Real-Time Fire and Smoke Detection Using Optimized YOLOv11 with DCNv3 and CIOU Loss

K.Pavani[1] , P.Jaswanth Venkata sai[2]

#1 Assistant Professor & Head of Department of MCA, SRK Institute of Technology, Vijayawada.

#2 Student in the Department of MCA, SRK Institute of Technology, Vijayawada

Abstract: Fire and smoke detection is a critical task for ensuring safety in residential, industrial, and environmental monitoring systems, where early identification can prevent severe damage and loss of life. Traditional detection methods often rely on handcrafted features and sensor-based approaches, which are limited in handling complex backgrounds, varying illumination, and irregular fire patterns. To address these challenges, this paper proposes an enhanced deep learning-based detection framework using an improved YOLOv11 architecture.

The proposed model replaces the conventional DCNv2 detection head with DCNv3 to enable dynamic and adaptive sampling of features, effectively capturing the irregular and amorphous shapes of fire and smoke. Additionally, the Intersection over Union (IOU) loss function is replaced with Complete Intersection over Union (CIOU) loss to improve bounding box regression accuracy and localization precision. The system is designed to process both static images and real-time video streams through a Flask-based web interface, enabling dynamic monitoring and user interaction.

Experimental evaluation on benchmark fire and smoke datasets demonstrates that the proposed YOLOv11-DH3-CIOU model achieves superior performance, with a mean average precision (mAP) of approximately 95% and improved detection accuracy for small and partially occluded targets. The results indicate that the proposed system outperforms existing models in terms of precision, robustness, and real-time applicability, making it suitable for deployment in intelligent surveillance and disaster management systems.

Index terms - — *Fire Detection, Smoke Detection, YOLOv11, DCNv3, CIOU Loss, Deep Learning, Object Detection, Real-Time Monitoring, Computer Vision, Video Analysis*

1. INTRODUCTION

Fire and smoke detection is a critical component of modern safety systems, as early identification of fire hazards can significantly reduce damage to life, property, and the environment. Rapid urbanization, industrial growth, and increased use of combustible materials have heightened the risk of fire-related incidents, making reliable detection systems essential. Traditional fire detection methods, such as heat sensors and smoke detectors, often suffer from delayed response times and are prone to false alarms caused by environmental factors like dust, fog, or lighting variations. Similarly, conventional image processing techniques rely on manually designed features such as color, texture, and motion, which are insufficient for handling complex and dynamic real-world scenarios.

With the advancement of deep learning, computer vision-based approaches have shown significant improvements in object detection tasks. Among these, the YOLO (You Only Look Once) family of models has gained popularity due to its ability to perform real-time detection with high accuracy. However, standard YOLO models still face challenges in detecting small objects and irregular shapes, particularly in cluttered environments where fire and smoke patterns are highly dynamic and unpredictable. These limitations reduce detection reliability in practical applications such as surveillance, industrial monitoring, and disaster management.

To overcome these challenges, this paper proposes an enhanced fire and smoke detection system based on an improved YOLOv11 architecture. The proposed model replaces the conventional DCNv2 detection head with DCNv3, which enables adaptive feature sampling and improves the detection of irregular and amorphous shapes of fire and smoke. Furthermore, the use of Complete Intersection over Union (CIOU) loss enhances bounding box regression accuracy, leading to better localization of detected regions. The system is also extended to support both image and real-time video processing through a web-based interface, enabling dynamic monitoring and faster response.

The main contribution of this work is the development of a robust, accurate, and real-time fire and smoke detection framework that performs effectively in complex environments. Experimental results demonstrate that the proposed model achieves improved precision and a higher mean average precision (mAP) compared to existing approaches, making it suitable for real-world deployment in safety-critical applications.

2. LITERATURE SURVEY

a) Improving Fire and Smoke Detection with You Only Look Once 11 and Multi-Scale Convolutional Attention:

Property, human health, and safety are all seriously threatened by fires. Conventional approaches find it difficult to satisfy the needs of fire detection due to their wasteful utilization of characteristics. You Only Look Once (YOLO) is an effective deep learning object identification framework that can quickly find and recognize smoke and fire items in visual imagery. However, there is still a dearth of study using the most recent YOLO11 for fire and smoke detection, and the practicality of detection models as well as the size variability of fire and smoke objects continue to be research priorities. In order to examine YOLO11's benefits in fire and smoke detection tasks, this study first contrasts it with traditional models in the YOLO series. We then suggest a Multi-Scale Convolutional Attention (MSCA) mechanism and integrate it into YOLO11 to develop YOLO11s-MSCA in order to address the issues of scale variability and model practicality. According to experimental data, YOLO11 beats other YOLO models by striking a balance between speed, accuracy, and usefulness. On the D-Fire dataset, the YOLO11s-MSCA model performs remarkably well, increasing smoke recognition accuracy by 2.8% and total detection accuracy by 2.6%. The model shows a better capacity to recognize tiny smoke and fire items. The model has significant durability and generalization capabilities, retaining effective detection performance in complex contexts, despite ongoing difficulty in managing occluded objects and diverse backdrops.

b) A Fire and Smoke Detection Model Based on YOLOv8 Improvement:

People's lives and property are protected by fire and smoke alerts. Research on using deep learning for smoke and fire warnings has been ongoing; in particular, the application of target detection algorithms has produced notable outcomes. A high-precision, lightweight enhancement based on the You Only Look Once (YOLO) model is created to improve the model's fire and smoke detection capabilities in various settings. Partial convolutions are used to simplify the model, and an attention block is added to enable cross-space learning. Bidirectional feature fusion is also achieved by redesigning the neck network. Experiments reveal that it has greatly decreased the model's size while also improving the outcomes for every metric in the public Fire-Smoke dataset. The enhanced model also performs the best when compared to other well-known target identification methods under the same circumstances. The heatmap studies, which also show that it is characterized by a lower leakage rate and more concentrated attention, are developed to provide a more visual comparison with the detectability of the original model.

c) An Improved Fire and Smoke Detection Method Based on YOLOv8n for Smart Factories:

Factories are essential to both social and economic advancement. However, factory fire catastrophes pose a serious risk to property and human life. Prior research on deep learning-based fire detection mostly concentrated on detecting wildfires, ignoring factory fires. Furthermore, many research concentrate on fire detection, but many algorithms fail to identify smoke, a crucial byproduct of a fire tragedy. This research suggests an enhanced fire and smoke detection approach based on YOLOv8n to better assist smart factories in monitoring fire catastrophes. A self-made dataset of over 5000 photos and their labels is produced in order to guarantee the quality of the algorithm and training procedure. Nine sophisticated algorithms are then chosen and put to the test using the dataset. When it comes to detecting speed and accuracy, YOLOv8n has the greatest results. The inter-channel feature competition is then improved by inserting ConNeXtV2 into the backbone. RepBlock and SimConv are chosen to take the place of the original Conv and enhance memory bandwidth and processing power. To guarantee a precise and effective bounding box, MPDIoU is used in place of CIoU for the loss function. Our enhanced algorithm performs better in all four accuracy-related metrics—precision, recall, F1, and mAP@50—according to ablation tests. The updated method reaches above 95% when compared to the original model, whose four metrics are about 90%. In particular, mAP@50 achieves 95.6%, showing a 4.5% improvement. The needs of real-time fire and smoke monitoring are met even while complexity increases.

d) Fire and Smoke Detection Using Fine-Tuned YOLOv8 and YOLOv7 Deep Models:

Wildfires are considered a major natural catastrophe that poses a major risk to human populations, animals, and forest ecosystems. The effects of global warming and human involvement with the environment have played major roles in the recent increase in the frequency of wildfire incidents. In order to address this issue, firefighters must be able to quickly recognize fires based on early smoke indicators so they can act and stop them from spreading. In this study, we modified and improved the YOLOv8 and YOLOv7 models, two current deep learning object identification models, for smoke and fire detection. We used a collection of more than 11,000 photos of flames and smoke. With a mAP:50 of 92.6%, an accuracy score of 83.7%, and a recall of 95.2%, the YOLOv8 models were successful at identifying smoke and fire. A YOLOv6 with big model, Faster-RCNN, and DEtection TRansformer were used to compare the outcomes. The results attest to the suggested models' potential for widespread use and promotion in the fire safety sector.

e) Early Wildfire Smoke Detection Using Different YOLO Models:

Smoke is an early warning sign of forest fires, which are a major ecological hazard. Only a small percentage of the overall smoke is captured in early smoke photos. Smoke identification is hindered by small pixel-based characteristics due to the erratic dispersion of smoke and the dynamic nature of the surrounding environment. A novel framework that reduces the sensitivity of several YOLO detection algorithms is presented in this work. Furthermore, we contrast the detection speed and performance of several YOLO models, including YOLOv3, YOLOv5, and YOLOv7, with earlier models, including Fast R-CNN and Faster R-CNN. Additionally, to assess the detection model's accuracy in identifying smoke targets, we employ a gathered dataset that characterizes three different detection areas: close, medium, and long distance. On a multi-oriented dataset, our model detects forest smoke with a mAP accuracy of 96.8% at an IoU of 0.5 using YOLOv5x, outperforming the gold-standard detection approach. The study's results also demonstrate a significant increase in detection accuracy through the use of various data-augmentation strategies. Furthermore, with a mAP accuracy of 95% as opposed to 94.8% when using an SGD optimizer, YOLOv7 performs better than YOLOv3. Numerous studies demonstrate that the proposed approach maintains a respectable performance level in difficult environmental settings while achieving noticeably better results than the most sophisticated object-detection algorithms when applied to smoke datasets from wildfires.

3. METHODOLOGY

i) Proposed Work:

The proposed work focuses on developing an advanced fire and smoke detection system using an improved YOLOv11 model to achieve high accuracy in complex and dynamic environments. The system enhances the standard YOLOv11 architecture by replacing the traditional DCNv2 detection head with DCNv3, which enables adaptive and flexible sampling of features. This improvement allows the model to effectively capture the irregular and amorphous shapes of fire and smoke, especially in cluttered backgrounds and small target regions. Additionally, the IOU loss function is replaced with the Complete Intersection over Union (CIOU) loss to improve bounding box regression and localization precision.

The proposed system, referred to as YOLOv11-DH3-CIOU, integrates these enhancements while maintaining real-time detection capability. It is designed to process both images and video streams, enabling dynamic monitoring of fire incidents. A Flask-based web interface is developed to allow users to upload images or videos and visualize detection results with bounding boxes and confidence scores. The system is evaluated using benchmark datasets, demonstrating improved precision, robustness, and mean average precision (mAP) compared to existing models, making it suitable for real-world safety and surveillance applications.

ii) System Architecture:

The system architecture is designed to perform efficient and real-time fire and smoke detection by integrating deep learning with a web-based interface. According to the architecture diagram (page 26), the system follows a structured pipeline where input data is processed through multiple stages to generate accurate detection results.

The architecture begins with user input, where images or videos are uploaded through a Flask-based web interface. The input data is then passed to the preprocessing module, which performs operations such as resizing, normalization, and frame extraction (for videos). After preprocessing, the data is fed into the YOLOv11-DH3-CIOU detection model, which is the core component of the system. This model utilizes DCNv3 for adaptive feature extraction and CIOU loss for improved bounding box localization.

The detection module analyzes the input and identifies fire and smoke regions by generating bounding boxes along with confidence scores. For video inputs, the system processes frames sequentially to ensure real-time detection. The results are then sent to the output visualization module, where detected regions are displayed on images or video frames.

Additionally, the architecture includes user authentication, ensuring secure access to the system, and a server module that manages requests and responses efficiently. The overall design ensures scalability, real-time performance, and accurate detection in complex environments, making it suitable for surveillance and safety applications.

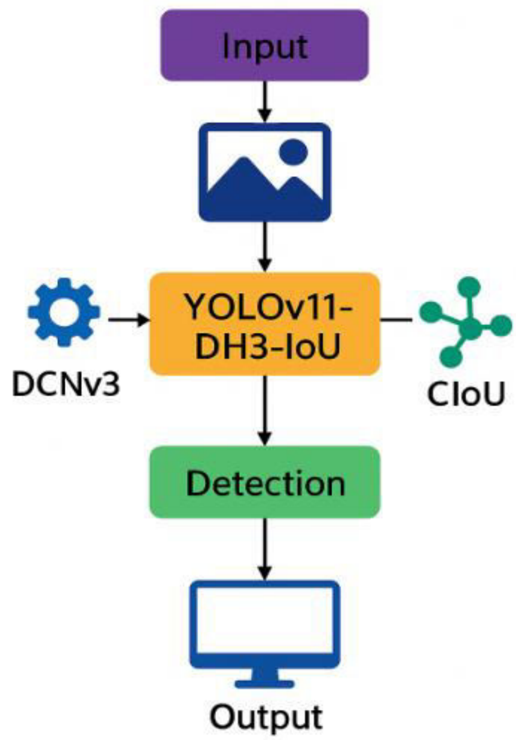

Fig1 proposed architecture

iii) Modules:

1. Importing Packages
Loads all required libraries such as deep learning frameworks, image processing tools, and visualization modules to support system execution.

2. DH3 Class (DCNv3 Implementation)
Implements the DCNv3 detection head, enabling adaptive sampling to capture irregular fire and smoke shapes more effectively.

3. Load Existing YOLOv11 Model
Loads the baseline YOLOv11 model to evaluate and compare performance with the proposed system.

4. Load Proposed YOLOv11-DH3-CIOU Model
Initializes the enhanced model incorporating DCNv3 and CIOU loss for improved detection accuracy.

5. Model Evaluation Module
Evaluates both models using metrics such as precision and mean average precision (mAP), and generates performance graphs.

6. Utility Functions for Image/Video Prediction
Handles preprocessing, model inference, and post-processing for both image and video inputs.

7. Flask Server Module
Provides a web-based interface that allows users to interact with the system by uploading media and viewing results.

8. User Authentication Module
Ensures secure login and access control, allowing only authorized users to use the system.

9. Fire & Smoke Detection (Image Module)
Processes uploaded images and detects fire and smoke regions with bounding boxes and confidence scores.

10. Fire & Smoke Detection (Video Module)
Processes video streams frame-by-frame to detect fire and smoke in real-time scenarios.

11. Output Visualization Module
Displays detection results clearly on images and videos, highlighting fire and smoke regions for user interpretation.

12. Logout Module
Handles user session termination to ensure system security after usage.

iv) Algorithms:

1. YOLOv11 (Existing Model)

YOLOv11 is a one-stage object detection algorithm that performs object localization and classification simultaneously in a single forward pass, enabling fast real-time detection. It uses convolutional neural networks with a DCNv2-based detection head to adapt to geometric variations in objects. The model predicts bounding boxes and class probabilities directly from input images, making it efficient for detecting fire and smoke. However, it has limitations in accurately detecting small objects and irregular shapes in complex environments.

2. YOLOv11-DH3-CIOU (Proposed Model)

The proposed YOLOv11-DH3-CIOU algorithm enhances the baseline model by integrating DCNv3 and CIOU loss for improved detection performance. DCNv3 enables dynamic and flexible sampling of feature maps, allowing better capture of irregular and amorphous fire and smoke patterns. The CIOU loss function improves bounding box regression by considering overlap area, distance between box centers, and aspect ratio, resulting in more precise localization. This model maintains real-time performance while significantly improving accuracy, especially in cluttered backgrounds and for small or partially visible fire regions.

4. EXPERIMENTAL RESULTS

The proposed fire and smoke detection system was evaluated using standard benchmark datasets, including the YOLO Fire and Smoke Monitoring dataset and the Baidu Paddle dataset. The performance of the enhanced YOLOv11-DH3-CIOU model was compared with the baseline YOLOv11 model using key evaluation metrics such as precision, recall, and mean average precision (mAP).

The experimental results demonstrate that the proposed model achieves a mean average precision (mAP) of approximately 95%, showing a significant improvement over the existing YOLOv11 model. The integration of DCNv3 enabled better feature extraction for irregular and small fire and smoke regions, while the CIOU loss function improved bounding box localization accuracy. As a result, the system produced fewer false positives and more accurate detections in complex and cluttered environments.

In addition to image-based evaluation, the system was tested on real-time video streams. The model successfully detected fire and smoke in dynamic scenes by processing video frames sequentially, maintaining high detection speed and accuracy. Visualization results include bounding boxes with confidence scores, clearly highlighting detected regions.

Overall, the experimental analysis confirms that the proposed YOLOv11-DH3-CIOU model outperforms existing approaches in terms of accuracy, robustness, and real-time performance, making it highly suitable for practical applications such as surveillance, industrial safety, and disaster management systems.

Accuracy: A test's accuracy is its capacity to distinguish healthy from ill cases. Find the percentage of instances with genuine positives and negatives to assess test accuracy.

Accuracy = TP + TN /(TP + TN + FP + FN)

$$Accuracy = \frac{(TN + TP)}{T}$$

Precision: Classification accuracy or positive cases constitute precision. The formula for accuracy is:

Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

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

Recall: A model's recall measures its ability to recognize all appropriate machine learning class instances. The ratio of accurately predicted positive observations to total positives indicates a model's class instance detection skill.

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

mAP: Mean Average Precision ranks quality. It considers the number and order of relevant ideas. Calculating MAP at K uses the arithmetic mean of each user or query's Average Precision (AP).

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k$$

$$AP_k = \text{the AP of class } k$$

$$n = \text{the number of classes}$$

F1-Score: A high F1 score suggests an accurate machine learning model. Integrating recall and precision improves model correctness. Accuracy measures how often a model predicts a dataset correctly.

$$F1 = 2 \cdot \frac{(Recall \cdot Precision)}{(Recall + Precision)}$$

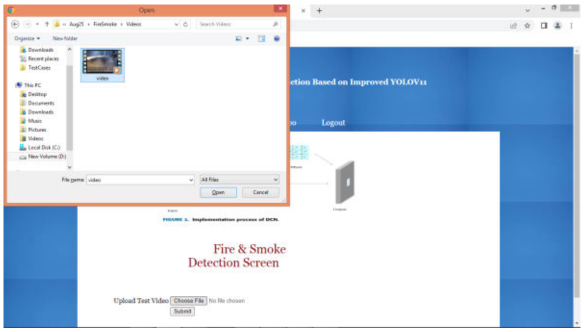

Fig2 Uploaded image

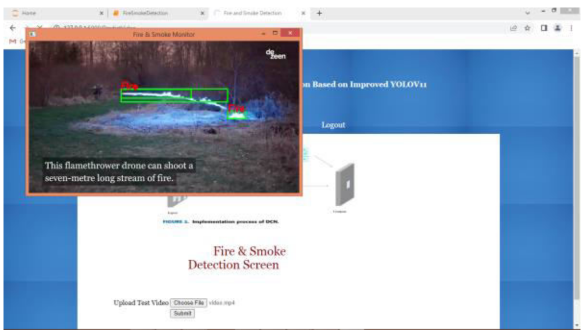

Fig3 Results

5. CONCLUSION

This paper presents an enhanced fire and smoke detection system based on an improved YOLOv11 architecture integrated with DCNv3 and CIOU loss. The proposed model effectively addresses the limitations of traditional and existing deep learning approaches by improving detection accuracy for small, irregular, and partially visible fire and smoke regions. The use of DCNv3 enables adaptive feature extraction, while CIOU loss enhances bounding box localization precision.

The system supports both image and real-time video detection through a user-friendly web interface, making it suitable for practical deployment in surveillance and safety applications. Experimental results demonstrate that the proposed YOLOv11-DH3-CIOU model achieves higher precision and approximately 95% mAP, outperforming the baseline YOLOv11 model. Overall, the system provides a robust, accurate, and efficient solution for real-time fire and smoke detection in complex environments.

6. FUTURE SCOPE

The proposed fire and smoke detection system can be further enhanced by integrating advanced technologies and expanding its application capabilities. Future improvements may include deploying the model on edge devices such as IoT-based surveillance systems and drones for faster, on-site fire detection without dependency on centralized servers. Incorporating thermal imaging and multi-sensor data fusion can improve detection accuracy under challenging conditions such as low visibility, fog, or nighttime environments.

Additionally, the system can be extended using transformer-based architectures or hybrid deep learning models to further enhance detection performance and generalization. Real-time alert systems such as SMS, alarms, or automated emergency response integration can be added for immediate action. The framework can also be scaled to detect multiple hazards such as gas leaks, explosions, or industrial anomalies, making it a comprehensive safety monitoring solution for smart cities and disaster management systems.

REFERENCES

[1] Li, Y., Nie, L., Zhou, F., Liu, Y., Fu, H., Chen, N., ... & Wang, L. (2025). Improving Fire and Smoke Detection with You Only Look Once 11 and Multi-Scale Convolutional Attention. *Fire*, *8*(5), 165.

[2] Gao, P. (2024). A Fire and Smoke Detection Model Based on YOLOv8 Improvement. *International Journal of Advanced Computer Science & Applications*, *15*(3).

[3] Zhang, Z., Tan, L., & Robert, T. L. K. (2024). An improved fire and smoke detection method based on YOLOv8n for smart factories. *Sensors*, *24*(15), 4786.

[4] Chetoui, M., & Akhloufi, M. A. (2024). Fire and smoke detection using fine-tuned YOLOV8 and YOLOV7 deep models. *Fire*, *7*(4), 135.

[5] Al-Smadi, Y., Alauthman, M., Al-Qerem, A., Aldweesh, A., Quaddoura, R., Aburub, F., ... & Alhmiedat, T. (2023). Early wildfire smoke detection using different yolo models. *Machines*, *11*(2), 246.

[6] H. Li and X. Li, ''Research on smoke filtering algorithm for forest fire images based on improved CycleGAN,'' Fire Sci. Technol., vol. 43, no. 11, pp. 1596–1602, 2024, doi: 10.20168/j.1009-0029.2024.11.1596.07.

[7] Y. Zhao and Y. Ban, ''GOES-R time series for early detection of wildfires with deep GRU-network,'' Remote Sens., vol. 14, no. 17, p. 4347, Sep. 2022, doi: 10.3390/rs14174347.

[8] J. Huang, J. Zhou, H. Yang, Y. Liu, and H. Liu, ''A small-target forest f ire smoke detection model based on deformable transformer for end to-end object detection,'' Forests, vol. 14, no. 1, p. 162, Jan. 2023, doi: 10.3390/f14010162.

[9] J. Lian, X. Pan, and J. Guo, ''An improved fire and smoke detection method based on YOLOv7,'' in Proc. 32nd Int. Conf. Comput. Commun. Netw. (ICCCN), Jul. 2023, pp. 1–7.

[10] R. Wang, R. Shivanna, D. Cheng, S. Jain, D. Lin, L. Hong, and E. Chi, ''DCN v2: Improved deep & cross network and practical lessons for web scale learning to rank systems,'' in Proc. Web Conf. (WWW), Ljubljana, Slovenia, Apr. 2021, pp. 1785–1797.

[11] H. Li, Y. Zhang, Y. Zhang, H. Li, L. Sang, and J. Zhu, ''DCNv3: Towards next generation deep cross network for CTR prediction,'' 2024, arXiv:2407.13349.

[12] S. Yang, Q. Huang, and M. Yu, ''Advancements in remote sensing for active fire detection: A review of datasets and methods,'' Sci. Total Environ., vol. 943, May 2024, Art. no. 173273, doi: 10.1016/j.scitotenv.2024.173273.

[13] R. Khanam and M. Hussain, ''YOLOv11: An overview of the key architectural enhancements,'' 2024, arXiv:2410.17725.

[14] B. Chen, M. Wei, J. Liu, H. Li, C. Dai, J. Liu, and Z. Ji, ''EFS-YOLO: A lightweight network based on steel strip surface defect detection,'' Meas. Sci. Technol., vol. 35, no. 11, Nov. 2024, Art. no. 116003.

Author Profiles

Mrs. K. Pavani is working as an Assistant and Head of Department of MCA, in SRK Institute of technology in Vijayawada. She completed her MCA and M.Tech in Computer Science. She has 10 years of teaching experience in SRK Institute of technology, Enikepadu, Vijayawada, NTR District. Her areas of interest include AI and ML, etc.

Mr.p.jaswanth is an MCA Student in the Department of Computer Application at SRK Institute Of Technology, Enikepadu, Vijayawada, NTR District. He has Completed Degree in B.Sc.(computers) from A.K.R.G Degree College Nallajerla. His area of interest are DBMS and Machine Learning with Python.