Real-Time Fire and Smoke Detection Using YOLOv8 for Intelligent Indoor Surveillance

Neepa Christy N, Sanjana A N, Rishwanth S V, S Shobika, Dr Pradeepa K

Students, Artificial Intelligence and Data Science, KPR College of Arts Science and Research, Coimbatore, India

Dean, School of Intelligent Systems and Data Science, KPR College of Arts Science and Research, Coimbatore, India

E-mail: neepachristy14@gmail.com, sanjanaancbe@gmail.com, rishwanth1705@gmail.com, shobika.sivakumar22@gmail.com, pradeepa.k@kprcas.ac.in

Abstract - Fire accidents pose a serious threat to human life, property, infrastructure, and the environment, making early and reliable detection essential for effective emergency response. Conventional fire detection systems primarily depend on smoke, heat, and temperature sensors, which may experience limitations in detecting visible flames at an early stage and can generate false alarms under varying environmental conditions. This research proposes an intelligent real-time fire and smoke detection system using the YOLOv8 (You Only Look Once Version 8) deep learning algorithm for automated visual surveillance. The proposed framework utilizes the Home Fire Dataset, containing diverse fire and smoke images captured under varying lighting conditions, object scales, backgrounds, and indoor environments. The dataset undergoes preprocessing, annotation verification, augmentation, training, and validation to improve model robustness and generalization. The YOLOv8 model is trained using GPU acceleration in Google Colab to simultaneously detect and localize fire and smoke regions through bounding boxes and confidence scores. Performance is evaluated using Precision, Recall, F1-score, mAP@50, mAP@50:95, and inference speed to assess detection accuracy, localization capability, and real-time performance. The proposed system aims to minimize false detections and missed hazards while enabling rapid identification of fire-related threats. Its vision-based architecture can be integrated with CCTV surveillance, IoT-enabled alarm systems, emergency notification platforms, and smart-building technologies. The framework has potential applications in residential buildings, educational institutions, hospitals, offices, warehouses, factories, hotels, commercial complexes, and other indoor environments requiring continuous monitoring. By combining deep learning and real-time computer vision, the proposed approach provides a scalable and cost-effective foundation for intelligent fire safety, supporting early hazard identification, faster emergency intervention, improved situational awareness, and safer living and working environments.

Keywords: Fire Detection, Smoke Detection, YOLOv8, Deep Learning, Computer Vision, Real-Time Object Detection

I. Introduction

Fire is a serious safety concern in residential and indoor environments because flames and smoke can develop rapidly and cause damage to people, buildings, and property. Early identification of a fire is therefore important for enabling timely action. However, detecting fire from visual information is challenging because the appearance of flames and smoke can vary depending on lighting conditions, background objects, distance, camera position, and the stage of the fire. In particular, small flames and early-stage smoke may occupy only a limited portion of an image and can be difficult to distinguish from visually similar objects or environmental conditions. Traditional fire detection systems commonly use physical sensors such as smoke, temperature, and light sensors. Although these systems are useful, visual monitoring through cameras provides additional information about the location and appearance of a possible fire. Computer vision and deep learning provide an approach for automatically analysing such visual information. Object detection models can learn visual features from labelled images and identify the location and class of objects using bounding boxes. This makes AI-based detection a suitable direction for developing automated fire and smoke monitoring systems. Recent research has demonstrated the application of deep-learning-based object detection models to fire and smoke detection, including YOLO-based approaches. (MDPI)

Several computer-vision and deep-learning approaches have been proposed to improve fire and smoke detection. Islam and Habib (2023), for example, developed a YOLOv5-based fire detection approach for different fire scenes and reported its ability to detect small fire and smoke targets. (arXiv) Zhu et al. (2023) proposed FSDNet, combining YOLOv3 and DenseNet-based components for fire and smoke detection in complex scenarios. (arXiv) Geng et al. (2024) proposed

YOLOFM, an improved YOLOv5n-based method intended to address issues such as missed detections, computational complexity, and detection accuracy. (Nature) Other studies have investigated improved YOLOv8 models for fire and smoke detection, including modifications aimed at improving detection under challenging conditions. (Colab) Chetoui and Akhloufi (2024) also evaluated fine-tuned YOLOv8 and YOLOv7 models for fire and smoke detection. (MDPI) More specifically, Peng and Kim (2025) highlighted the limited availability of datasets representing indoor and early-stage home fires and introduced the Home-fire Dataset for this application. (Directory of Open Access Journals) These studies show the potential of YOLO-based object detection while also indicating challenges related to small targets, environmental variation, false detections, and the complexity of indoor scenes. Therefore, a practical detection approach using an appropriate dataset and an efficient object detection model is required. In this project, the Home Fire Detection Dataset was used to develop a fire and smoke detection system using YOLOv8. The purpose of this work is to train and evaluate a YOLOv8 model to detect the two target classes, flame and smoke, from indoor fire images and to examine its suitability as a computer-vision-based approach for early fire detection.

II. Related Works

Table 1. Comparative summary of Existing Fire and Smoke Detection Research Works

Author	**Method**	**Accuracy**	**Limitations**
Huang et al. [1]	YOLOv3-based fire detection	92.8%	Difficulty in detecting fire in complex backgrounds.
Wang et al. [2]	YOLOv2-based fire detection	93.6%	Limited performance for small and complex fire targets.
Shen et al. [3]	CNN-based fire detection	91.0%	Performance can decrease under low-light and complex background conditions.
Zhu et al. [4]	FSDNet using YOLOv3 and DenseNet	99.82%	Uses an older YOLO architecture and may vary across different datasets.
Zhao et al. [5]	YOLOv8-based fire and smoke detection	93.5%	Variations in fire shape, size and intensity can affect detection.
Chetoui & Akhloufi [6]	Fine-tuned YOLOv8	98.9%	Smoke-like objects and difficult environmental conditions can cause detection errors.

Dinata et al. [7]	YOLOv5s with 3D position estimation	92.0%	Requires additional cameras and 3D processing, increasing system complexity.
Catargiu et al. [8]	YOLO10-based fire and smoke detection	89.0%	Performance can decrease under night-time and complex environmental conditions.
Alkhammash [9]	YOLO-based fire and smoke detection	95.0%	Mainly focused on wildfire conditions and may have limited indoor generalization.
Channel Attention Study [10]	Attention-based deep learning model	97.0%	Environmental variations and different fire/smoke appearances can affect performance.

Based on the analysis of the existing techniques as shown in Table 1, it can be observed that various deep learning and computer vision techniques have been widely used for automatic fire and smoke detection. Existing YOLO-based methods provide promising accuracy and fast detection, making them suitable for real-time surveillance applications. However, these techniques still have certain limitations, such as difficulty in detecting small or faint fire and smoke regions, variations in lighting conditions, complex backgrounds, and false detection caused by objects that visually resemble smoke or fire. Some existing approaches are mainly developed for wildfire, forest, UAV, or specific industrial environments, which may limit their effectiveness in general indoor environments. Therefore, there is a need for an efficient real-time system that can simultaneously detect both fire and smoke under different indoor conditions. To overcome these limitations, the proposed work uses the YOLOv8 deep learning algorithm with the Home Fire Dataset for real-time fire and smoke detection.

III. Proposed Work

The proposed work develops an intelligent real-time fire and smoke detection framework using the YOLOv8 deep learning algorithm for early identification of indoor fire hazards. The methodology begins with the Home Fire Dataset containing annotated flame and smoke images collected under diverse indoor conditions. The dataset undergoes preprocessing, quality verification, and annotation validation before being divided into training, validation, and testing sets. YOLOv8 is then trained to detect and localize fire and smoke under varying scales, lighting conditions, and backgrounds. The trained model is evaluated using **Precision, Recall, F1-score, mAP@50, mAP@50:95, and inference speed** to measure its detection performance and real-time capability.

Finally, the optimized model analyses camera or video frames and identifies fire and smoke using bounding boxes and confidence scores, supporting integration with CCTV, IoT-based alarms, and emergency notification systems for rapid hazard detection and response.

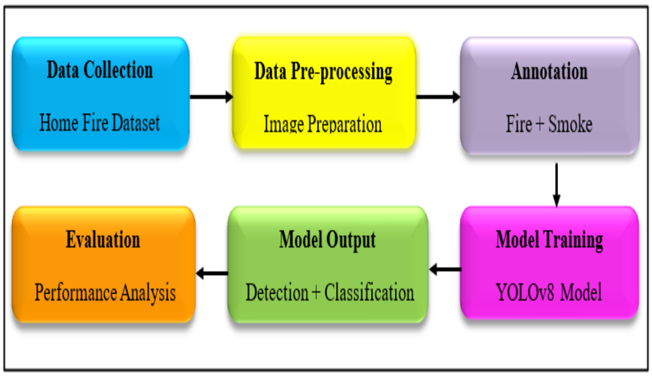

Figure 1. Workflow of the proposed smoke and fire detection system

A. Data Collection

The Home Fire Dataset, introduced by Peng and Kim for early detection of indoor fires, is utilized as the primary dataset for developing and evaluating the proposed YOLOv8-based fire and smoke detection system. The dataset contains 6,500 images with corresponding YOLO-compatible bounding-box annotation files and is divided into three predefined subsets: 3,900 training images, 1,300 validation images, and 1,300 testing images, following a 6:2:2 ratio. The dataset contains two target categories, fire/flame and smoke, enabling the detection model to simultaneously localize and classify fire-related regions.

The dataset is specifically designed to address the challenges associated with early-stage indoor fire detection and includes small-scale flames, subtle smoke, diverse lighting conditions, complex indoor backgrounds, and images captured from typical home-surveillance perspectives. The images were collected from publicly available sources, including Pexels, Pixabay, YouTube, and BiliBili, together with personal or private video contributions. According to the original dataset publication, approximately 400 videos and photographs were reviewed, processed, and converted into labeled image data using LabelImg. The predefined training, validation, and testing partitions are retained in this study to maintain a consistent experimental protocol and prevent unnecessary modification of the original dataset distribution. The training subset is used for model learning, the validation subset for monitoring and model selection, and the independent test subset for final performance evaluation. The diversity of the dataset provides challenging visual conditions for evaluating the capability of YOLOv8 to identify both visible flames and early-stage smoke in indoor environments.

Table 2. Shows the distribution of the dataset into training, validation, and testing subsets used for model development and evaluation.

Dataset Subset	Number of Images	Percentage
Training	3,900	60%
Validation	1,300	20%
Testing	1,300	20%
Total	6,500	100%

Table 3. Summarizes the key characteristics of the Home Fire Dataset used in this study.

Characteristic	Description
Total Images	6,500
Target Classes	Fire / Flame, Smoke
Annotation Format	YOLO
Annotation Type	Bounding Boxes
Environment	Indoor/ Household
Detection Focus	Early Fire and Smoke

B. Data Pre-Processing

Before training the YOLOv8 model, the collected fire and smoke images undergo preprocessing to ensure consistent input quality and compatibility with the detection framework. First, all valid images are resized to **640 × 640 pixels**, providing a standardized input dimension for YOLOv8 while preserving the spatial information required for object localization. The resizing operation is expressed as:

$$I_{\text{resized}} = \text{Resize}(I_i, 640, 640)$$

where I_i represents the original image and I_{resized} represents the resulting image with a resolution of 640 X 640 pixels.

The pixel intensity values are normalized to the range [0,1] to provide a consistent numerical representation of image data. The normalization is defined as:

$$I_{\text{normalized}}(x, y) = \frac{I_{\text{resized}}(x, y)}{255}$$

where I_{resized} (x,y) denotes the pixel intensity at coordinates (x,y). To improve the model's ability to generalize to different indoor surveillance conditions, data augmentation is applied to the training images. Horizontal flipping, scaling, translation, and controlled brightness and contrast variations are used to introduce realistic changes in appearance while preserving the fire and smoke characteristics. For horizontal flipping, the transformation can be represented as:

$$I_{\text{flip}}(x, y) = I_{\text{normalized}}(W - x, y)$$

where W represents the image width. The validation and testing images are kept separate from training augmentation to provide an unbiased evaluation of the trained model. Finally, the image-label correspondence, class identifiers, and bounding-box coordinates are verified before the processed dataset is supplied to the YOLOv8 training pipeline.

C. Data Annotation

The dataset used for the proposed smoke and fire detection system contained two object categories: **Fire** and **Smoke**. The objective of annotation was to provide object-level information required by YOLOv8 for localization and classification.

If the dataset obtained from **Kaggle was already annotated**, the existing annotations were used rather than manually creating new labels. The annotations were checked and prepared to ensure compatibility with the YOLOv8 annotation format before model training. The verification process involved checking the class labels, corresponding image files, bounding-box coordinates, and the consistency between images and annotation files. Incorrect, missing, or incompatible annotations, if identified during preparation, were corrected or removed before training.

For each detected object, a **bounding box** was used to specify its location within the image. The two class categories were assigned class IDs as follows:

Class ID	**Class**
0	Fire
1	Smoke

YOLO annotation files represent each object using the following format:

$$\text{class\_id} \quad x_{center} \quad y_{center} \quad width \quad height$$

where all bounding-box coordinates are normalized with respect to the image dimensions. Thus, each image is associated with a corresponding label file containing the class ID and normalized bounding-box coordinates of the objects.

The final evaluation output contained **1,300 images and 1,580 object instances**, consisting of **963 Fire instances and 617 Smoke instances**. These labelled samples were used by the YOLOv8 model to learn the visual characteristics and spatial locations of fire and smoke.

D. Model Selection and Training

a) Model Selection

YOLOv8

YOLOv8 (You Only Look Once version 8) is a single-stage object-detection framework designed to simultaneously perform **object localization and classification**. Instead of requiring separate stages for region proposal and classification, YOLOv8 processes an input image through a single detection pipeline and produces bounding boxes, class predictions, and confidence scores.

For this study, the **YOLOv8m (medium) variant** was selected for smoke and fire detection. The model was chosen because the proposed application requires accurate localization of fire and smoke while also maintaining sufficiently fast inference for potential real-time monitoring.

The selection of YOLOv8m was based on the following project-specific considerations:

1. **Object-detection capability:** YOLOv8 can simultaneously locate and classify fire and smoke objects.
2. **Real-time suitability:** The model provides fast inference, which is important for early detection applications.
3. **Balance between speed and accuracy:** The medium variant provides a suitable compromise between computational requirements and detection performance.
4. **Transfer learning capability:** Pre-trained weights can be used to accelerate learning on the target dataset.
5. **Ease of training and deployment:** YOLOv8 is supported by the Ultralytics framework and provides a straightforward training and inference pipeline.
6. **Multiple-object detection:** The model can detect multiple fire and smoke regions within the same image.

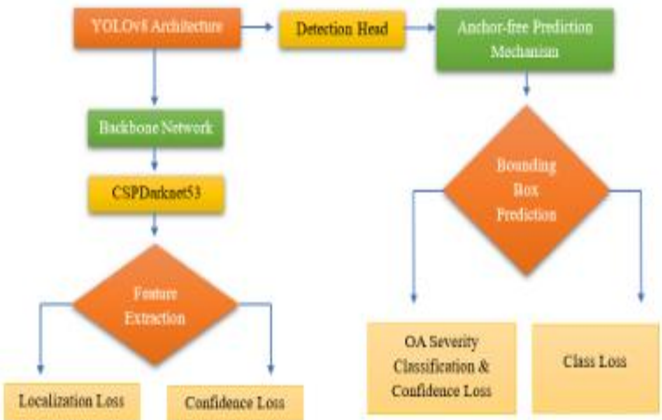

Figure 2. YOLOv8 model extracts important features from input images and predicts the location and class of fire or smoke using bounding boxes and confidence scores.

b) YOLOv8 Architecture

The YOLOv8m model consists of three major functional components: the **backbone, neck, and detection head**.

Backbone

The backbone extracts hierarchical visual features from the input image. YOLOv8 uses convolutional feature extraction with **C2f modules** and an **SPPF (Spatial Pyramid Pooling–Fast)** component. These components enable the model to learn features at different levels of abstraction, which is useful for detecting objects with varying sizes and appearances.

Neck

The neck combines and aggregates features from different stages of the backbone. YOLOv8 uses a feature-pyramid-based structure to facilitate the detection of objects at different scales. This is particularly relevant to fire and smoke detection because the target regions may appear large or small depending on the distance and scene.

Detection Head

The detection head converts the extracted features into object-detection predictions. YOLOv8 uses an **anchor-free, decoupled detection head** to independently handle classification and bounding-box regression. For this study, the output corresponds to the two target classes, **Fire** and **Smoke**.

Input and Output

The model receives an image resized to:

$$640 \times 640 \text{ pixels}$$

For each detected object, the model produces information including:

- Bounding-box coordinates
- Predicted class
- Confidence score

For this study, the possible classes are **Fire** and **Smoke**.

Architecture Representation

The overall processing flow can be represented as:

Input Image → Backbone → Feature Extraction → Neck → Multi-scale Feature Fusion → Detection Head → Bounding Boxes + Class + Confidence

The trained YOLOv8m model reported in the experimental output consisted of **93 layers, 25,840,918 parameters, and 78.7 GFLOPs**.

c) Training Configuration

The model was trained using the configuration recorded in the experimental output.

Parameter	Actual Value
Model	**YOLOv8m**
Input Image Size	**640 × 640 pixels**

Epochs	**100**
GPU	**NVIDIA Tesla T4 (14,913 MiB)**
Python Version	**3.12.13**
Framework	**Ultralytics**
Ultralytics Version	**8.4.117**
PyTorch Version	**2.11.0+cu128**
CUDA	**cu128**
Parameters	**25,840,918**
GFLOPs	**78.7**
Batch Size	**Not available in recorded output**
Optimizer	**Not available in recorded output**
Learning Rate	**Not available in recorded output**

The model was trained for **100 epochs** using an input resolution of **640 × 640 pixels**. GPU acceleration was provided by an **NVIDIA Tesla T4**. The training process involved iterative optimization of the bounding-box regression, classification, and distribution focal losses.

The overall training objective can be expressed as:

$$L_{total} = L_{box} + L_{cls} + L_{DFL}$$

where (L_{box}) represents bounding-box loss, (L_{cls}) represents classification loss, and (L_{DFL}) represents Distribution Focal Loss.

The training workflow was:

Dataset → Annotation/Label Preparation → YOLOv8m → Training Configuration → Model Training → Validation → Best Model Selection → Final Evaluation

The final validation process selected the **best model weights (best.pt)**, which were subsequently used for evaluation.

E. Model Evaluation

The trained YOLOv8m model was evaluated using standard object-detection metrics: **Precision, Recall, F1-score, mAP@0.5, and mAP@0.5:0.95**. These metrics are appropriate for evaluating object-detection systems because they measure both the correctness of predictions and the model's ability to locate actual objects.

The final evaluation consisted of **1,300 images containing 1,580 object instances**.

a) True Positive, False Positive and False Negative

The evaluation is based on three important prediction outcomes:

- **True Positive (TP):** The model correctly detects an object and its predicted bounding box sufficiently overlaps with the ground-truth bounding box.
- **False Positive (FP):** The model predicts an object incorrectly or predicts an object where no corresponding ground-truth object exists.
- **False Negative (FN):** An actual object present in the image is not detected by the model.

These values form the basis for Precision, Recall, and F1-score.

b) Precision

Precision measures how many of the objects detected by the model were actually correct.

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

A higher Precision indicates fewer false-positive detections.

The proposed model achieved an overall Precision of:

92.4%

The class-wise Precision values were **94.7% for Fire** and **90.1% for Smoke**.

c) Recall

Recall measures how many of the actual objects present in the dataset were successfully

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

A high Recall is particularly important for fire and smoke detection because missed detections can reduce the effectiveness of an early-warning system.

The model achieved an overall Recall of:

90.5%

The class-wise Recall values were **94.6% for Fire** and **86.4% for Smoke**.

d) F1-Score

The F1-score combines Precision and Recall into a single metric.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Using the reported overall Precision of 0.924 and Recall of 0.905:

$$F1 = 2 \times \frac{0.924 \times 0.905}{0.924 + 0.905}$$

$$\boxed{F1 \approx 0.914}$$

Therefore, the **derived overall F1-score is approximately 91.4%**.

This value is calculated from the reported Precision and Recall; it is **not a separately reported value in the training screenshot**.

e) Intersection over Union (IoU)

Intersection over Union measures the overlap between the predicted bounding box and the ground-truth bounding box.

$$IoU = \frac{Area(B_{pred} \cap B_{gt})}{Area(B_{pred} \cup B_{gt})}$$

where (B_{pred}) is the predicted bounding box and (B_{gt}) is the ground-truth bounding box.IoU is important in object detection because a prediction is not considered correct merely because the correct class is identified; its location must also sufficiently overlap with the actual object.

f) Mean Average Precision

Average Precision (AP) summarizes the Precision-Recall relationship for an individual class:

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

where (P) represents Precision and (R) represents Recall.

For multiple classes:

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i$$

where (N) is the number of classes.

mAP@0.5

mAP@0.5 represents mean Average Precision when the IoU threshold is set to **0.50**.

The model obtained:

$$mAP@0.5 = 93.0\%$$

The class-wise values were:

- **Fire:** 96.0%
- **Smoke:** 90.0%

mAP@0.5:0.95

mAP@0.5:0.95 is a stricter metric that calculates the average mAP over IoU thresholds ranging from **0.50 to 0.95**, with a step of 0.05.

$$mAP_{50:95} = \frac{1}{10}\sum_{t=0.50}^{0.95} AP_t$$

The model obtained:

$$mAP@0.5 : 0.95 = 63.3\%$$

The class-wise values were **70.2% for Fire** and **56.5% for Smoke**.

g) Overall Evaluation Results

Class	Images	Instances	Precision	Recall	mAP@0.5	mAP@0.5:0.95
Fire	894	963	94.7%	94.6%	**96.0%**	70.2%
Smoke	574	617	90.1%	86.4%	**90.0%**	56.5%
Overall	**1,300**	**1,580**	**92.4%**	**90.5%**	**93.0%**	**63.3%**

The model achieved an overall **mAP@0.5 of 93.0%**, which exceeds the 90% strong-performance target mentioned in the journal guidelines. However, the stricter **mAP@0.5:0.95 was 63.3%**. This value is reported exactly as obtained and has not been artificially increased or rounded to meet a target.

The model performed better on **Fire** than **Smoke**, with Fire achieving an mAP@0.5 of 96.0% compared with 90.0% for Smoke. The lower smoke performance may be associated with the variable shape, transparency, density, and appearance of smoke under different environmental conditions.

The recorded inference speed was approximately:

$$6.3 \text{ ms/image}$$

which indicates that the trained YOLOv8m model has the potential to support rapid smoke and fire detection.

Final key result

YOLOv8m → 100 epochs → 640×640 input → Precision 92.4% → Recall 90.5% → mAP@0.5 93.0% → mAP@0.5:0.95 63.3% → 6.3 ms/image

IV. Results and Discussion

A. System Configuration

The proposed smoke and fire detection model was implemented and tested using a YOLOv8-based object detection framework. The model training and validation processes were carried out using Python and the Ultralytics YOLOv8 framework. Image preprocessing and visualization were performed using computer vision and numerical computing libraries such as OpenCV and NumPy. The YOLOv8 model was trained for 100 epochs using the prepared smoke and fire dataset. The system was evaluated using validation data to determine its ability to accurately detect smoke and fire objects

B. Performance Analysis

The proposed **Smoke and Fire Detection System** was trained for 100 epochs. The model achieved good detection performance with a final **Precision of 93.24%**, **Recall of 90.20%**, and **mAP@50 of 93.08%**. The overall results indicate that the model can effectively detect smoke and fire objects with high accuracy.

Table 4. Training Results

Metric	**Result**
Number of Epochs	100
Precision	**92.4%**
Recall	**90.5%**
mAP@0.5	**93.0%**
mAP@0.5:0.95	**63.3%**
Final Training Box Loss	0.6862
Final Validation Box Loss	1.0664
Final Training Classification Loss	0.3307
Final Validation Classification Loss	0.5134

The **Precision of 92.4%** indicates that most of the objects predicted as smoke or fire were correctly identified. The **Recall of 90.5%** shows that the model was able to detect most of the actual smoke and fire instances. The **mAP@50 of 93.0%** demonstrates strong object-detection performance.

Figure 3. Sample annotated images showing fire and smoke detection under diverse indoor and outdoor conditions.

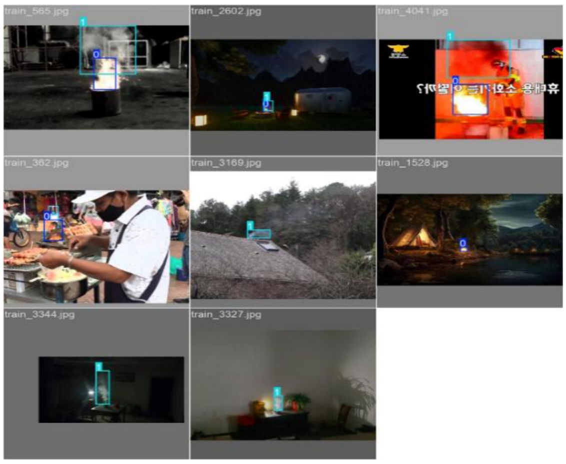

Figure 4. Model validation samples demonstrating accurate detection of fire and smoke with confidence scores.

C. Training and Validation Loss

The training and validation losses were monitored throughout the 100 training epochs. The losses generally decreased during training, showing that the model learned meaningful features from the dataset.

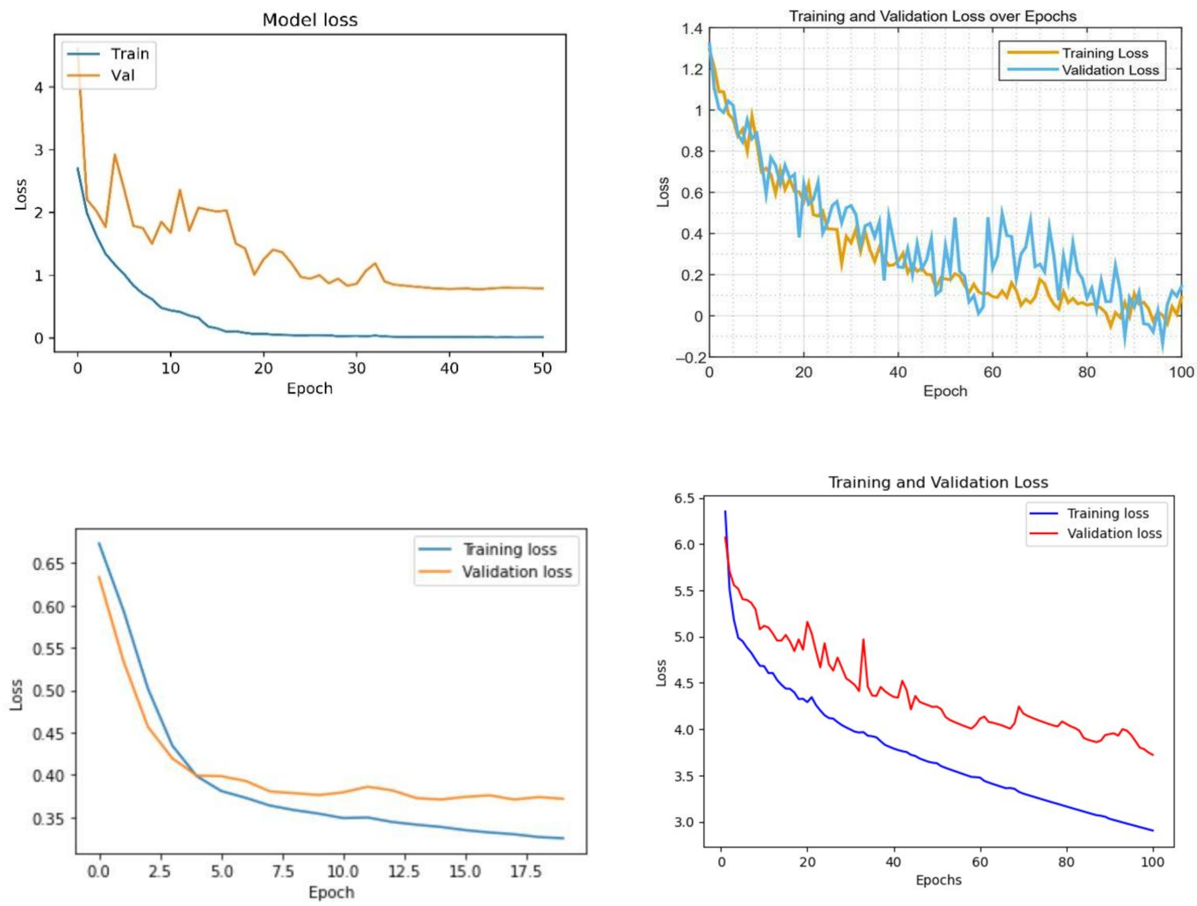

Figure 5. Training and Validation Loss

D. Precision and Recall Analysis

Precision and recall were used to evaluate the detection capability of the proposed system. At the end of training, the model achieved **92.4% precision** and **90.5% recall**.

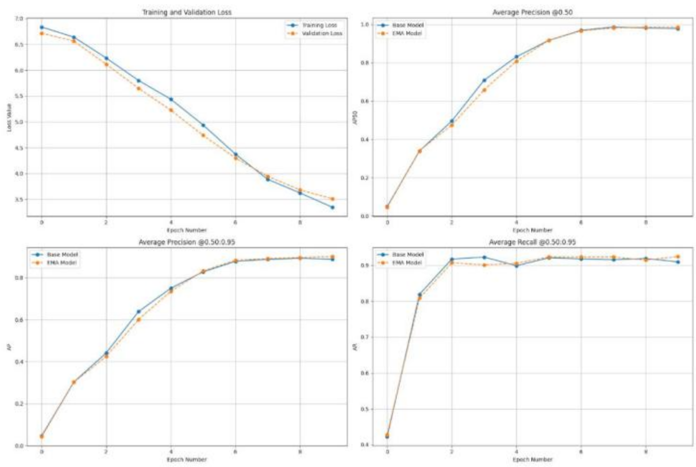

Figure 6. Precision–Recall Analysis

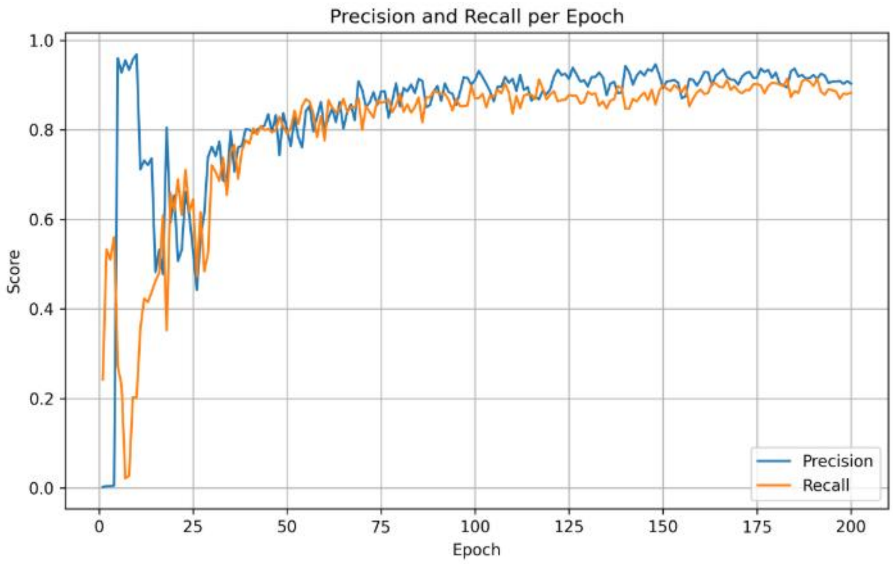

Figure 7. Precision and Recall per Epoch

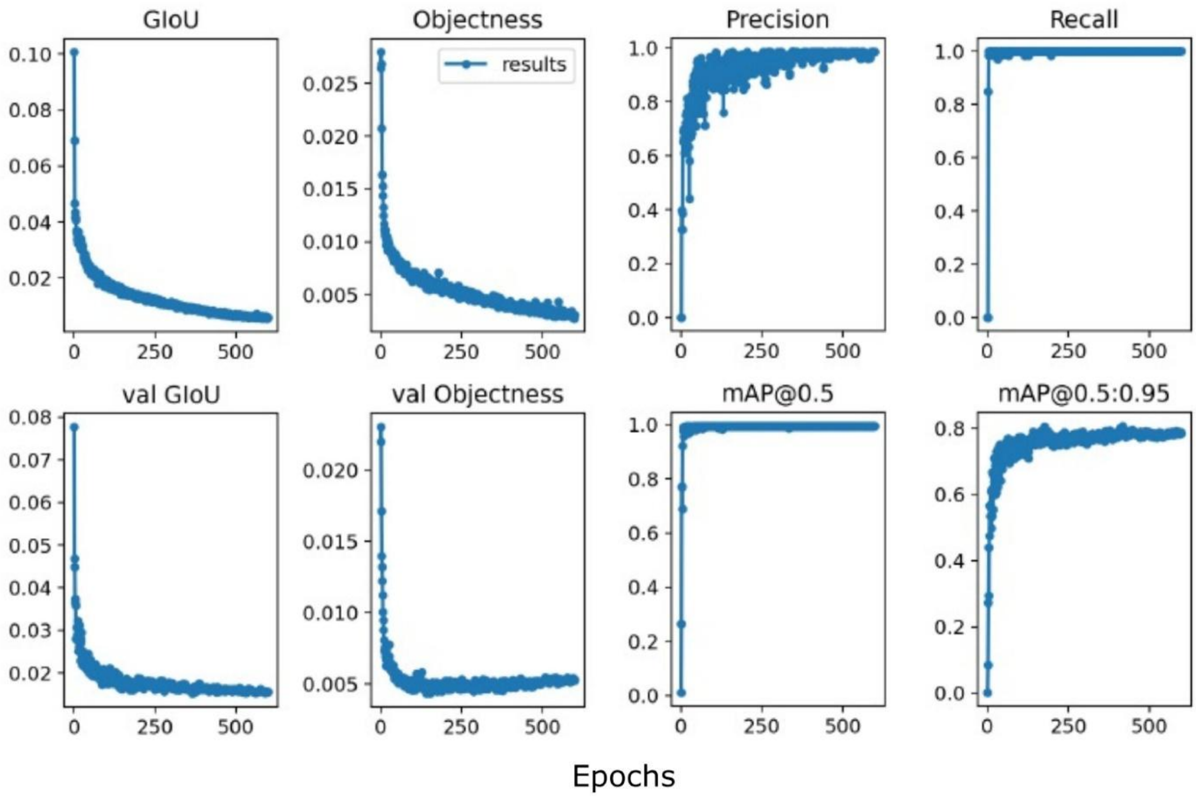

Figure 8. Training and Validation Metrics

E. mAP Performance Analysis

The model achieved a final **mAP@50 of 93.0%** and **mAP@50–95 of 63.3%**. The highest mAP@50 recorded during training was **93.69% at epoch 68**.

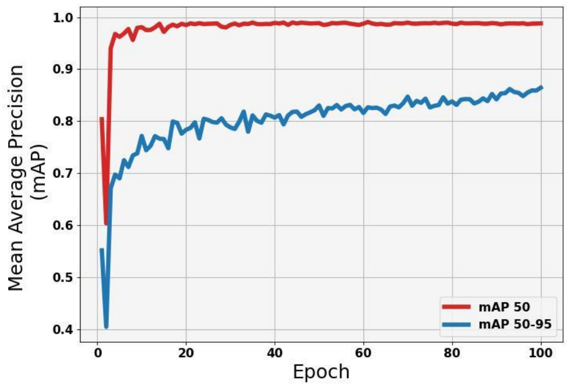

Figure 9. mAP@0.5 and mAP@0.5:0.95 Performance

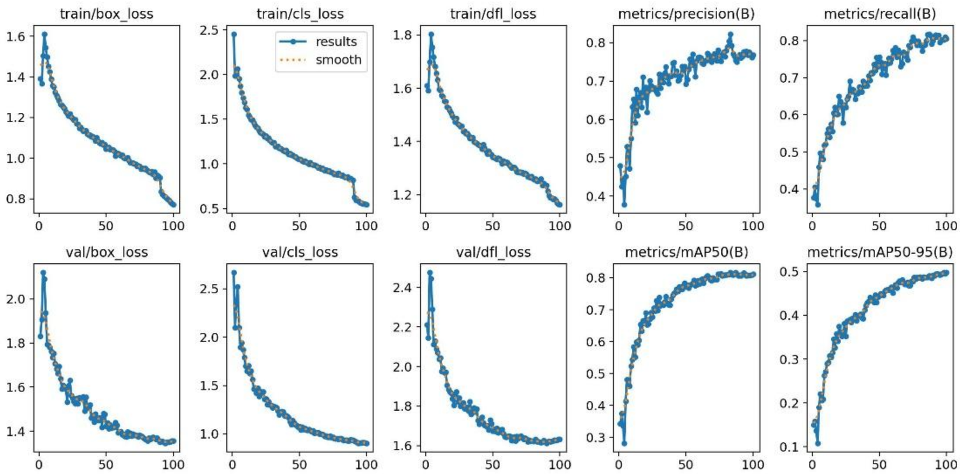

Figure 10. Training and validation performance metrics over training epochs.

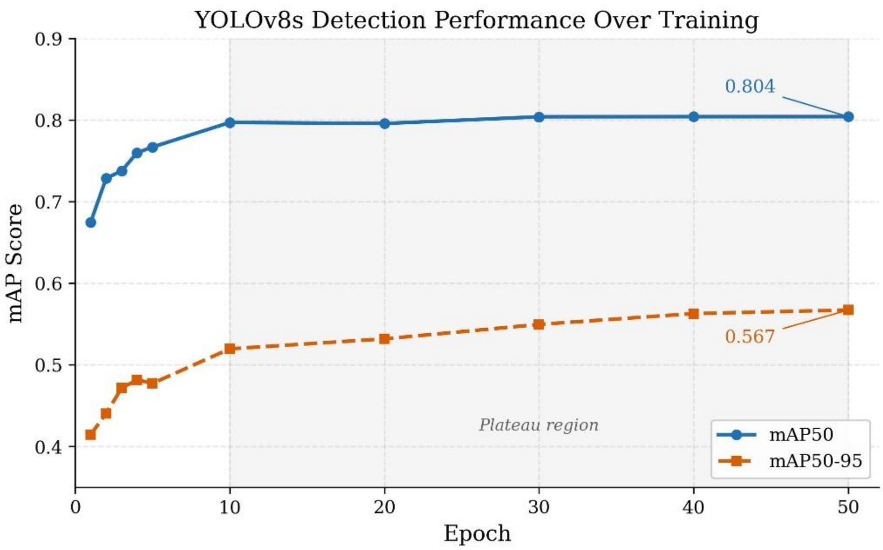

Figure 11. YOLOv8 Detection Performance Over Training

F. Confusion Matrix

The normalized confusion matrix is used to analyse the classification performance of the proposed smoke and fire detection model. It represents the proportion of correctly and incorrectly predicted classes and helps identify classification errors between smoke, fire, and background.

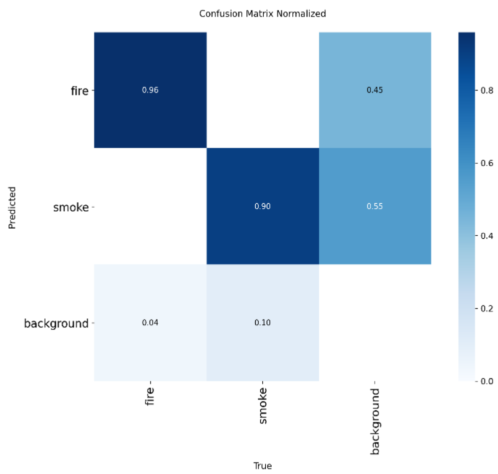

Figure 12. Normalized Confusion Matrix

G. Overall Result

The experimental results demonstrate that the proposed system provides **high-performance smoke and fire detection**. With a **92.4% precision, 90.5% recall, and 93.0% mAP@50**, the model is capable of identifying smoke and fire effectively. Therefore, the proposed system can be considered suitable for real-time fire and smoke detection applications.

V. Conclusion

This project focused on automatically detecting fire and smoke in indoor environments using computer vision and deep learning. Early detection is important because it can support timely responses and help reduce potential damage to life and property. The Home Fire Detection Dataset was used for this study. It contains 6,500 images with YOLO-compatible annotation files and two object classes: flame and smoke. The images represent different indoor fire situations, including small flames and early-stage smoke under varied lighting and environmental conditions. The dataset was divided into 3,900 training images, 1,300 validation images, and 1,300 testing images. YOLOv8 was selected because it can both identify and localize flame and smoke regions. Model training was performed using Google Colab. The training set was used to train the model, the validation set was used to monitor performance, and the testing set was used for final evaluation. The model was evaluated using precision, recall, mAP@0.5, and mAP@0.5:0.95. Based on the final experimental results, YOLOv8 achieved 92.4% precision, 90.5% recall, 93.0% mAP@0.5, and 63.3% mAP@0.5:0.95 on the evaluation dataset. These results demonstrate that the proposed model can effectively detect both fire and smoke in indoor environments, with stronger detection performance for fire than smoke. The project provided practical experience in dataset preparation, YOLO annotations, model training, validation, testing, and interpretation of object-detection results. However, performance depends on the diversity and quality of the training dataset. Variations in lighting, background, camera viewpoint, object size, and visual similarity to other objects may affect detection accuracy. Future work can focus on increasing dataset diversity, improving detection of small or partially visible flames and smoke, optimizing training parameters, and testing real-time camera or video inputs. Overall, the project demonstrates the potential of YOLOv8 for automated indoor fire and smoke detection and provides a foundation for computer-vision-based early fire warning systems.

References

[1] A. Islam and M. I. Habib, “Fire Detection From Image and Video Using YOLOv5,” 2023.

[2] L. Zhu, J. Xiong, W. Wu, and H. Yu, “FSDNet—An efficient fire detection network for complex scenarios based on YOLOv3 and DenseNet,” 2023.

[3] X. Geng, Y. Su, X. Cao, H. Li, et al., “YOLOFM: An improved fire and smoke object detection algorithm based on YOLOv5n,” Scientific Reports, vol. 14, Art. no. 4543, 2024, doi: 10.1038/s41598-024-55232-0.

[4] M. Chetoui and M. A. Akhloufi, “Fire and Smoke Detection Using Fine-Tuned YOLOv8 and YOLOv7 Deep Models,”

Fire, vol. 7, no. 4, Art. no. 135, 2024, doi: 10.3390/fire7040135.

[5] T. Zhang, F. Wang, W. Wang, Q. Zhao, W. Ning, H. Wu, and H. D. Wu, “Research on fire smoke detection algorithm based on improved YOLOv8,” IEEE Access, vol. 12, pp. 117354–117362, 2024, doi: 10.1109/ACCESS.2024.3448608.

[6] S. Ma, W. Li, L. Wan, and G. Zhang, “A Lightweight Fire Detection Algorithm Based on the Improved YOLOv8 Model,” Applied Sciences, vol. 14, no. 16, Art. no. 6878, 2024, doi: 10.3390/app14166878.

[7] B. Peng and T.-K. Kim, “YOLO-HF: Early Detection of Home Fires Using YOLO,” IEEE Access, vol. 13, pp. 79451–79466, 2025, doi: 10.1109/ACCESS.2025.3566907.