International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Peer Reviewed, Referred & Indexed Journal

Research Paper

Fire and Smoke Detection Based on Improved YOLOV11

Muddangula Lavanya
PG , Department of computer science and engineering,
Sree Dattha institutions of Engineering & Science.
Hyderabad, Telangana, India-501510.
Email: muddangullavanya@gmail.com

Dr. GNV Vibhav Reddy
Associate Professor, Department of computer science and Engineering,Sree Dattha institutions of Engineering & Science. Hyderabad, Telangana, India-501510.
Email: vicechairman@sreedattha.ac.in

***Abstract*— Fire and smoke detection plays a vital role in safeguarding human life, infrastructure, and ecological systems by enabling early identification of hazardous events. Conventional vision-based detection techniques rely heavily on handcrafted features, which often fail to deliver reliable performance under complex environmental conditions such as varying illumination, background clutter, and smoke density. To address these limitations, a deep learning–based fire and smoke detection framework is developed using advanced YOLO architectures. A curated image dataset containing fire and smoke scenes is utilized, where bounding box annotations are prepared in YOLO format for object localization. Preprocessing includes image normalization, dataset structuring, and configuration through a unified data file to support both classification and detection pipelines. Multiple YOLO variants, including YOLOv5, YOLOv8, YOLOv9, YOLOv10, YOLOv11s, and an enhanced YOLOv11-DH3 model, are trained and evaluated. In addition, the latest YOLOv11x model is incorporated to assess performance improvements. Model evaluation is conducted using standard metrics such as precision, recall, and mean average precision (mAP). Experimental results demonstrate that YOLOv11x achieves the best performance, attaining a precision of 0.930, recall of 0.981, and mAP of 0.967. The system is further integrated into a Flask-based web interface to enable real-time fire and smoke detection, enhancing practical applicability and deployment efficiency.**

"Keywords— Fires, Feature extraction, Shape, YOLO, Accuracy, Detectors, Monitoring, Image color analysis, Web and internet services, Deep learning."

I. INTRODUCTION

Fire incidents remain one of the most destructive hazards affecting modern society, posing severe threats to human safety, infrastructure, and ecological systems. With the rapid growth of urbanization and the increasing frequency of large-scale public events, fire risks have expanded from industrial zones to densely populated residential areas and natural environments such as forests. Fireworks, electrical faults, industrial activities, and climatic factors contribute to the rising probability of fire outbreaks, making prevention and monitoring essential components of public safety management [1]. Historical events demonstrate that uncontrolled fires can lead to extensive loss of life, large-scale property destruction, and long-term environmental damage, emphasizing the importance of early detection and timely intervention [2].

Despite continuous advancements in fire safety systems, significant challenges persist in achieving reliable early detection under complex real-world conditions. Conventional fire monitoring methods often rely on physical or chemical sensing mechanisms, which are susceptible to environmental interference and limited spatial coverage [3]. These systems may generate false alarms or fail to distinguish between fire-related and non-fire-related phenomena, particularly in dynamic or visually cluttered environments [4]. Moreover, many existing detection approaches struggle to provide robust performance across diverse scenarios, including variations in lighting, weather, background complexity, and fire scale, revealing a critical gap in achieving accurate and adaptable fire and smoke monitoring solutions [5].

The objective of this work is to develop an intelligent visual detection framework capable of identifying fire and smoke events with high reliability and responsiveness. The proposed approach aims to enhance detection accuracy while maintaining operational efficiency suitable for real-time monitoring applications. By leveraging large-scale visual data and advanced learning-based detection paradigms, the framework seeks to address the limitations of traditional systems and improve robustness across diverse environments [6]. The main contributions include the systematic evaluation of multiple detection paradigms, comprehensive performance analysis under unified conditions, and the establishment of a scalable monitoring solution applicable to both urban and natural settings [7].

The significance of this work lies in its potential to strengthen early warning capabilities and support rapid emergency response, thereby minimizing casualties and economic losses caused by fire disasters. Improved fire and smoke detection can enhance situational awareness for safety authorities, assist in evacuation planning, and provide critical decision support during emergencies [8]. Furthermore, the deployment of intelligent visual monitoring systems can reduce maintenance costs, lower false alarm rates, and enable continuous large-area surveillance [9]. Ultimately, this contribution supports the advancement of intelligent fire prevention infrastructure, promoting safer cities and more resilient environmental protection strategies [10].

IJERST International Journal of Engineering Research and Science & Technology

II. LITERATURE REVIEW

Recent advances in computer vision have significantly improved the capability of fire and smoke detection systems by leveraging learning-based visual representations. Lin et al. introduced a video-based smoke detection approach using three-dimensional convolutional neural networks, demonstrating strong temporal feature extraction and improved detection accuracy in dynamic scenes [11]. However, the high computational cost of 3D models limits their real-time applicability. Li and Li focused on smoke filtering in forest fire imagery using generative learning, achieving enhanced visual clarity under complex backgrounds, yet their method is primarily oriented toward image enhancement rather than direct detection [12].

To address efficiency constraints, Jeong et al. proposed a lightweight recurrent framework for real-time wildfire smoke detection, emphasizing reduced model complexity and faster inference [13]. While effective for early smoke identification, recurrent models often struggle with spatial localization precision. Zhao and Ban explored time-series satellite data for early wildfire detection, highlighting the potential of remote sensing for large-area monitoring [14]. Despite its broad coverage, satellite-based detection faces challenges related to temporal resolution, cloud interference, and delayed response, which restrict its use in urgent scenarios.

Earlier works also investigated traditional feature-driven approaches. Jamali et al. presented a saliency-based fire detection method combining texture and color features, offering interpretability and low computational demand [15]. Nonetheless, handcrafted features are highly sensitive to environmental variations and lack robustness in complex scenes. More recent studies have shifted toward end-to-end detection paradigms. Huang et al. developed a small-target forest fire smoke detection model emphasizing improved sensitivity to distant or weak smoke patterns [16]. Although effective for small targets, such approaches often require complex model designs and extensive training data.

Single-stage object detection frameworks have gained popularity due to their balance between speed and accuracy. Lian et al. enhanced a real-time detection framework for fire and smoke recognition, achieving improved performance in comparison to earlier detectors [17]. Foundational work by Redmon et al. established unified real-time object detection, enabling fast and efficient visual recognition [18]. Subsequent architectural improvements, such as deformable convolution mechanisms, further enhanced spatial adaptability and feature alignment in complex visual environments [19]. Additionally, cross-network learning strategies have been explored to improve feature interaction and generalization performance in large-scale learning tasks [20].

Despite these advancements, existing approaches still face challenges in maintaining high detection accuracy across diverse environments while ensuring real-time performance and deployment efficiency. Many methods are either computationally intensive, sensitive to environmental noise, or limited in generalization across fire scenarios. These gaps motivate the need for a more robust and scalable detection framework that balances accuracy, speed, and adaptability. The present work builds upon recent detection paradigms to address these limitations, aiming to enhance detection reliability and practical applicability in real-world fire and smoke monitoring systems.

III. MATERIALS AND METHODS

The proposed system is designed to achieve efficient and reliable fire and smoke detection using deep learning–based object detection models trained on a curated image dataset annotated in YOLO format for precise localization. The overall workflow follows a unified detection pipeline encompassing dataset organization, normalization, and configuration to ensure consistent training and evaluation. Multiple YOLO-based architectures, including YOLOv5, YOLOv8, YOLOv9, YOLOv10, and an improved YOLOv11-DH3 model, are employed to learn discriminative spatial representations of fire and smoke under diverse visual conditions. The YOLOv11-DH3 model incorporates architectural refinements that enhance feature representation and robustness in complex scenes, while the integration of the latest YOLOv11x model further improves detection accuracy and generalization. System performance is evaluated using precision, recall, and mean average precision metrics. To support real-world applicability, the detection framework is deployed through a Flask-based web interface, enabling real-time inference and user interaction. This integrated approach enhances early hazard identification, localization accuracy, and practical deployment efficiency for safety monitoring systems.

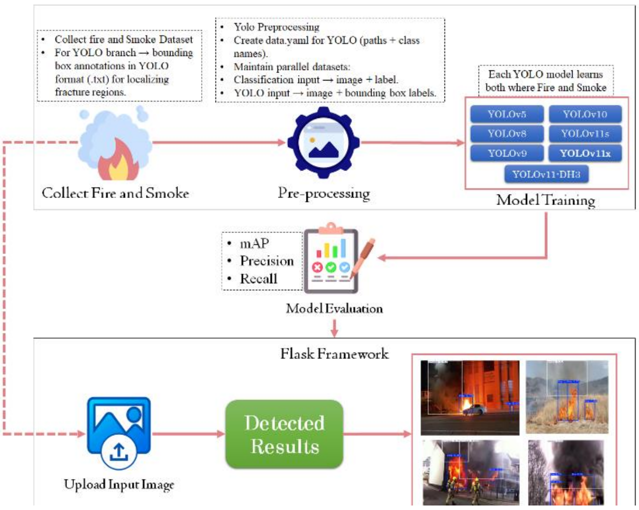


Fig. 1. Proposed Architecture

The system architecture is organized into four layers: data, experimentation, core model, and application. The data layer handles dataset acquisition, YOLO-based annotation, and configuration. The experiment and training layer benchmarks multiple YOLO variants through comparative evaluation. The core model layer performs real-time inference on image or video inputs, generating bounding boxes, class labels, and confidence scores. Finally, the application layer deploys the trained model via a web-based interface, enabling visualization, user interaction, and real-time fire and smoke detection.

A) Dataset Collection

The experiments are conducted using a curated fire and smoke image dataset obtained from a publicly available online repository. The dataset comprises diverse visual samples containing fire and smoke instances captured under varying environmental conditions, backgrounds, and illumination levels. Each image is manually annotated with bounding boxes in YOLO format, covering multiple object classes related to fire and smoke. The dataset exhibits high variability in scene complexity and object scale, making it well suited for training and evaluating robust object detection models. This diversity supports effective generalization and reliable performance in real-world fire monitoring scenarios.

Fig. 2. Dataset

B) Pre-Processing

Dataset preprocessing ensures structured, standardized, and well-annotated data for effective learning. It prepares images and labels for both classification and detection tasks, improving training stability, accuracy, and generalization performance.

C) Data Preprocessing:

The preprocessing stage standardizes the dataset to support consistent and efficient model training across multiple learning pipelines. Images are normalized to reduce variations caused by illumination and scale differences, while class labels are carefully aligned with annotation files. A unified configuration file is created to define dataset paths and class information for detection tasks. Parallel dataset structures are maintained for classification and detection, ensuring compatibility across models and improving training stability, convergence speed, and generalization capability in real-world scenarios.

D) Algorithms

YOLOv5: YOLOv5 serves as an efficient single-stage object detector that enables fast and accurate localization by jointly predicting bounding boxes and class probabilities. Its balanced design supports real-time inference while maintaining reliable detection performance across varied visual conditions.

YOLOv8: YOLOv8 enhances detection robustness through improved feature aggregation and anchor-free prediction mechanisms. By effectively capturing multi-scale spatial information, it improves localization accuracy and generalization, particularly in complex and dynamically changing visual environments.

YOLOv9: YOLOv9 focuses on refined feature representation and precise bounding box regression to improve detection reliability. Its multi-scale detection capability supports accurate object identification under challenging backgrounds, varied object sizes, and diverse environmental conditions.

YOLOv10: YOLOv10 emphasizes architectural optimization for faster inference and improved accuracy. By integrating efficient feature fusion and detection heads, it supports precise localization and consistent performance in real-time monitoring and safety-critical applications.

YOLOv11s: YOLOv11s is a lightweight detection architecture designed for rapid inference with minimal performance degradation. It effectively balances computational efficiency and detection accuracy, ensuring robust object recognition under occlusion, illumination variation, and cluttered scenes.

YOLOv11-DH3: YOLOv11-DH3 introduces architectural refinements to strengthen feature extraction and detection stability. Enhanced feature fusion improves robustness and localization precision, supporting reliable object detection in complex and visually challenging environments.

YOLOv11x: YOLOv11x represents an advanced detection architecture combining refined feature extraction and multi-scale representation. It achieves high precision and strong generalization, enabling accurate and stable real-time object detection across diverse and demanding scenarios.

IV. Experimental Results

Precision: Precision evaluates the fraction of correctly classified instances or samples among the ones classified as positives. Thus, the formula to calculate the precision is given by:

$$Precision = \frac{\text{True Positive}}{\text{True Positive } + \text{False Positive}} \quad (1)$$

Recall: Recall is a metric in machine learning that measures the ability of a model to identify all relevant instances of a particular class. It is the ratio of correctly predicted positive observations to the total actual positives, providing insights into a model's completeness in capturing instances of a given class.

$$Recall = \frac{\text{TP}}{\text{TP } + \text{ FN}} \quad (2)$$

mAP: Mean Average Precision (MAP) is a ranking quality metric. It considers the number of relevant recommendations and their position in the list. MAP at K is calculated as an arithmetic mean of the Average Precision (AP) at K across all users or queries.

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k \quad (3)$$

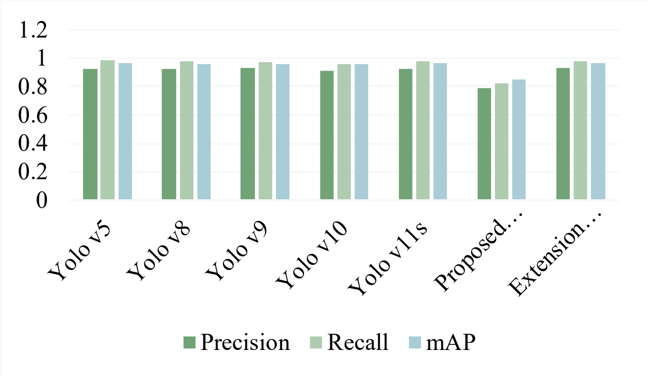
Table. 1. Performance Evaluation

ML Model	Precision	Recall	mAP
Yolo v5	0.925	0.983	0.964
Yolo v8	0.925	0.979	0.962
Yolo v9	0.929	0.973	0.961
Yolo v10	0.909	0.959	0.960
Yolo v11s	0.927	0.980	0.963

Proposed Yolo 11-DH3	0.792	0.822	0.852
Extension Yolo v11x	0.930	0.981	0.967

The performance evaluation compares multiple YOLO models using precision, recall, and mAP. YOLOv11x achieves the best overall results, indicating superior detection accuracy, robustness, and generalization compared to other variants.

Fig. 3. Comparison Graph



The comparison graph illustrates precision, recall, and mAP across YOLO variants, highlighting YOLOv11x as the top-performing model, while YOLOv11-DH3 shows comparatively lower performance among evaluated architectures.

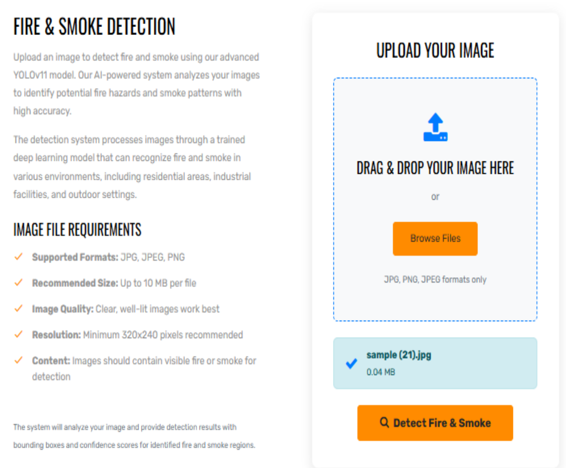


Fig.4 Upload a Image

The interface displays an image upload module for fire and smoke detection, allowing users to submit images and initiate analysis through an interactive web-based system for real-time hazard identification.

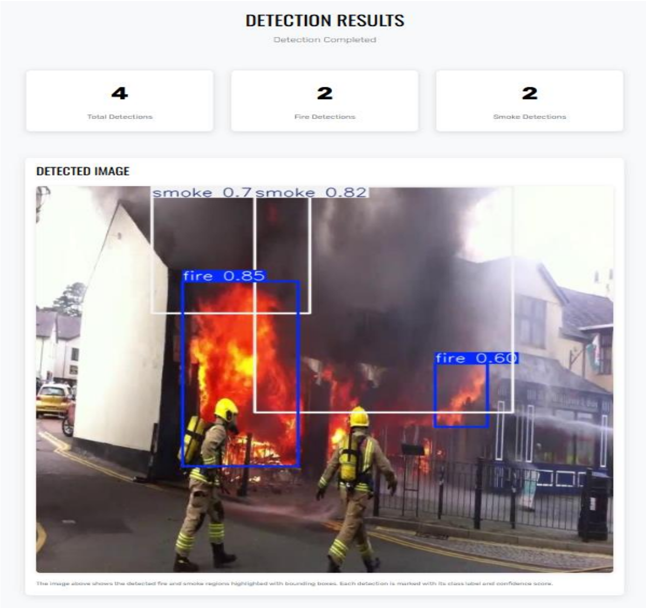


Fig.5 Predicted Result

The output image presents detected fire and smoke regions highlighted with bounding boxes and confidence scores, demonstrating accurate localization and classification of hazardous areas to support effective emergency response and safety monitoring.

V. Conclusion

In conclusion, this work focuses on developing an efficient and reliable fire and smoke detection system to enhance early hazard identification and improve safety in complex environments. The proposed approach utilizes a curated dataset comprising annotated fire and smoke images and employs advanced deep learning–based object detection techniques for accurate localization and classification. Multiple YOLO architectures, including YOLOv5, YOLOv8, YOLOv9, YOLOv10, YOLOv11s, and an enhanced YOLOv11-DH3 model, are systematically trained and evaluated to analyze their detection capabilities. The framework leverages standardized preprocessing, structured data configuration, and robust training strategies to ensure consistency and reliability across models. In addition, the latest YOLOv11x model is incorporated to further strengthen detection accuracy and generalization performance. To support practical deployment, a lightweight Flask-based web interface is integrated, enabling users to upload images and obtain real-time detection results through an accessible platform. Experimental evaluation demonstrates that the proposed system achieves superior performance, with YOLOv11x attaining a precision of 0.930, recall of 0.981, and mAP of 0.967. Overall, the developed framework offers a reliable, efficient, and scalable solution for automated fire and smoke detection, contributing significantly to safety monitoring and real-world decision support systems.

The system's future scope includes further enhancing detection accuracy and adaptability across diverse and complex environments by integrating multi-sensor data such as infrared or thermal imagery. Optimization techniques, including lightweight model architectures, can enable deployment on edge devices for real-time, low-latency applications. Incorporating explainable AI methods could improve interpretability, allowing users to understand model decisions and increase trust. Additionally, extending the framework to handle video streams or continuous monitoring can support proactive hazard management. Cloud-based deployment and mobile integration can further improve accessibility and scalability, enabling widespread adoption in urban, industrial, and ecological safety monitoring scenarios.

REFERENCES

[1] Alkhammash, E. H. (2025). A Comparative Analysis of YOLOv9, YOLOv10, YOLOv11 for Smoke and Fire Detection. Fire (2571-6255), 8(1).

[2] 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.

[3] Hu, J., He, Y., Zeng, M., Qian, Y., & Zhang, R. (2024). Smoke and fire detection based on yolov7 with convolutional structure reparameterization and lightweighting. IEEE Sensors Letters, 8(8), 1-4.

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

[5] Geng, X., Su, Y., Cao, X., Li, H., & Liu, L. (2024). YOLOFM: An improved fire and smoke object detection algorithm based on YOLOv5n. Scientific Reports, 14(1), 4543.

[6] Government Set to Revise Total Number of Hectares Destroyed During Bushfire Season, 9News, Sydney, NSW, Australia, Feb. 2020.

[7] 58 People Died in the High-Rise Residential Fire in Jing'an District, Shanghai, China News Service, Beijing, China, 2010.

[8] Theon-Site Rescue Work of the Fire and Explosion Accident at Jilin Dehui Poultry Industry Company Has Basically Ended, Xinhua News Agency, Beijing, China, 2013.

[9] V. K. R. Kodur and A. M. Shakya, ''Factors governing the shear response of prestressed concrete hollowcore slabs under fire conditions,'' Fire Saf. J., vol. 88, pp. 67–88, Mar. 2017, doi: 10.1016/j.firesaf.2017.01.003.

[10] N. Zheng, X. Guo, Y. Tie, N. Dong, L. Qi, and L. Guan, ''Incremental generalized multiple maximum scatter difference with applications to feature extraction,'' J. Vis. Commun. Image Represent., vol. 55, pp. 67–79, Aug. 2018, doi: 10.1016/j.jvcir.2018.04.012.

[11] G.Linetal.,''Smokedetectiononvideosequencesusing3Dconvolutional neural networks,'' Fire Technol., vol. 55, pp. 1827–1847, 2019, doi: 10.1007/s10694-019-00832-w.

[12] 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.

[13] M. Jeong, M. Park, J. Nam, and B. C. Ko, ''Light-weight student LSTM for real-time wildfire smoke detection,'' Sensors, vol. 20, no. 19, p. 5508, Sep. 2020, doi: 10.3390/s20195508.

[14] 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.

[15] M.Jamali, N. Karimi, and S. Samavi, ''Saliency based fire detection using texture andcolorfeatures,'' in Proc.28thIranianConf.Electr.Eng.(ICEE), Aug. 2020, pp. 1–5.

[16] 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.

[17] J.Lian,X.Pan,andJ.Guo,''Animprovedfireandsmokedetectionmethod based on YOLOv7,'' in Proc. 32nd Int. Conf. Comput. Commun. Netw. (ICCCN), Jul. 2023, pp. 1–7.

[18] 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. (CVPR), Jun. 2016, pp. 779–788, doi: 10.1109/CVPR.2016.91.

[19] J. Dai, H. Qi, Y. Xiong, Y. Li, G. Zhang, H. Hu, and Y. Wei, ''Deformable convolutional networks,'' in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Oct. 2017, pp. 764–773, doi: 10.1109/ICCV.2017.89.

[20] 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.

[21] 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.

[22] P. Foggia, A. Saggese, and M. Vento, ''Real-time fire detection for video-surveillance applications using a combination of experts based on color, shape, and motion,'' IEEE Trans. Circuits Syst. Video Technol., vol. 25, no. 9, pp. 1545–1556, Sep. 2015, doi: 10.1109/TCSVT.2015. 2392531.

[23] 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.

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

[25] R. M. Haralick, K. Shanmugam, and I. Dinstein, ''Textural features for image classification,'' IEEE Trans. Syst. Man, Cybern., vol. SMC-3, no. 6, pp. 610–621, Nov. 1973, doi: 10.1109/TSMC.1973.4309314.

[26] 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.

[27] S. Hao, X. Li, W. Peng, Z. Fan, Z. Ji, and I. Ganchev, ''YOLO-CXR: A novel detection network for locating multiple small lesions in chest X-ray images,'' IEEE Access, vol. 12, pp. 156003–156019, 2024.

[28] M. Tan, R. Pang, and Q. V. Le, ''EfficientDet: Scalable and efficient object detection,'' in Proc. IEEE/CVFConf.Comput.Vis.PatternRecognit. (CVPR), Jun. 2020, pp. 10778–10787.

[29] K. Gajendiran, S. Kandasamy, and M. Narayanan, ''Influences of wildfire on the forest ecosystem and climate change: A comprehen sive study,'' Environ. Res., vol. 240, Jan. 2024, Art. no. 117537, doi: 10.1016/j.envres.2023.117537.

[30] K. Muhammad, J. Ahmad, Z. Lv, P. Bellavista, P. Yang, and S. W. Baik, ''Efficient deep CNN-based fire detection and localization in video surveillance applications,'' IEEE Trans. Syst. Man, Cybern. Syst., vol. 49, no. 7, pp. 1419–1434, Jul. 2019.