International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

REAL-TIME FIRE DETECTION WITH YOLOV10: IMPROVED TECHNIQUES FOR SMOKE AND FLAME IDENTIFICATION

Sadiya Sultana[1], Dr. Lalitha Saroja[2]
[1]PG Scholar, Department of CSE, Shadan Women's College of Engineering and Technology, Hyderabad
Sadiya1sultana8@gmail.com
[2]Assoc Professor, Department of CSE, Shadan Women's College of Engineering and Technology, Hyderabad
lalithasarojathota@gmail.com

ABSTRACT:
In order to improve fire detection systems for increased safety and prevention, this study presents a sophisticated technique for detecting flames and smoke utilising the YOLOv10 algorithm. Significant improvements are incorporated into the suggested method to handle issues including overlapping objects, low visibility, cluttered backdrops, and variable fire intensities. The technology catches greater details of smoke and flames by utilising an improved feature extraction mechanism. Additionally, it has an advanced attention system that suppresses unimportant background information while highlighting important fire-related parts in an image. The system employs an enhanced bounding box regression technique to substantially improve detection accuracy, guaranteeing accurate localisation of flames and smoke even in difficult or thick settings. These enhancements increase the system's resilience, flexibility, and ability to effectively handle high-resolution photos and movies. According to experimental results, the YOLOv10-based system works far better in terms of precision, recall, and mean average precision (mAP) than earlier techniques like YOLOv5s. The system can be used for large-scale applications in public areas, businesses, and homes because it processes data in real-time. The suggested approach provides a workable way to enhance fire safety and prevention measures by guaranteeing early and accurate detection of flames and smoke.

1. INTRODUCTION:

Because fires represent serious risks to property and human life, there is a rising need for efficient fire detection systems. In complicated or large-scale situations, traditional fire detection systems—such as smoke alarms and heat detectors—have limitations when it comes to identifying flames and smoke, particularly when visibility and fire intensity circumstances alter. This research presents a sophisticated fire detection system based on the YOLOv10 (You Only Look Once) algorithm, which uses cutting-edge deep learning techniques to improve real-time flame and smoke detection and localisation.

The foundation of our system is YOLOv10, which is perfect for dynamic, real-world applications since it can quickly and accurately recognise many items in an image at once. To overcome obstacles including crowded backdrops, overlapping objects, and changing firing circumstances, the system employs improved bounding box regression, an attention mechanism, and increased feature extraction. These enhancements guarantee that the system can consistently identify smoke and flames even in intricate, poorly visible, or crowded settings. The suggested system includes real-time processing in addition to precise fire detection, which is crucial for prompt notifications in the event of a fire outbreak. This system is intended to improve fire safety and prevention in a variety of settings, including public areas, residential areas, industrial sites, and even large-scale woods. The main objective of the project is to create a reliable, scalable, and effective fire detection system that can greatly lower the risk of fire-related damages and save lives by facilitating faster reactions to fire situations. To further improve its accuracy and efficiency, the suggested system also integrates cutting-edge machine learning methods including transfer learning and hyperparameter optimisation. Even with a somewhat smaller dataset, the system can adjust to the unique difficulties of flame and smoke detection by fine-tuning a pre-trained YOLOv10 model. The model's resilience in a variety of environmental circumstances, such as shifting illumination, weather, or backdrops, is further guaranteed by the application of data augmentation techniques. In addition to improving detection performance, these improvements make the system extremely flexible and scalable, enabling its deployment in a variety of environments, from tiny residential neighbourhoods to expansive industrial complexes. The project's ultimate objective is to develop a thorough, flexible, and dependable fire detection system that offers speed and accuracy, guaranteeing that fire occurrences are discovered early and reducing the dangers connected with delayed fire detection.

1.1 OBJECTIVE:

The goal of this project is to create a sophisticated fire detection system that uses the YOLOv10 algorithm to precisely identify and locate smoke and flames in real time. In order to overcome obstacles like crowded backdrops, fluctuating fire intensities, and overlapping objects, the system incorporates sophisticated approaches including improved feature extraction, attention mechanisms, and improved bounding box regression methods. The research also aims to make the system resilient enough to operate

Received: 06-10-2026 | Accepted: 19-11-2026 | Published: 25-11-2026

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

well in challenging situations with poor visibility or crowded surroundings. Transfer learning and data augmentation will be used to optimise the system for real-time applications, guaranteeing quick and accurate detection. The objective is to develop a flexible and scalable solution that may be used in a variety of contexts, including public spaces, residential neighbourhoods, industrial facilities, and wildfire management. The project's ultimate goal is to increase fire safety by offering early and precise fire detection, facilitating prompt responses to stop fire-related damage, and saving lives.

2. EXISTING SYSTEM:

The YOLOv5s algorithm is used by the current system in this project to identify smoke and flames in pictures and videos. A popular and effective object recognition model, YOLOv5s is renowned for its quickness and precision in real-time applications. When it comes to fire detection, YOLOv5s can identify a variety of things in the scene, such as smoke and flames, and localise them by drawing bounding boxes around them.

The current system makes use of YOLOv5s' capacity to identify and categorise objects in a single forward pass, which makes it appropriate for applications that call for real-time fire detection, like public, residential, and industrial surveillance systems. It performs object detection using a typical feature extraction procedure and the model's pre-trained weights. Although YOLOv5s can handle complicated situations, it may have trouble with some issues, like crowded backgrounds, fluctuating fire intensities, and low visibility, which can make it less accurate at identifying smoke and flames.

Performance-wise, YOLOv5s is comparatively quick and effective; yet, the system may encounter difficulties in complex or dense situations where objects overlap or when the smoke or fire is not easily seen. Notwithstanding these difficulties, YOLOv5s offers a strong basis for fire detection applications by striking a fair balance between accuracy and speed. However, the proposal of YOLOv10 as an improvement to overcome these limits and improve overall detection accuracy and reliability resulted from the necessity for more sophisticated algorithms to improve detection performance in difficult real-world settings.

2.1 EXISTING SYSTEM DISADVANTAGES:

- Lower Detection Accuracy: YOLOv5s may miss smaller or partially obscured flames and smoke, leading to false negatives, whereas YOLOv10 provides higher precision in complex environments.
- Limited Feature Extraction: YOLOv5s struggles with capturing fine details of flames and smoke, while YOLOv10 enhances feature extraction for better detection in cluttered scenarios.
- Inability to Prioritize Critical Features: YOLOv5s processes the entire image equally, while YOLOv10 uses an attention mechanism to focus on fire-related areas, improving accuracy.
- Bounding Box Regression Limitations: YOLOv5s may provide less accurate bounding box localization, whereas YOLOv10's improved method ensures better detection, even in dense environments.
- Performance in Dense Environments: YOLOv5s may fail in detecting flames and smoke in crowded or obstructed scenes, while YOLOv10 excels in these situations.
- Real-Time Detection Limitations: YOLOv5s may struggle with high-resolution images in real-time, whereas YOLOv10 is optimized for faster processing and real-time detection.

3. LITERATURE SURVEY:

Coordinate attention for efficient mobile network design, Q. Hou, D. Zhou, and J. Feng

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.

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

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.

CSPNet: A new backbone that can enhance learning capability of CNN, C.-Y. Wang, H.-Y. Mark Liao, Y.-H. Wu, P.-Y. Chen, J.-W. Hsieh, and I.-H. Yeh

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.

4. RELATED WORK:

Recent developments in artificial intelligence and computer vision have greatly enhanced image-based forest fire detection systems. Image-based techniques offer more coverage, quicker response times, and more precise environmental data than conventional sensor-based techniques. Due to its superior accuracy over traditional image processing methods, Convolutional Neural Networks (CNNs) have been frequently used for smoke and fire detection. The efficacy of deep learning models in detecting fire events in intricate forest ecosystems has been shown in a number of studies.

Many strategies have been put forth by researchers to enhance the effectiveness of fire detection. To improve fire recognition accuracy, Pan et al. created AddNet, which combines image segmentation and classification approaches. A cascading classification method that can identify flames in both local and global picture regions was presented by Li et al. In a similar vein, drone-based surveillance systems that employ ensemble learning techniques and YOLO-based models have demonstrated encouraging outcomes in raising detection accuracy while lowering false alarm rates.

Even with these developments, it is still very difficult to identify low-concentration smoke and small-scale fires. Conventional CNN-based models frequently have tiny receptive fields, which limits their capacity to capture minor object features and long-range relationships. Researchers have investigated methods like multi-scale feature learning, contextual information integration, scale normalisation, and data augmentation to overcome these constraints. These methods aid in enhancing the identification of distant and minor fire areas in a variety of environmental circumstances.

Advanced architectures like Transformer-based object identification models, PANet, ASFF, and Feature Pyramid Networks (FPN) have also been studied recently. These techniques improve the model's capacity to capture global contextual information and improve feature extraction across many scales. In particular, transformer-based methods use self-attention processes to get over the drawbacks of traditional CNNs, making it possible to detect small fire and smoke targets in intricate forest environments with greater accuracy.

5. PROPOSED SYSTEM:

The YOLOv10 algorithm for improved flame and smoke detection in photos and videos is introduced in the suggested system. YOLOv10 is a more sophisticated version of the YOLO series that was created to outperform its predecessors by utilising a number of cutting-edge methods that solve the shortcomings of earlier versions, such as YOLOv5s. This method is especially made to manage complicated settings with crowded backdrops, fluctuating fire intensities, and poor visibility—all of which are typical problems in fire detection situations. The incorporation of an improved feature extraction process that captures finer details of flames and smoke, increasing detection accuracy, is one of the main enhancements in the suggested system. A complex attention process is also used, which suppresses unimportant background information while giving priority to important regions of a picture that are probably going

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

to contain smoke or flames. By ensuring that the model concentrates on the most crucial elements, this approach enhances speed and accuracy.

The application of an enhanced bounding box regression technique, which guarantees more accurate localisation of flames and smoke even in congested or difficult conditions, is another noteworthy development. Even when several items overlap or when smoke and flames are partially veiled, the system is still able to precisely locate and identify fire-related objects in photos. Real-time detection, which is crucial for fire safety applications, is another feature of the suggested YOLOv10-based system. The system may instantly notify users in a variety of locations, including public areas, industrial facilities, and residential areas, by swiftly and effectively analysing high-resolution photos and videos. The system is resilient and flexible, able to function under a variety of circumstances, such as fluctuating light levels, environmental elements, and fire intensities.

The suggested approach significantly improves real-time fire detection by outperforming the current YOLOv5s algorithm in terms of precision, recall, and mean average precision (mAP). It offers a more dependable and flexible way to improve fire safety and prevention practices in various settings.

5.1 PROPOSED SYSTEM ADVANTAGES:

- Higher Detection Accuracy: YOLOv10 offers improved accuracy, effectively detecting flames and smoke, even in challenging environments with low visibility or partial obstructions.
- Enhanced Feature Extraction: YOLOv10 captures finer details of flames and smoke, improving its ability to detect fire-related objects in cluttered and dynamic settings.
- Advanced Attention Mechanism: YOLOv10 prioritizes fire-related regions, ensuring faster and more accurate detection by suppressing irrelevant background information.
- Improved Bounding Box Localization: YOLOv10's enhanced bounding box regression method ensures more precise localization of flames and smoke, even in dense or overlapping environments.
- Better Performance in Dense Environments: YOLOv10 excels in detecting flames and smoke in complex, object-rich environments, where YOLOv5s may struggle.
- Optimized Real-Time Detection: YOLOv10 processes high-resolution images more efficiently, ensuring faster real-time detection suitable for large-scale applications in fire safety.

6. MODULES:

Object Detection Engine:

This module, which serves as the system's foundation, is in charge of identifying and locating smoke and flames in pictures and video streams. This engine detects the objects and pinpoints their precise location in the image using the sophisticated YOLOv10 architecture. This module is perfect for real-time fire detection applications because of its fast speed and capacity to detect several items at once, guaranteeing that dangerous circumstances are quickly identified.

Model Adaptation through Pre-trained Networks:

This module makes use of pre-trained networks and transfer learning to improve the model's performance. A pre-trained YOLOv10 model that has been trained on sizable, broad datasets serves as the system's starting point. The system can then effectively learn fire-specific features by fine-tuning the model using a smaller, domain-specific fire dataset. This method speeds up the training process and reduces the requirement for a lot of labelled data.

Performance Tuning Module:

The goal of this module is to optimise the model's performance by adjusting its hyperparameters. To make sure the model converges fast and precisely, important parameters like learning rate, batch size, and number of epochs are changed. Optimising these parameters increases the accuracy of detection and guarantees that the system can function well under various circumstances, adjusting to all kinds of fire incidents.

Data Expansion and Augmentation:

The data augmentation module uses a variety of modifications, including as rotation, flipping, scaling, and lighting or contrast adjustments, to artificially expand the size of the training dataset. This procedure makes the system more resilient, allowing it to adapt better to novel and unexpected circumstances. This module makes sure that the fire detection system works in a variety of settings by modelling various real-world situations.

Model Training and Validation:

The process of training the fire detection model on labelled datasets and verifying its effectiveness on independent test data falls under the purview of this module. Using the training data, the model gains the ability to identify and categorise fire-related characteristics. In order to make sure the model is not overfitting and is capable of detecting fires in real-world situations, the validation step evaluates the model's capacity to generalise to new data.

Real-time Fire & Smoke Detection and Localization:

This module enables the system to identify smoke and flames in real time once it has been trained. It instantly detects the existence of fire-related threats after processing incoming video or picture streams. In the event of a fire, it also helps automated systems or emergency response teams respond quickly and

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

precisely by outputting the location of these threats inside the image or video frame.

Performance Evaluation and Monitoring:

Using common criteria like accuracy, precision, recall, and F1-score, the assessment module focuses on determining how successful the fire detection system is. It guarantees that the system is operating as planned and offers suggestions for enhancements. In order to ensure that the system continues to be effective and scalable in production settings, this module also monitors the system's capacity to manage massive amounts of data in real-time.

7. TECHNIQUE USED:

7.1 EXISTING TECHNIQUE: YOLOV5s

YOLOv5s, a real-time object detection variant of the YOLO (You Only Look Once) paradigm, is the current algorithm used in this research. YOLOv5s is renowned for being extremely quick and efficient because it can complete both localisation and classification in a single forward pass. It creates a grid out of an image and forecasts bounding boxes and class probabilities for every object in the grid. Convolutional layers are used by the model to extract features at various abstraction levels, which are subsequently routed to the detection head for object classification and localisation. YOLOv5s is a popular option for real-time applications like industrial monitoring and surveillance because of its performance optimisation.

YOLOv5s is useful for broad object detection jobs, however it has drawbacks when used for more complicated and specialised issues like fire detection. YOLOv5s may have trouble picking out delicate details like light smoke or faint flames in congested areas or low-visibility situations, such as when detecting smoke or in pictures with different fire intensities. This is because its architecture prioritises processing speed over detailed or high-resolution feature extraction. Furthermore, YOLOv5s lacks an attention mechanism, which prevents it from highlighting important areas of an image. This results in reduced detection accuracy when the fire or smoke is small, far away, or partially concealed. Its reliance on anchor boxes, which might not always capture the proper aspect ratios for irregularly shaped fire and smoke items and result in inaccurate bounding box predictions, is another drawback. The use of more sophisticated techniques, like YOLOv10, which can more successfully handle these issues, is required since these factors restrict YOLOv5s' performance in detecting fires and smoke under difficult real-world settings.

7.2 PROPOSED TECHNIQUE: YOLOv10:

YOLOv10, an improved version of the YOLO (You Only Look Once) model, is the suggested algorithm in this project. It improves object detection task performance, especially in smoke and fire detection. Compared to earlier iterations, especially YOLOv5s, YOLOv10 offers a number of enhancements that make it more efficient for specialised jobs like smoke and flame detection in difficult conditions. One of the main enhancements is the addition of an improved feature extraction method that records finer details of smoke and flames, which are frequently challenging to identify because of their small size, subtle look, or fluctuating intensity. This makes it possible for the model to differentiate between smoke and fire even in situations where sight is poor or congested.

The incorporation of an advanced attention mechanism is another noteworthy improvement in YOLOv10. By reducing unimportant background information, this method enables the model to concentrate on crucial regions of a picture that are more likely to contain smoke or flames. This selective attention enhances detection accuracy and lessens the influence of cluttered or loud backgrounds, especially in settings with overlapping items or other distractions. Furthermore, YOLOv10 employs an enhanced bounding box regression technique that guarantees more accurate localisation of smoke and flames, even in dense or challenging-to-understand photos.

Additionally, YOLOv10 is well-suited for real-time detection in large-scale applications, such as fire safety monitoring in industrial sites, residential neighbourhoods, or public spaces, because it is built to handle high-resolution photos and videos more effectively. Compared to its predecessors, YOLOv10 is more resilient and able to adjust to a variety of environmental situations thanks to these developments. It is the best option for precise, real-time fire and smoke detection because it works better than older models, such as YOLOv5s, in terms of precision, recall, and mean average precision (mAP).

YOLOv10's optimised architecture, which lowers computing complexity while preserving excellent detection accuracy, is another important benefit. Effective feature fusion and better training techniques are incorporated into the model to boost generalisation performance and convergence speed. This makes it possible for YOLOv10 to reliably identify smoke and fire in a variety of settings, such as urban areas, industrial sites, woodlands, and low light. It is ideal for deployment on edge devices, surveillance cameras, and drone-based monitoring systems because to its ability to strike a balance between speed and precision. It offers dependable early fire detection and aids in minimising potential damage through quick response.

8. SYSTEM ARCHITECTURE:

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

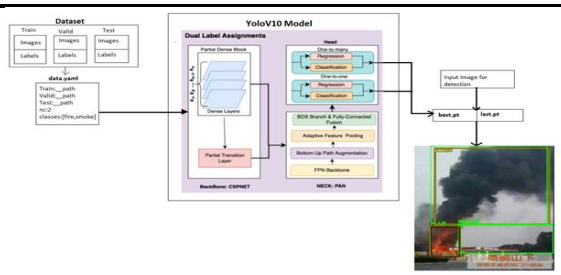

Fig 1 :System Architecture

9. RESULTS



Fig 2 : main website



Fig 2 : Register page

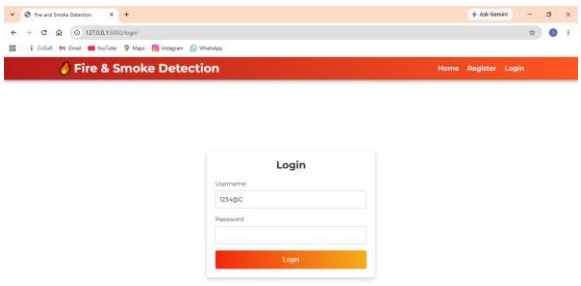

Fig 3: Login Page

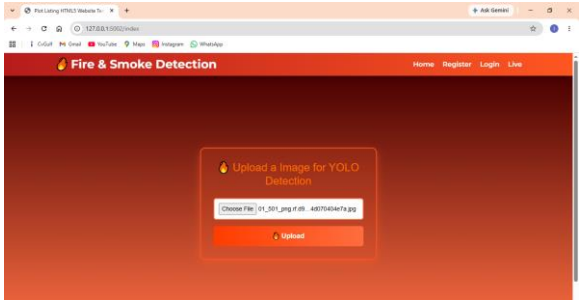

Fig 4:Image Upload

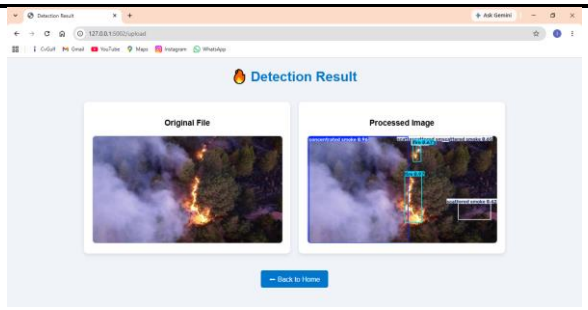

Fig 5: Result

10. DISCUSSION AND CONCLUSION:

In summary, this study shows how the YOLOv10 algorithm may be used to improve fire detection systems by offering a highly accurate and efficient way to identify smoke and flames in real time. The suggested solution greatly enhances detection performance in comparison to earlier techniques by tackling typical issues such crowded backdrops, fluctuating fire intensities, and poor visibility. Even in complicated and difficult situations, the system can precisely localise flames and smoke because to the incorporation of sophisticated algorithms including better bounding box regression, increased feature extraction, and attention processes.

The system is a useful tool for improving fire safety and prevention because of its real-time capabilities and adaptability to various circumstances. The system offers an efficient method of early fire detection in public areas, residential areas, and industrial facilities, potentially saving lives and minimising property loss. The system's stability and efficiency are further improved by the use of data augmentation and transfer learning, which makes it scalable and appropriate for a variety of applications.

All things considered, this project advances the development of more dependable and flexible fire detection technologies, opening the door for future developments that might include edge computing, multi-hazard recognition, and IoT integration to further enhance safety and reaction times in the event of a fire.

11. FUTURE ENHANCEMENTS:

The accuracy, scalability, and adaptability of the fire detection system can be further improved in a number of ways in the future. To expand the system's safety capabilities, one possible improvement would be to incorporate deep learning models that particularly target the detection of additional dangerous circumstances, including gas leaks or chemical spills. The system may offer a more complete safety solution for a variety of settings, such as urban regions and industrial sites, by fusing fire detection with multi-hazard recognition. Additionally, integrating edge computing could boost real-time performance and drastically cut down on

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

processing time. The technology may instantly offer fire alerts without depending on cloud servers by processing data locally on edge devices. This would be especially helpful in distant or low-connectivity areas. Drone or aerial monitoring, which uses UAVs with thermal and visual sensors to detect smoke and flames over vast, hard-to-reach areas like woods or industrial zones, could be another improvement. The use of multi-modal sensors, such as temperature, gas, and smoke detectors, to supplement image-based detection and create a more reliable and robust detection system is another area that could be improved in the future. Internet of Things (IoT) technology integration may provide smart integration with fire suppression equipment for prompt response and automatic alerts to emergency personnel.

Lastly, adaptive algorithms that take into account environmental factors like fog, heavy rain, or changing lighting conditions might be incorporated into the system to optimise it for use in a variety of weather circumstances. This would further improve safety and preventive measures in a variety of real-world settings by guaranteeing that the fire detection system operates dependably in all circumstances.

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. Gavrila 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] P. Shamsolmoali, M. Zareapoor, E. Granger, J. Chanussot, and J. Yang, ‘‘Enhanced single-shot detector for small object detection in remote sensing images,’’ in Proc. IEEE Int. Geosci. Remote Sens. Symp., 2022, pp. 1716–1719.

[8] D. Jiang, B. Sun, S. Su, Z. Zuo, P. Wu, and X. Tan, ‘‘FASSD: A feature fusion and spatial attention-based single shot detector for small object detection,’’ Electronics, vol. 9, no. 9, p. 1536, Sep. 2020.

[9] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, ‘‘YOLOv4: Optimal speed and accuracy of object detection,’’ 2020, arXiv:2004.10934.

[10] J. Redmon and A. Farhadi, ‘‘YOLO9000: Better, faster, stronger,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jul. 2017, pp. 6517–6525.

[11] Z. Ge, S. Liu, F. Wang, Z. Li, and J. Sun, ‘‘YOLOX: Exceeding YOLO series in 2021,’’ 2021, arXiv:2107.08430.

[12] X. Huang, X. Wang, W. Lv, X. Bai, X. Long, K. Deng, Q. Dang, S. Han, Q. Liu, X. Hu, D. Yu, Y. Ma, and O. Yoshie, ‘‘PP-YOLOv2: A practicalobject detector,’’ 2021, arXiv:2104.10419.

[13] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, ‘‘You only look once:Unified, real-time object detection,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 779–788.

[14] J. Redmon and A. Farhadi, ‘‘YOLOv3: An incremental improvement,’’ 2018, arXiv:1804.02767.

[15] Y.-L. Chung, H.-Y. Chung, and C.-W. Chou, ‘‘Efficient flame recognition method based on a deep convolutional neural network and image processing,’’ in Proc. IEEE 8th Global Conf. Consum. Electron. (GCCE), Oct. 2019, pp. 573–574.

[16] C. Chaoxia, W. Shang, and F. Zhang, ‘‘Information-guided flame detection based on faster R-CNN,’’ IEEE Access, vol. 8, pp. 58923–58932, 2020.

[17] H. Yar, Z. A. Khan, F. U. M. Ullah, W. Ullah, and S. W. Baik, ‘‘A modified YOLOv5 architecture for efficient fire detection in smart cities,’’ Exp. Syst. Appl., vol. 231, Nov. 2023, Art. no. 120465.

[18] M. N. Uddin, M. Sakibul Islam Sakib, S. Nawer, and R. T. Mohona, ‘‘Improved fire detection by YOLOv8 and YOLOv5 to enhance fire safety,’’ in Proc. 26th Int. Conf. Comput. Inf. Technol. (ICCIT), Dec. 2023, pp. 1–6.

[19] Z. Dai, ‘‘Image flame detection method based on improved YOLOv3,’’ IOP Conf. Ser., Earth Environ. Sci., vol. 693, Mar. 2021, Art. no. p. 012012.

[20] Y.-Y. Qin, J.-T. Cao, and X.-F. Ji, ‘‘Fire detection method based on depthwise separable convolution and YOLOv3,’’ Int. J. Autom. Comput., vol. 18, no. 2, pp. 300–310, Apr. 2021.

[21] X. Wu, Y. Cao, X. Lu, and H. Leung, ‘‘Patchwise dictionary learning for video forest fire smoke detection in wavelet domain,’’ Neural Comput. Appl., vol. 33, no. 13, pp. 7965–7977, Jul. 2021.

[22] Y. Cao, Q. Tang, X. Wu, and X. Lu, ‘‘EFFNet: Enhanced feature foreground network for video smoke source prediction and detection,’’ IEEE Trans.

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

Circuits Syst. Video Technol., vol. 32, no. 4, pp. 1820–1833, Apr. 2022.
[23] F. Yang, Q. Xue, Y. Cao, X. Li, W. Zhang, and G. Li, ‘‘Multi-temporal dependency handling in video smoke recognition: A holistic approach spanning spatial, short-term, and long-term perspectives,’’ Exp. Syst. Appl., vol. 245, Jul. 2024, Art. no. 123081.
[24] Y. Liu, Z. Shao, and N. Hoffmann, ‘‘Global attention mechanism: Retain information to enhance channel-spatial interactions,’’ 2021, arXiv:2112.05561.
[25] Q. Hou, D. Zhou, and J. Feng, ‘‘Coordinate attention for efficient mobile network design,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Jun. 2021, pp. 13713–13722.
[26] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, ‘‘ECA-net: Efficient channel attention for deep convolutional neural networks,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2020, pp. 11531–11539.
[27] S. Liu, D. Huang, and Y. Wang, ‘‘Learning spatial fusion for single-shot object detection,’’ 2019, arXiv:1911.09516.
[28] X. Wang, T. Xiao, Y. Jiang, S. Shao, J. Sun, and C. Shen, ‘‘Repulsion loss: Detecting pedestrians in a crowd,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jul. 2017, pp. 7774–7783.
[29] C.-Y. Wang, H.-Y. Mark Liao, Y.-H. Wu, P.-Y. Chen, J.-W. Hsieh, and I.-H. Yeh, ‘‘CSPNet: A new backbone that can enhance learning capability of CNN,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), Jun. 2020, pp. 1571–1580.
[30] T. Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, ‘‘Feature pyramid networks for object detection,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jul. 2017, pp. 2117–2125.
[31] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, ‘‘Path aggregation network for instance segmentation,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2018, pp. 8759–8768.
[32] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, ‘‘CBAM: Convolutional block attention module,’’ in Proc. Eur. Conf. Comput. Vis. (ECCV). Cham, Switzerland: Springer, 2018, pp. 3–19.
[33] X. Zhang, X. Zhou, M. Lin, and J. Sun, ‘‘ShuffleNet: An extremely efficient convolutional neural network for mobile devices,’’ in Proc. IEEE/CVFConf. Comput. Vis. Pattern Recognit., Jun. 2018, pp. 6848–6856.

Received: 06-10-2026 | Accepted: 19-11-2026 | Published: 25-11-2026