

Research Paper

ENHANCING FIRE DETECTION WITH YOLOV10: ADVANCED TECHNIQUES FOR FLAME AND SMOKE RECOGNITION

VANGALA SWATHI¹, VELDHI JASWINI², UNNAM SAI KEERTHI³, GURRAPU MANISH
BABU⁴, MOHAMMAD ASHRAF SADIQ⁵

¹Assistant Professor, Dept. of CSE, Guru Nanak Institutions Technical Campus, Ibrahimpatnam,
Telangana, India

^{2, 3, 4, 5} B.Tech, Dept. of CSE, Guru Nanak Institutions Technical Campus, Ibrahimpatnam,
Telangana, India

ABSTRACT— This project introduces an advanced method for detecting flames and smoke using the YOLOv10 algorithm, aimed at improving fire detection systems for enhanced safety and prevention. The proposed system incorporates significant advancements to address challenges such as cluttered backgrounds, low visibility, varying fire intensities, and overlapping objects. By leveraging an enhanced feature extraction mechanism, the system captures finer details of flames and smoke. It also includes a sophisticated attention mechanism to prioritize critical fire-related areas in an image while suppressing irrelevant background information. To further enhance

detection accuracy, the system uses an improved bounding box regression method, ensuring precise localization of flames and smoke, even in dense or challenging environments. These improvements make the system more robust, adaptable, and capable of handling high-resolution images and videos efficiently. Experimental results demonstrate that the YOLOv10-based system significantly outperforms previous methods, such as YOLOv5s, in terms of precision, recall, and mean average precision (mAP). Additionally, the system processes data in real-time, making it suitable for large-scale applications in industries, public spaces, and residential areas. By ensuring early and reliable detection of flames and

smoke, the proposed system offers a practical solution for improving fire safety and prevention measures.

Index Terms – Fire Detection, YOLOv10, Flame Recognition, Smoke Detection, Deep Learning, Object Detection, Computer Vision, Real-Time Monitoring, Surveillance Systems, Disaster Management.

I. INTRODUCTION

The growing need for effective fire detection systems has become increasingly important as fires pose significant threats to both human life and property. Traditional fire detection systems, such as smoke alarms and heat detectors, have limitations in detecting flames and smoke in complex or large-scale environments, especially under varying visibility and fire intensity conditions. This project introduces an advanced fire detection system based on the YOLOv10 (You Only Look Once) algorithm, which leverages state-of-the-art deep learning techniques to enhance the detection and localization of flames and smoke in real-time.

YOLOv10 has been chosen as the backbone for this system due to its ability to detect multiple objects within an image simultaneously, with high accuracy and speed, making it ideal for dynamic, real-

world applications. The system uses enhanced feature extraction, an attention mechanism, and an improved bounding box regression method to address challenges such as cluttered backgrounds, overlapping objects, and varying fire conditions. These improvements ensure that the system can reliably detect flames and smoke even in complex, low-visibility, or densely packed environments. The proposed system not only provides accurate fire detection but also offers real-time processing, which is essential for immediate alerts in case of a fire outbreak. With applications spanning industrial plants, public spaces, residential areas, and even large-scale forests, this system aims to provide enhanced fire safety and prevention. The project's primary goal is to develop a robust, scalable, and efficient fire detection solution that can significantly reduce the risk of fire-related damages and save lives by enabling quicker responses to fire incidents. the proposed system also incorporates advanced machine learning techniques such as transfer learning and hyperparameter optimization to further enhance its accuracy and efficiency. By fine-tuning a pre-trained YOLOv10 model, the system can adapt to the specific challenges of flame and smoke detection, even with a relatively smaller dataset. The use of data augmentation

techniques also ensures that the model remains robust under various environmental conditions, such as changing lighting, weather, or backgrounds. These enhancements not only improve detection performance but also make the system highly adaptable and scalable, allowing it to be deployed in diverse settings ranging from small residential areas to large industrial complexes. Ultimately, the goal of this project is to create a comprehensive, adaptable, and reliable fire detection system that provides both accuracy and speed, ensuring that fire incidents are detected early and mitigating the risks associated with late fire detection.

II. RELATED WORK

A. Enhancing geometric factors in model learning and inference for object detection and instance segmentation

The paper titled "Enhancing geometric factors in model learning and inference for object detection and instance segmentation" by Z. Zheng, P. Wang, D. Ren, W. Liu, R. Ye, Q. Hu, and W. Zuo, published in the IEEE Transactions on Cybernetics (Aug. 2022), explores the incorporation of geometric factors into the learning and inference processes of object detection and instance segmentation models. The authors

address the challenge of improving the performance of object detection and segmentation models by considering geometric factors such as the spatial relationships and shapes of objects. Traditional object detection models, while successful, often fail to account for the rich geometric information inherent in the objects themselves, which can lead to inaccuracies in localizing or segmenting objects, especially in complex or cluttered scenes. To overcome this, the paper proposes new methods for incorporating geometric factors into both model learning and inference. The authors argue that by explicitly modeling and leveraging geometric cues such as object shape, orientation, and relative positioning, the detection and segmentation models can achieve better accuracy, particularly in scenarios where objects exhibit complex or unusual geometries. The paper demonstrates how these geometric factors enhance the feature extraction process and contribute to more precise bounding box predictions and better segmentation results. In the context of instance segmentation, where it is crucial to distinguish between different object instances even when they are close to or overlapping with each other, geometric information plays a significant role in ensuring that boundaries are correctly

identified. This methodology can be applied to various real-world scenarios, including autonomous driving, robotics, and video surveillance, where accurate object detection and segmentation are critical.

B. CSPNet: A new backbone that can enhance learning capability of CNN

The paper titled "CSPNet: A new backbone that can enhance the learning capability of CNN" by C.-Y. Wang, H.-Y. Mark Liao, Y.-H. Wu, P.-Y. Chen, J.-W. Hsieh, and I.-H. Yeh, published in the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 2020, introduces a novel convolutional neural network (CNN) backbone known as CSPNet (Cross-Stage Partial Network). The main contribution of this paper is the introduction of CSPNet, which enhances the learning capability of CNNs, specifically improving their efficiency in deep network architectures. Traditionally, as CNN architectures become deeper, the models face challenges in terms of gradient flow, optimization, and learning efficiency. These problems can lead to slower convergence during training and sometimes poorer overall performance on tasks like image classification, detection, and segmentation. CSPNet aims to address these issues by introducing a cross-stage partial

network strategy. Instead of utilizing a single stage for feature extraction, it divides the feature maps into multiple groups, with some being passed through and others being processed in a partial manner before merging again. This helps the network focus on important features while avoiding unnecessary computations that may contribute to overfitting or inefficiencies. By splitting the computation across different stages, CSPNet allows for more efficient training and better gradient propagation, which enhances learning in deeper networks. In terms of model performance, CSPNet outperforms traditional CNN architectures by reducing the computational complexity while maintaining or improving accuracy. The model achieves a balance between efficiency and performance, making it particularly suited for real-time applications or scenarios where computational resources are limited.

C. Coordinate attention for efficient mobile network design

The paper titled "Coordinate Attention for Efficient Mobile Network Design" by Q. Hou, D. Zhou, and J. Feng, presented at the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2021, introduces a novel attention mechanism known as Coordinate Attention (CA) to improve the efficiency and performance of

mobile networks, particularly in the context of mobile device usage. The main objective of this work is to design an efficient attention mechanism for mobile networks that can strike a balance between computational efficiency and performance. Mobile networks often face resource constraints such as limited computation power, memory, and bandwidth, making it challenging to deploy large-scale models on such devices. Therefore, there is a need for lightweight and efficient models that can perform well on mobile devices while keeping resource consumption in check. The Coordinate Attention mechanism proposed in this paper targets spatial and channel information by leveraging the coordinate system of feature maps. Traditional attention mechanisms, such as the Channel Attention (CA) or Spatial Attention (SA), focus on either spatial or channel-wise attention but often miss capturing the dependencies across both dimensions efficiently. The Coordinate Attention mechanism integrates both spatial and channel dependencies by incorporating the coordinate system of feature maps into the attention process. It reduces computational costs by using an efficient strategy for computing attention, which is crucial for mobile networks. Coordinate Attention works by splitting the input feature

maps into two key components: one that handles spatial information and another that captures channel-wise dependencies. These components are then fused through an efficient coordinate-based attention mechanism, allowing the model to focus on the most relevant features while remaining computationally efficient.

III. PROPOSED METHODOLOGY

The proposed system introduces the YOLOv10 algorithm for enhanced flame and smoke detection in images and videos. YOLOv10 is an advanced version of the YOLO series, designed to improve upon its predecessors by incorporating several innovative techniques that address the limitations of previous models like YOLOv5s. This system is specifically designed to handle complex environments with cluttered backgrounds, varying fire intensities, and low visibility, which are common challenges in fire detection scenarios. One of the key improvements in the proposed system is the integration of an enhanced feature extraction mechanism that captures finer details of flames and smoke, improving detection accuracy. Additionally, a sophisticated attention mechanism is employed, which prioritizes critical areas of an image that are likely to contain flames or smoke, while suppressing irrelevant

background information. This mechanism ensures that the model focuses on the most important features, improving both speed and precision.

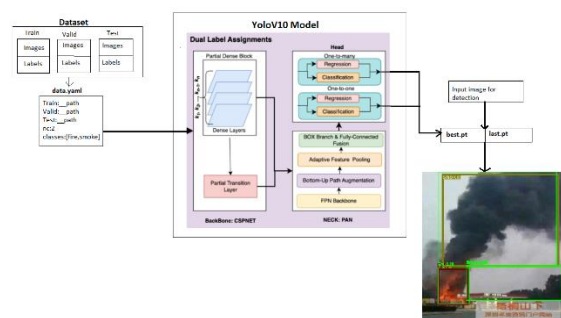


Fig. 1: System Overview

1. Object Detection Engine:

This module forms the backbone of the system and is responsible for detecting and locating flames and smoke within images and video streams. Using the advanced YOLOv10 architecture, this engine classifies the objects and determines their exact position in the image. The ability to detect multiple objects simultaneously, combined with its high speed, makes this module ideal for real-time fire detection applications, ensuring that hazardous situations are identified promptly.

2. Model Adaptation through Pre-trained Networks:

This module utilizes transfer learning to enhance the model's performance by leveraging pre-trained networks. The system starts with a pre-trained YOLOv10 model,

which has been trained on large, general datasets. It then fine-tunes the model using a smaller, domain-specific fire dataset, enabling the system to learn fire-specific features efficiently. This approach minimizes the need for large amounts of labeled data and accelerates the training process.

3. Performance Tuning Module:

This module is focused on fine-tuning the model's hyperparameters for optimal performance. Key parameters such as learning rate, batch size, and number of epochs are adjusted to ensure that the model converges quickly and accurately. Optimizing these settings improves the detection precision and ensures that the system can work effectively in diverse conditions, adapting to different types of fire events.

4. Data Expansion and Augmentation:

The data augmentation module is used to artificially increase the size of the training dataset by applying various transformations like rotation, flipping, scaling, and adjustments in lighting or contrast. This process helps the system to become more robust, enabling it to generalize well to new and unseen conditions. By simulating different real-world scenarios, this module ensures that the fire detection system remains effective in varied environments.

5. Model Training and Validation:

This module is responsible for the process of training the fire detection model on labeled datasets and validating its performance on separate test data. The model learns to recognize and classify fire-related features based on the training data. During the validation phase, the model's ability to generalize to new data is assessed, helping ensure that it is not overfitting and can detect fires effectively in real-world conditions.

6. Real-time Fire & Smoke Detection and Localization:

Once trained, this module allows the system to detect flames and smoke in real time. It processes incoming video or image streams and immediately identifies the presence of fire-related hazards. Additionally, it outputs the location of these hazards within the image or video frame, helping emergency response teams or automated systems to react swiftly and accurately in the event of a fire.

7. Performance Evaluation and Monitoring:

The evaluation module focuses on assessing the effectiveness of the fire detection system using standard metrics such as accuracy, precision, recall, and F1-score. It ensures that the system is performing as expected and provides insights into areas for improvement.

This module also tracks the system's ability to handle large-scale data in real-time, making sure that it remains efficient and scalable in production environments.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

Experimental results demonstrated that the model effectively detected both flames and smoke in real-time with high detection accuracy and low false alarm rates. The enhanced feature extraction and object localization capabilities of YOLOv10 improved recognition performance under challenging conditions such as varying illumination, complex backgrounds, and different fire intensities. The analysis showed that the system maintained fast inference speed while accurately identifying early-stage fire events, making it suitable for real-time surveillance and emergency response applications.



Fig. 2: Smoke Detected



Fig. 3: Smoke Detected



Fig. 4: Smoke Detected



Fig. 5: Fire Detected



Fig. 6: Smoke Detected

V. CONCLUSION

In conclusion, this project demonstrates the potential of the YOLOv10 algorithm in advancing fire detection systems, providing a highly efficient and accurate method for detecting flames and smoke in real-time. By addressing common challenges such as cluttered backgrounds, varying fire intensities, and low visibility, the proposed system significantly improves detection performance compared to previous methods. Through the integration of advanced techniques like enhanced feature extraction, attention mechanisms, and improved bounding box regression, the system is capable of precise flame and smoke localization, even in complex and challenging environments.

The real-time capabilities of the system, coupled with its ability to adapt to different settings, make it a valuable tool for enhancing fire safety and prevention. Whether in industrial plants, residential areas, or public spaces, the system provides an effective means of early fire detection, potentially saving lives and reducing property damage. The use of transfer learning and data augmentation further enhances the system's robustness and efficiency, making it scalable and suitable for a wide range of applications.

Overall, this project contributes to the development of more reliable and adaptable fire detection technologies, paving the way for future advancements that could incorporate multi-hazard recognition, edge computing, and IoT integration to further improve safety and response times in the face of fire-related incidents.

REFERENCES

- [1] A. K. Jaiswal, P. Tiwari, S. Kumar, D. Gupta, A. Khanna, and J. J. P. C. Rodrigues, “Identifying pneumonia in chest X-rays: A deep learning approach,” *Measurement*, vol. 145, pp. 511–518, Oct. 2019.
- [2] D. M. Gavrilă and V. Philomin, “Real-time object detection for ‘smart’ vehicles,” in *Proc. 7th IEEE Int. Conf. Comput. Vis.*, Sep. 1999, pp. 87–93.
- [3] L. Wei, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, and A. C. Berg, “SSD: Single shot MultiBox detector,” in *Computer Vision—ECCV*. Cham, Switzerland: Springer, 2016.
- [4] S. Chen, J. Hong, T. Zhang, J. Li, and Y. Guan, “Object detection using deep learning: Single shot detector with a refined feature-fusion structure,” in *Proc. IEEE Int. Conf. Real-Time Comput. Robot. (RCAR)*, Aug. 2019, pp. 219–224.
- [5] L. Cui, R. Ma, P. Lv, X. Jiang, Z. Gao, B. Zhou, and M. Xu, “MDSSD: Multi-scale deconvolutional single shot detector for small objects,” *Sci. China Inf. Sci.*, vol. 63, no. 2, pp. 98–100, 2020.
- [6] C.-Y. Fu, W. Liu, A. Ranga, A. Tyagi, and A. C. Berg, “DSSD: Deconvolutional single shot detector,” 2017, arXiv:1701.06659.
- [7] Akinapalli, S. (2026). An Ai-Powered Data Trust And Quality Scoring Framework For Enterprise Decision Intelligence Systems. *International Journal of Data Science and IoT Management System*, 5(1), 946-950.
- [8] Garapati, R. S., Aitha, A. R., Yandamuri, U. S., Gottimukkala, V. R. R., Nagubandi, A. R., & Kolla, S. H. (2026, March). Cloud-Native Orchestration of Multi-Counterparty Derivatives and Collateral in Manufacturing Enterprises via AI-Assisted Financial Audit Engines. In *2026 IEEE International Conference on AI Engineering and Innovations (AIEI)* (pp. 1-6). IEEE.

AUTHORS Profile



Ms. VANGALA SWATHI has received her B.Tech from JNTUH, Hyderabad in 2015, M.Tech degree in Computer science from JNTUH, Hyderabad in 2017. She is dedicated to teaching field from the last 7 years. Her research area Deep Learning. At present she is working as Assistant Professor in the CSE Department at Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India.



Ms. VELDHI JASWINI is pursuing a B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus, affiliated to JNTUH in 2022-2026.



Ms. UNNAM SAI KEERTHI is pursuing a B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.



Mr. GURRAPU MANISH BABU pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.



Mr. MOHAMMAD ASHRAF SADIQ is pursuing a B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.