and Core brightness/ Micro flickers, Heat distortion.

C. Environmental Variation and Image Characteristics

Urban: buildings, vehicles

Rural: fields, farms

Forested: wildlands, parks

Industrial: factories, chemical plants

The weather ranges from sunny to overcast, rainy, and foggy, ensuring diverse backgrounds.

Resolution: 360p–1080p, detailed enough for small flames and faint smoke.

Flame Colors: From deep red to bright yellow, white-hot, occasional blue (gas/chemical), and multi-hued transitions.

Smoke Attributes: From transparent or light-gray plumes to thick black industrial smoke, often occluding backgrounds or overlapping flames.

Examples are shown in Fig. 1–8.

D. Data Collection and Generation

To build a diverse dataset suitable for fire and smoke detection tasks, 100,000 images were collected from multiple sources, including public footage, mobile captures, and synthetically generated samples. Data acquisition emphasized visual diversity, resolution quality, and coverage of varied real-world conditions.

Fig. 1. Indoor fire example in a living room.

Fig. 2. Indoor fire example in a kitchen.

Fig. 3. Outdoor fire example in an urban area.

Fig. 4. Outdoor fire example in a forest.

Fig. 5. Controlled flame example.

Fig. 6. Fire with dense black smoke light.

Fig. 7. Stage example under bright spotlights (non-fire light source).

Fig. 8. Close-up example of a lit matchstick.

1) Public online video footage: Part of the dataset was created by extracting frames from publicly available videos of fire and smoke incidents, including vehicle fires, wildfires, residential, and industrial events. Frames were sampled at different intervals to capture variations in flame appearance, intensity, and smoke development.

2) High-resolution fire imagery via mobile capture: To expand the dataset with both controlled and spontaneous scenes, high-resolution images were captured indoors and outdoors using mobile devices. These photos include varied ignition sources, faint smoke, and diverse lighting conditions, offering close-range details that are not commonly found in web-sourced datasets.

Examples include:

Indoor scenes (candles, cigarettes, lighters)

Outdoor scenes with subtle smoke (burning paper, small debris)

Urban and residential environments (streets, homes, rooftops) are captured at various times of day under natural and artificial lighting.

These variations help incorporate low visibility and early-stage smoke conditions that are often underrepresented in public datasets.

3) Synthetic sample generation: To cover rare or difficult-to-capture scenarios (dense smoke in confined spaces, industrial flare-ups, and flame reflections), synthetic images were generated using generative models and compositing. The selected samples were used to supplement underrepresented cases in the dataset. Each sample, regardless of origin, was manually reviewed for quality, clarity, and relevance before annotation and inclusion in the final dataset (Table II).

TABLE II. CATEGORIES AND ENVIRONMENTAL VARIATIONS

Category	Subcategories/ Variations
Fire Scenarios	- Small flames: candles, matchsticks, lighters - Medium flames: kitchen stoves, barbeques - Large fires: building fires, wildfires
Smoke Variations	-Transparent/light smoke (early ignition, cigarettes) -Dense black smoke (industrial, vehicular) - Hazy smoke under varied lighting (indoor low-light, outdoor daylight)
Background Complexity	- Urban: traffic, buildings, busy streets - Rural/Natural: forests, grasslands, mountains - Indoor: kitchens, living rooms, garages
Lighting Weather	- Lighting: daytime vs nighttime - Weather: rainy, snowy, foggy scenes

This diversity helps expose the model to a wide range of fire and smoke appearances, which can support better performance across varied scene conditions. The FS-100K dataset includes 27,846 fire-only images representing 27.85% of the dataset, 19,624 smoke-only images representing 19.62%, 27,391 images containing both fire and smoke, representing 27.39%, and 25,139 background images without visible fire or smoke, representing 25.14%. The background samples were included to improve the model's ability to distinguish real fire and smoke from visually similar non-hazardous scenes and to reduce false-positive detections.

E. Data Annotation and Labeling

Each image was manually annotated using pixel-level segmentation as illustrated in Fig. 9.

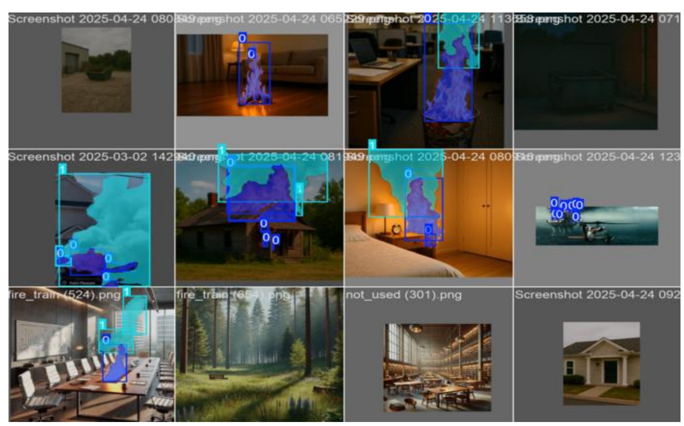


Fig. 9. Pixel-level segmentation illustrating annotated regions of fire and smoke showing fire and smoke masks under different indoor and outdoor conditions.

This fine-grained annotation supports detailed modeling of fire and smoke shapes, including transparent, thin, or partially occluded regions that are poorly represented by bounding boxes. Segmentation masks provide more precise supervision for training both the detection and segmentation components.

Polygon manual annotation was performed, where annotators outlined fire and smoke regions to generate pixel-level masks. Manual refinement was applied for complex cases such as faint smoke, overlapping fire–smoke regions, or low-visibility scenes to maintain annotation quality across diverse conditions.

The model was trained entirely from scratch without pretrained weights, allowing the learned representations to focus specifically on fire and smoke characteristics rather than generic visual features from unrelated datasets.

To evaluate the YOLOv8-FireSmoke model, it performed an assessment on a 20K-image validation split of the FS-100k dataset. The next section presents the evaluation of the YOLOv8-FireSmoke model using both performance metrics and data analysis. In addition to standard measures such as precision–recall curves and F1-score, also analyze the distribution of bounding boxes and object locations in the dataset. The results show that most objects are concentrated near the center of the images and are generally small in size. These patterns are important because they can affect how the model learns and performs. By combining model performance with dataset analysis, this section provides a clearer and more reliable understanding of the model's behavior on the FS-100k validation set.

F. Architecture of the Model

The network consists of three main components: a backbone. The proposed model consists of a backbone for hierarchical feature extraction, a bidirectional neck for multi-scale feature fusion, and a multi-scale decoupled segmentation head for final prediction. The backbone integrates C2fFireMFA blocks to enhance feature representation using multi-scale feature aggregation (MFA), while C2fCBAM blocks are employed in deeper layers to improve attention through channel and spatial refinement. The neck adopts a top-down and bottom-up structure to effectively combine semantic and spatial information across multiple feature levels. Finally, the head operates on four scales (P2–P5) to handle objects of varying sizes, producing bounding box, classification, and segmentation outputs (Fig. 10).

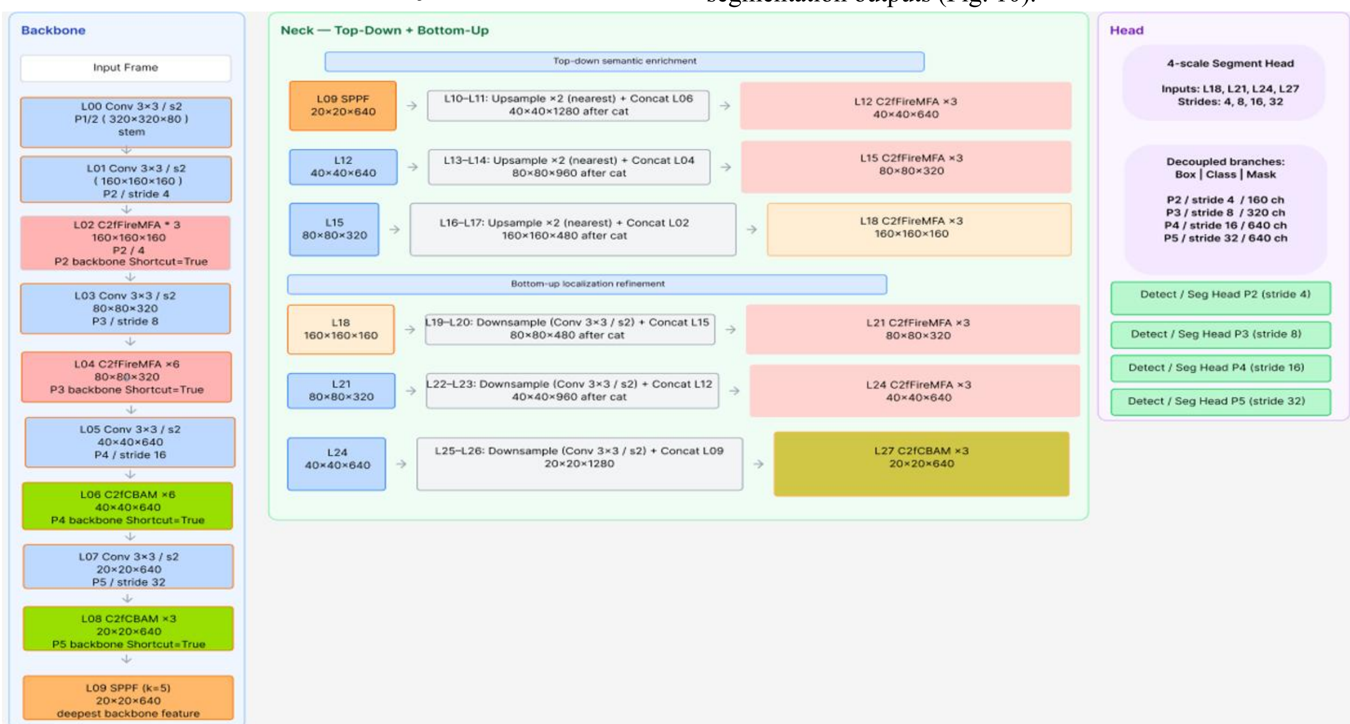


Fig. 10. Model architecture of fire-smoke.

IV. RESULTS AND ANALYSIS

A. Accuracy Table

Table III presents a comparative summary of the key performance metrics of the proposed YOLOv8-FireSmoke model against the baseline YOLO model. The evaluation includes precision, recall, and mAP@50 measured across both fire and smoke detection tasks.

The results show an enhancement across all metrics with the Fire-Smoke model. The precision increased from 94% to 98.2% and the recall from 92% to 97.3%, indicating enhanced detection accuracy and reduced false predictions. Similarly, mAP@0.5 improved from 89% to 94%, it is reflecting better overall localization and detection performance.

These improvements were achieved with a moderate increase in model complexity, as seen in the higher number of layers (232 to 337) and parameters (71.75M to 73.53M), along with a slight rise in computational cost (GFLOPs). Overall, the proposed model achieves a balance between accuracy and efficiency, making it reliable for fire and smoke detection applications and providing a controlled evaluation under the same dataset and evaluation protocol, the original YOLOv8x baseline and the proposed Fire-Smoke model were trained and evaluated using the same FS-100K training and validation splits (Fig. 11). Table III reports the comparison in terms of model complexity and detection performance. This controlled setting allows a direct assessment of the performance improvement achieved by the proposed architectural modifications, as shown in Table III.

TABLE III. CONTROLLED COMPARISON BETWEEN YOLOV8X BASELINE AND THE PROPOSED FIRE-SMOKE MODEL ON THE FS-100K VALIDATION SPLIT

Model	Baseline Of Yolo	Fire-Smoke Model
Layer	232	337
Parameters	71,752,774	73.53M
Precision	94%	98.2%
Recall	92%	97.3%
Map@0.5	89%	94%
GFLOPs	329	333.3

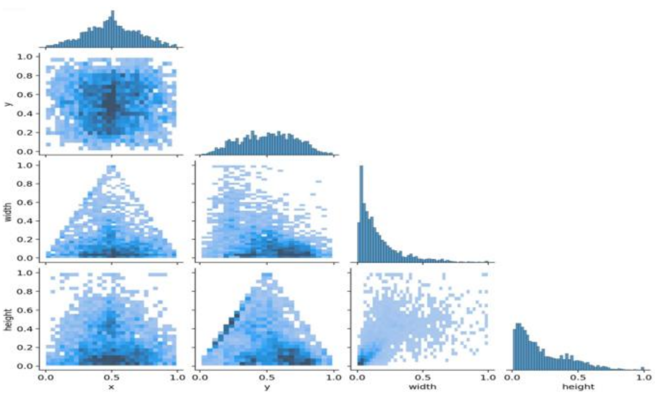


Fig. 11. Joint distributions of object center coordinates and bounding box sizes, highlighting spatial concentration and scale variability across the dataset.

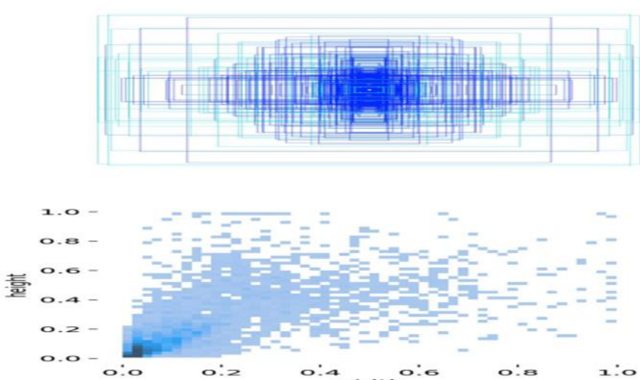


Fig. 12. Statistical analysis of annotated bounding boxes. The upper visualization depicts the spatial aggregation of bounding boxes, indicating a strong central tendency in object placement. The lower heatmap represents the joint distribution of normalized width.

As shown in Fig. 15, the precision and recall of three recent detectors are compared with the results reported in their original papers. Our YOLOv8-FireSmoke model achieves strong performance on the FS100k dataset, with a precision of 98.2% and a recall of 97.3%. Since the compared models were evaluated on different datasets, the comparison provides a general context rather than a strict benchmark. The direct comparison between the model and the YOLOv8x baseline is presented separately in Table III, where both models are evaluated using the same FS-100K validation split and evaluation protocol. Even so, the proposed YOLOv8-FireSmoke model achieves performance levels that are consistent with recent high-performing fire detection models (Fig. 14).

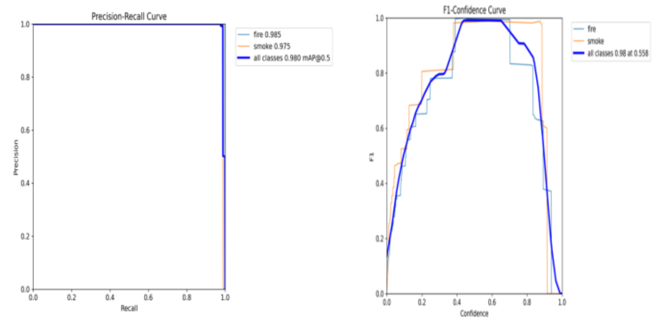


Fig. 13. Dual-metric visualization summarizing detector performance. Left: F1-score versus confidence threshold for fire, smoke, and overall classes—highlighting a peak F1 ≈98% at confidence ≈ 0.55. Right: Precision-recall curves at IoU 0.5.

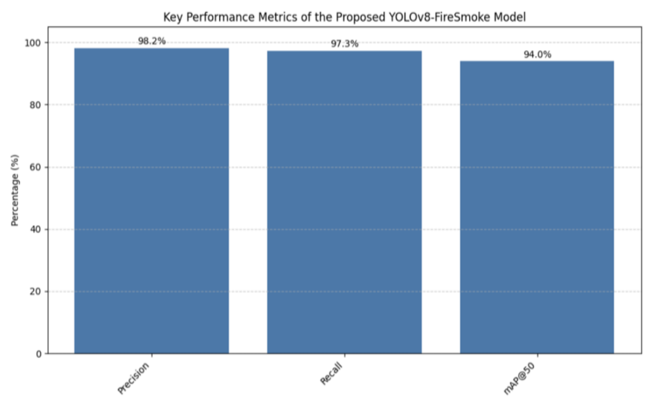


Fig. 14. Bar chart summarizing the proposed YOLOv8-FireSmoke model's precision, recall, and mAP@50 performance.

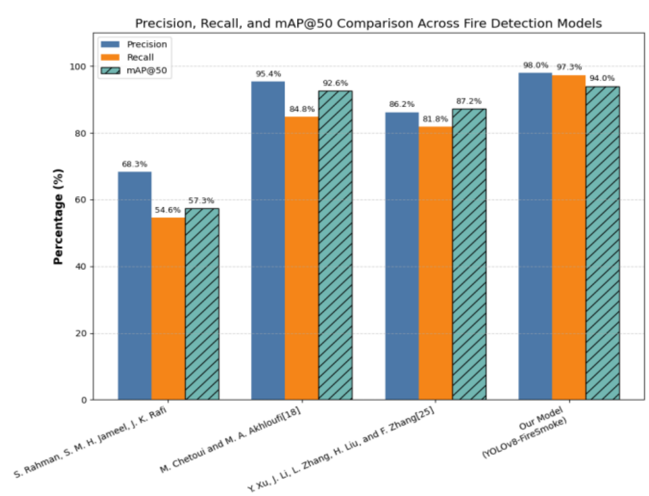


Fig. 15. Comparative evaluation of fire and smoke detection models based on precision, recall, and mAP@50.

B. Model Performance and Configuration

The results indicate that the model can localize and detect fire. The results provide both visual and numerical information about the performance of the YOLOv8-FireSmoke model. From the distribution plots (Fig. 12), object locations are more concentrated near the center of the images, while bounding box sizes vary. This suggests that the dataset contains objects at different positions and scales.

The performance curves (Fig. 13) show that the model maintains a high F1-score across different confidence levels, with a peak of about 0.98 at a confidence of around 0.55. The precision-recall curves for fire and smoke also indicate stable performance without sharp drops.

Based on Table III, the proposed model performs better than the baseline. Precision increases from 94% to 98.2%, recall from 92% to 97.3%, and mAP@0.5 from 89% to 94%. This improvement comes with a small increase in computational cost.

The training configuration was selected to support stable learning and improve the model's ability to detect fire and smoke under different visual conditions. The model was trained for 100 epochs using an input image resolution of 768 × 768 pixels, which provides sufficient spatial detail for detecting small fire and smoke regions. A batch size of 16 was used to balance training stability with memory requirements.

AdamW was selected as the optimizer because it combines adaptive gradient-based optimization with decoupled weight decay, helping to control overfitting during training. The initial learning rate was set to 1×10^{-3}, and a cosine learning-rate schedule was applied to gradually reduce the learning rate as training progressed. A 5-epoch warm-up stage was used at the beginning of training with a warm-up momentum of 0.9 and a warm-up bias learning rate of 0.01, allowing the model to start learning gradually before full optimization. The momentum parameter was set to 0.937, and the weight decay was set to 5×10^{-4}. Data augmentation included HSV color adjustment, horizontal flipping, translation, scaling, mosaic augmentation, copy-paste augmentation, and, for segmentation, overlapping masks were enabled, and the mask ratio was set to 4. The random seed was set to 42.

Overall, the results suggest that the YOLOv8-FireSmoke framework may be suitable for real-time applications such as smart building monitoring, wildfire observation systems, industrial safety, and UAV-based inspection, where timely detection is important.

C. Threat Level Algorithm

A threat level post-processing step was designed to differentiate small, controlled ignition sources from larger Dangerous fires. The algorithm evaluates factors such as flame size, surrounding smoke density, and the spatial relationship between fire and smoke. Small, localized ignition sources such as lighters, matchsticks, or brief household ignition events are treated as low threat cases, while larger fires with substantial smoke development are assigned higher threat levels.

This refinement helps reduce false alarms in scenes where minor ignition sources are visually similar to early-stage fire events.

1) Threat level post-processing: A post-processing module was introduced to differentiate small, controlled ignition sources (lighters, matchsticks, brief household ignition events) from larger fire events that exhibit substantial flame growth or smoke development. The module uses visual cues such as flame size, spatial extent, and surrounding smoke density to assign a threat category to each detected fire region. This additional step helps reduce false alarms in situations where small ignition sources may appear visually similar to early-stage fire events.

2) Step of threat level algorithm: Once a fire region is detected by the model, a secondary analysis evaluates contextual information to refine its threat category. The refinement considers attributes such as the relative size of the flame, proximity and density of visible smoke, and the spatial consistency of the fire across frames. Based on these cues, each detected fire instance is assigned to one of three levels, as illustrated in Fig. 16–19.

3) The decision is based on: The threat level estimation uses three main cues extracted from the detection output:

Fire area (number of pixels of the fire).

Smoke proximity: Distance from the fire and smoke.

Smoke Density.

4) Parameter definitions: The T-safe fire area threshold separates a small harmless flame from a large, dangerous flame.

D-crit – maximum center-to-center distance, in pixels, for treating smoke near the fire.

S-safe minimum smoke area ratio regarded as dense, indicating active combustion

D. Threat Level Rules

1) Normal fire – No threat

As shown in Fig. 16.

Fig. 16. Testing in real time for the proposed model for (normal fire).

Condition fire area < T-safe and (no smoke within D-crit or smoke ratio < S-safe).

Examples: candle flame, lighter, matchstick.

2) Dangerous fire – High threat

As shown in Fig. 17.

Fig. 17. Testing in real time for the proposed model (normal fire and dangerous fire).

Condition: fire area ≥ T-safe and (no smoke within D-crit or smoke ratio < S- safe).

Examples: stove flare-up, unattended large flame.

3) Real Threat

As shown in Fig. 18-19.

Fig. 18. Testing in real time for the proposed model (real threat)(a).

Fig. 19. Testing in real time for the proposed model (real threat)(b).

Condition: centroid distance between fire and smoke ≤ D-crit, and either fire area ≥ T-safe or smoke area ≥ T-safe. Examples: small ignition producing thick smoke, large fire with nearby smoke, spreading indoor fire.

As shown in: https://youtu.be/JXU0-0EI7Ns

The threat-level classification in the current experimental configuration was based on the mask area of the detected fire and smoke regions and on the spatial relationship between them. The safe area threshold, T-safe, was set to 1000 mask pixels. Therefore, any isolated fire or smoke region with a mask area below 1000 pixels was treated as a small, low-risk detection, while any isolated fire or smoke region with a mask area equal to or greater than 1000 pixels was treated as a dangerous, high-risk detection. The critical distance threshold, D-crit, was set to 90 pixels. This value was used as the maximum centroid-to-centroid distance for considering a fire region and a smoke region spatially associated. When fire and smoke were detected within this distance, they were treated as a related fire-smoke pair. If both regions in the pair had mask areas below T-safe, the case was classified as a normal low-risk fire-smoke detection. However, if the fire and smoke regions were spatially associated and one of them had a mask area equal to or greater than T-safe, the case was classified as a real threat. Before threat classification, same-class detections were merged. Fire regions were merged with fire regions, and smoke regions were merged with smoke regions, when their IoU was 0.5 or higher or when the centroid distance between them was 30 pixels or less. The final threat level was then assigned after calculating the merged mask areas and centroid distances. If multiple detections were present in the same frame, the highest detected threat level was assigned to the frame. The FS-100K dataset, containing over 100,000 images, is available from the corresponding author upon reasonable request.

E. Threat Level Algorithm Flowchart (Fig. 20)

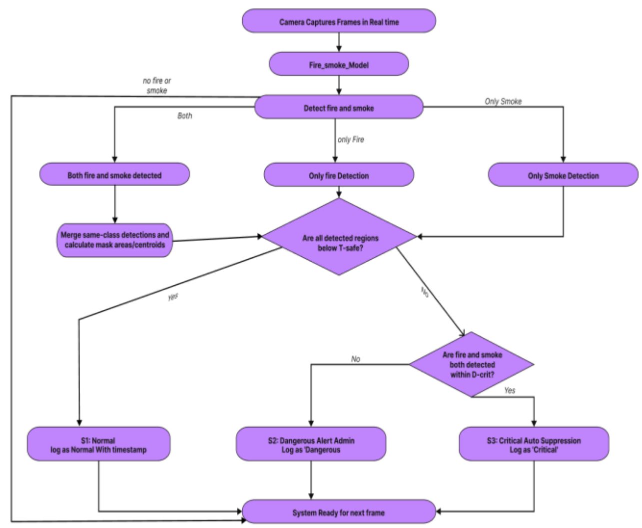


Fig. 20. An overview of the proposed intelligent fire detection threat level algorithm.

V. Conclusion

This work presented YOLOv8-FireSmoke, a real-time deep learning model for fire detection. The approach integrates three main components:

An optimized architecture consisting of a backbone, neck, and multi-scale detection head. The backbone is responsible for hierarchical feature extraction, while the neck employs a top-down and bottom-up feature aggregation strategy to enhance semantic and spatial information across multiple scales. The detection head further processes these features to enable accurate localization.

Three-class post-processing logic that separates normal ignition sources, dangerous fires, and fires–smoke threats. This step helps reduce false alarms while ensuring that hazardous events are escalated quickly.

FS-100K dataset, consisting of 100,000 diverse images from real and synthetic sources and covering indoor, outdoor, low light, daylight, and challenging edge case scenarios. On a 20,000-image validation split, the proposed model achieved 94% mAP@50 and precision above 98%.

Real-time inference speed, processing video at approximately 57 FPS on a single RTX 4070 Super GPU, including post-processing. This performance exceeds the requirements of real-time surveillance and early warning systems.

Overall, the results show that combining rich training data with context-sensitive post-processing yields an accurate and practical framework suitable for deployment in urban, industrial, and wildland environments. Vision-based systems also offer advantages over fixed heat or smoke sensors, including earlier visual response and fewer nuisance alarms in complex spaces. As fire risk increases in many regions, deep learning–based detection may support the transition from reactive response to more proactive monitoring. Future work may incorporate thermal or multispectral inputs to further improve robustness in extremely low visibility conditions.

VI. Future Work

Multispectral and Thermal Fusion.

Integrate infrared or thermal cameras with RGB streams to improve detection at low visibility in nighttime or very hard smoke conditions in darkness.

3D Localization and Multi Camera Geometry.

Employ 3D bounding boxes or stereo/multi-camera rigs to refine distance and volumetric growth estimates, enabling better threat assessment and resource allocation.

Predictive spread Modelling decouples the visual detector with atmospheric and fuel load data to forecast the likely spread direction or growth rate, turning early warning into actionable predictions.

References

[1] O. Haas, I. C. Prentice, and S. P. Harrison, “Wildfires on a changing planet,” Nat. Commun., vol. 17, p. 1599, 2026, doi: 10.1038/s41467-025-68176-4.

[2] C. X. Cunningham et al., “Wildfires will intensify in the wildland–urban interface under near-term warming,” Commun. Earth Environ., vol. 6, p. 542, 2025, doi: 10.1038/s43247-025-02475-y.

[3] D. I. Kelley et al., “State of Wildfires 2024–2025,” Earth Syst. Sci. Data, vol. 17, pp. 5377–5488, 2025, doi: 10.5194/essd-17-5377-2025.

[4] S.-Y. Kim and A. Muminov, “Forest fire smoke detection based on deep learning approaches and unmanned aerial vehicle images,” Sensors, vol. 23, no. 12, p. 5702, 2023, doi: 10.3390/s23125702.

[5] J. Sandoval Díaz et al., “Community resilience to wildfires: A systematic review of impacts, coping strategies, indicators, and governance challenges,” Prog. Disaster Sci., vol. 27, p. 100447, 2025, doi: 10.1016/j.pdisas.2025.100447.

[6] Z. A. Mani, A. Khorram-Manesh, and K. Goniewicz, “Global health impacts of wildfire disasters from 2000 to 2023: A comprehensive analysis of mortality and injuries,” Disaster Med. Public Health Prep., vol. 18, p. e230, 2024, doi: 10.1017/dmp.2024.150.

[7] S. P. Naik, V. Lohbande, S. Hambir, R. Korade, and R. Hatkar, “Fire detection with image processing,” Int. J. Res. Appl. Sci. Eng. Technol., vol. 11, no. 5, pp. 2123–2128, 2023, doi: 10.22214/ijraset.2023.52073.

[8] C. Jin, T. Wang, N. Alhusaini, S. Zhao, H. Liu, K. Xu, and J. Zhang, “Video fire detection methods based on deep learning: Datasets, methods, and future directions,” Fire, vol. 6, no. 8, p. 315, 2023, doi: 10.3390/fire6080315.

[9] J. Carnicer, A. Alegría, C. Giannakopoulos, F. Di Giuseppe, A. Karali, N. Koutsias, P. Lionello, M. Parrington, and C. Vitolo, “Global warming is shifting the relationships between fire weather and realized fire-induced CO_2 emissions in Europe,” Sci. Rep., vol. 12, p. 10365, 2022, doi: 10.1038/s41598-022-14480-8.

[10] S. McElroy, L. Schwarz, H. Green, I. Corcos, K. Guirguis, A. Gershunov, and T. Benmarhnia, “Defining heat waves and extreme heat events using sub-regional meteorological data to maximize benefits of early warning systems to population health,” Sci. Total Environ., vol. 721, p. 137678, 2020, doi: 10.1016/j.scitotenv.2020.137678.

[11] A. Salaria and A. Singh, “A guide for selection of wireless communication technology for effective and robust early forest fire detection system,” Bull. Electr. Eng. Inform., vol. 14, no. 2, pp. 1026–1035, 2025, doi: 10.11591/eei.v14i2.8613.

[12] M. Bugarić, D. Krstinić, L. Šerić, and D. Stipanicev, “Current trends in wildfire detection, monitoring and surveillance,” Fire, vol. 8, no. 9, p. 356, 2025, doi: 10.3390/fire8090356.

[13] S. Khan and A. Khan, “FFireNet: Deep learning-based forest fire classification and detection in smart cities,” Symmetry, vol. 14, no. 10, p. 2155, 2022, doi: 10.3390/sym14102155.

[14] B. M. Diaconu, “Recent advances and emerging directions in fire detection systems based on machine learning algorithms,” Fire, vol. 6, no. 11, p. 441, 2023, doi: 10.3390/fire6110441.

[15] J. Pesonen et al.,Detecting wildfires on UAVs with real-time segmentation trained by larger teacher models, in Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV), 2025.

[16] P. Barmpoutis, P. Papaioannou, K. Dimitropoulos, and N. Grammalidis, “A review on early forest fire detection systems using optical remote sensing,” Sensors, vol. 20, no. 22, p. 6442, 2020, doi: 10.3390/s20226442.

[17] N. D. Ismail, R. Ramli, and M. N. A. Rahman, “A systematic literature review of vision-based fire detection, prediction and forecasting,” Jurnal Kejuruteraan, vol. 37, no. 1, pp. 191–218, 2025, doi: 10.17576/jkukm-2025-37(1)-14.

[18] S. A. Elhanashi, S. Essahraui, P. Dini, and S. Saponara, “Early fire and smoke detection using deep learning: A comprehensive review of models, datasets, and challenges,” Appl. Sci., vol. 15, no. 18, p. 10255, 2025, doi: 10.3390/app151810255.

[19] L. Deng, S. Wu, S. Zou, and Q. Liu, “Large-space fire detection technology: A review of conventional detector limitations and image-based target detection techniques,” Fire, vol. 8, no. 9, p. 358, 2025, doi: 10.3390/fire8090358.

[20] U. U. Deshpande et al., “Real-time fire and smoke detection system for diverse indoor and outdoor industrial environmental conditions using a vision-based transfer learning approach,” Front. Comput. Sci., vol. 7, p. 1636758, 2025, doi: 10.3389/fcomp.2025.1636758.

[21] L. He, Y. Zhou, L. Liu, Y. Zhang, and J. Ma, “Research and application of deep learning object detection methods for forest fire smoke recognition,” Sci. Rep., vol. 15, p. 16328, 2025, doi: 10.1038/s41598-025-98086-w.

[22] A. Abozeid and R. Alanazi, “An intelligent approach for early smoke/fire detection using vision sensors in smart cities,” Sci. Rep., vol. 16, p. 11387, 2026, doi: 10.1038/s41598-026-42762-y.

[23] S. Danish et al., “Vision-based fire management system using autonomous unmanned aerial vehicles: A comprehensive survey,” Artif. Intell. Rev., vol. 59, p. 16, 2026, doi: 10.1007/s10462-025-11415-3.

[24] M. Chetoui and M. A. Akhloufi, “Fire and smoke detection using fine-tuned YOLOv8 and YOLOv7 deep models,” Fire, vol. 7, no. 4, p. 135, 2024, doi: 10.3390/fire7040135.

[25] M. Prakash, S. Neelakandan, M. Tamilselvi, S. Velmurugan, S. B. Priya, and E. O. Martinson, “Deep learning-based wildfire image detection and

classification systems for controlling biomass," Int. J. Intell. Syst., vol. 2023, Art. no. 7939516, pp. 1–18, 2023, doi: 10.1155/2023/7939516.

[26] R. P. Vignesh, P. Darshan Kumar, S. Rajasekar, and K. Prakash, "Fire and smoke detection using the YOLO algorithm," Int. J. Mod. Eng. Res., vol. 14, no. 3, pp. 26–30, 2024.

[27] S. Rahman, S. M. H. Jamee, J. K. Rafi, J. S. Juthi, S. A. Aziz, and J. Uddin, "Real-time smoke and fire detection using You Only Look Once v8-based advanced computer vision and deep learning," Int. J. Adv. Appl. Sci., vol. 13, no. 4, pp. 987–999, 2024, doi: 10.11591/ijaas.v13.i4.pp987-999.

[28] M. Mukhiddinov, A. B. Abdusalomov, and J. Cho, "A wildfire smoke detection system using unmanned aerial vehicle images based on the optimized YOLOv5," Sensors, vol. 22, no. 23, p. 9384, 2022, doi: 10.3390/s22239384.

[29] N. Borges et al., "A fire management intelligent system for the Brazilian Cerrado biome based on a deep learning two-phase detection method," J. Reliab. Intell. Environ., vol. 11, no. 1, p. 5, 2025, doi: 10.1007/s40860-025-00244-4.

[30] Y. Xu, J. Li, L. Zhang, H. Liu, and F. Zhang, "CNTCB YOLOv7: An effective forest fire detection model based on ConvNeXtV2 and CBAM," Fire, vol. 7, no. 2, p. 54, 2024, doi: 10.3390/fire7020054.

[31] D. Y. T. Chino, L. P. Avalhais, J. F. Rodrigues Jr., and A. J. M. Traina, "BoWFire dataset: Fire and non-fire still images for pixel-based detection," Univ. of São Paulo (GBDI-ICMC). [Online]. Available: https://bitbucket.org/gbdi/bowfire-dataset/downloads/

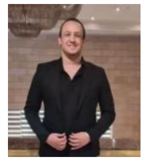
[32] Environmental Sciences UMR CNRS 6134 SPE Lab., "Corsican Fire Database," Univ. of Corsica Pasquale Paoli. [Online]. Available: https://cfdb.univ-corse.fr/

[33] H.-Y. Cheng, J.-L. Yin, B.-H. Chen, and Z.-M. Yu, "Smoke100k: A database for smoke detection." [Online]. Available: https://github.com/BigMM/Smoke100k

[34] D-Fire, "An image dataset for fire and smoke detection," GitHub repository. [Online]. Available: https://github.com/gaiasd/DFireDataset

[35] W. Wahyono, A. Harjoko, A. Dharmawan, F. D. Adhinata, G. Kosala, and K. H. Jo, "Region-based annotation data of fire images for intelligent surveillance systems," Zenodo, 2022, doi: 10.5281/zenodo.5893854.

AUTHORS' PROFILE



Mohamed Youssef, He earned his B.Sc. in Computer Science and AI from MTI University and is currently pursuing an M.Sc. in Computer Science, specializing in Data Science. He also holds a Master of Business Administr action (MBA) from New Mexico and a professional cybersecurity (CC) certificate. His background spans data science, AI, and cybersecurity, with research interests in deep learning, computer vision, and AI applications.

MOHAMED I. MARIE is a Professor of Information Systems at the Faculty of Computers and Artificial Intelligence, Capital University. He received his B.Sc. degree from Cairo University in 1994, followed by his M.Sc. and Ph.D. degrees from Helwan University in 1999 and 2004, respectively. He has held a range of academic and administrative positions throughout his career, including his service at Jazan University from 2012 to 2019. His scholarly contributions include more than 30 published research papers, in addition to reviewing over 70 manuscripts for academic venues. He also maintains research profiles and academic records on several platforms, including Publons, Web of Science, ORCID, Academia, and Google Scholar.

Sarah Naiem, she is a faculty member at the Faculty of Computer Science and AI, Capital Unive rsity, where she lectures and conducts research in computer science and information systems. Her interests include cloud computing, data science, text and opinion mining, software engineering, and information systems.