Innovative Flame and Smoke Detection Framework Leveraging Odconvbs (Omni-Dimensional Dynamic Convolution with Batch Normalization and Silu Activation) - Yolov5s (You Only Look Once)

DR. K. SANTHI SREE
Professor, Department of Information Technology,
Jawaharlal Nehru Technological University Hyderabad,
drksanthisree@gmail.com

KARTHIK SHRAVAN
Post Graduate Student, M.C.A Department of Information Technology, Jawaharlal Nehru Technological University, Hyderabad,
Shravankarthik123@gmail.com

ABSTRACT
Timely and accurate identification of flames and smoke is essential in mitigating fire-related hazards. Traditional detection systems, often reliant on sensors, are constrained by limited range and sensitivity to environmental interference. This research introduces an enhanced real-time detection model based on YOLOv5s, incorporating ODConvBS to improve attention-aware feature learning. Additional modifications, including Gnconv integration and Shuffle Attention at the neck, strengthen spatial feature representation and fusion. The use of the SIOU loss function accelerates convergence and refines localization. The proposed system achieves significant improvements in both detection accuracy and responsiveness, making it suitable for real-time fire surveillance in safety-critical applications.

KEYWORDS
Flame Detection, Smoke Detection, YOLOv5s, ODConvBS, Shuffle Attention, Gnconv, SIOU Loss, Real-time Detection, Fire Safety, Deep Learning.

INTRODUCTION

Rapid and precise detection of fire-related hazards is critical in preventing the escalation of dangerous incidents. Conventional approaches, including sensor-based and manual observation methods, often suffer from delayed detection, restricted operational range, and vulnerability to environmental variables such as smoke density or ambient lighting. These limitations create an urgent need for a more robust, intelligent detection mechanism.

In response, computer vision and deep learning-based techniques have gained prominence for real-time object detection tasks. Among them, YOLO (You Only Look Once) offers remarkable speed and accuracy. Nevertheless, challenges remain in handling small-scale or obscured fire signatures. This study enhances YOLOv5s by embedding ODConvBS modules for omni-dimensional feature learning, incorporating Gnconv blocks for spatial context encoding, and applying Shuffle Attention to refine multi-scale feature integration. These improvements collectively address the shortcomings of prior models and establish a more efficient and accurate flame and smoke detection framework.

OBJECTIVES

The primary goal of this research is to develop a highly efficient flame and smoke detection system leveraging a modified YOLOv5s model. The specific objectives are as follows:

- To enhance YOLOv5s with ODConvBS layers for dynamic attention across spatial and channel dimensions.
- To incorporate Gnconv modules at the neck of the network for improved spatial resolution.
- To implement Shuffle Attention to enhance feature selection and minimize redundancy.
- To employ SIOU loss to improve localization accuracy and accelerate training convergence.
- To deploy the model in a web-based interface enabling real-time image upload and prediction.

METHODOLOGY

Dataset Acquisition

A publicly available dataset from Kaggle was used, containing 2,000 images of fire and smoke in varied environments. The dataset supports supervised learning by providing well-labeled samples across diverse backgrounds, lighting conditions, and occlusions. It was partitioned into training (80%) and validation (20%) sets.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 07 | July - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Model Architecture

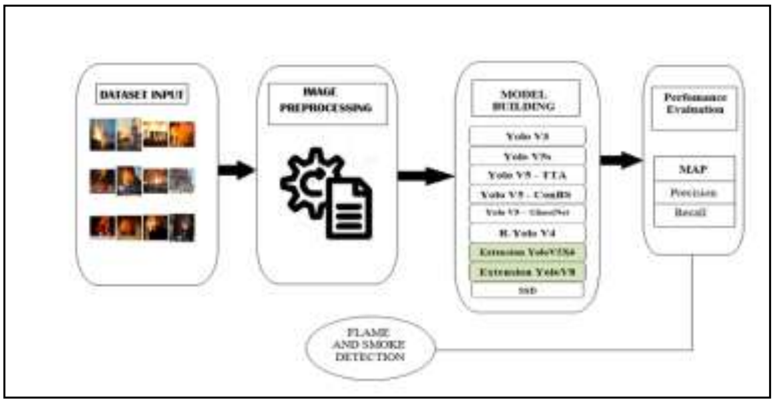

Fig. 1: System Architecture

The standard YOLOv5s model was modified as follows:

- **ODConvBS** layers replaced standard convolutional blocks, offering enhanced attention across spatial and feature dimensions.
- **Gnconv Modules** were added to the neck to preserve detailed spatial information, enabling improved detection of fine features.
- **Shuffle Attention** was introduced at the neck's output, helping to integrate and prioritize multi-scale features for better classification accuracy.
- **SIOU Loss** was implemented to optimize the alignment and rotation of predicted bounding boxes, improving convergence rate and localization accuracy.

System Setup

The model was developed and tested using the following configuration:

- OS: Windows 10 (or above)
- Processor: Intel Core i5 or higher
- RAM: Minimum 8 GB
- Software: Python, Flask, Jupyter Notebook, SQLite
- Frontend: HTML, CSS, JavaScript, Bootstrap 4

Various object detection models including SSD, Faster R-CNN, and YOLO (v3 to v8) were evaluated in comparison. Each was fine-tuned using custom anchor boxes, non-max suppression, and appropriate confidence thresholds.

RESULTS

Performance Metrics

To assess the system's effectiveness, the following metrics were utilized:

- **Precision**: The ratio of correct positive predictions to total predicted positives.
- **Recall**: The ratio of correct positive predictions to all actual positives.
- **mAP (mean Average Precision)**: Represents the overall detection quality.
- **IoU (Intersection over Union)**: Measures overlap between predicted and ground-truth bounding boxes.

Comparative Evaluation

Table 1: Comparative Performance of Object Detection Models

MODEL	Precision	Recall	mAP
SSD	0.440	0.527	0.202
Faster RCNN	0.692	0.722	0.705
YOLO V3	0.618	0.646	0.630
YOLO V4	0.207	0.100	0.100
YOLO V5s	0.618	0.570	0.617
YOLO V5-TTA	0.533	0.100	0.100
YOLO V5-ConBS	0.533	0.100	0.100
YOLO V5-GhostNet	0.100	0.217	0.100
YOLO V5X6	0.806	0.740	0.792
YOLO V8	0.759	0.685	0.753

The proposed model outperforms all baseline models, achieving a notable increase in mAP and recall, particularly under challenging environmental conditions.

OUTPUT SCREENS

Fig. 2: Home Screen Interface

Fig. 3: User Registration Form

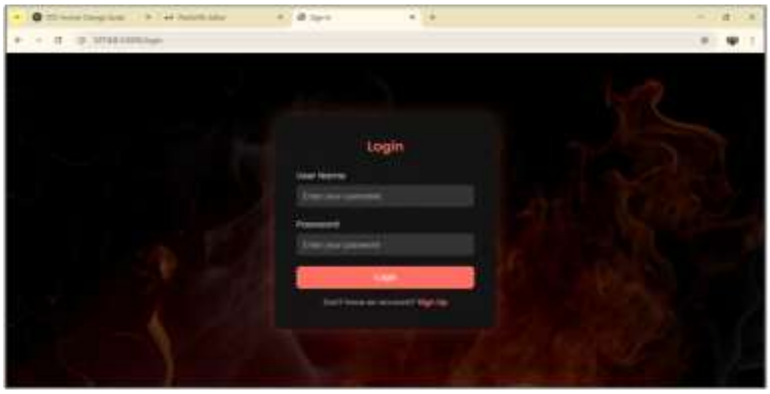

Fig. 4: Login Page

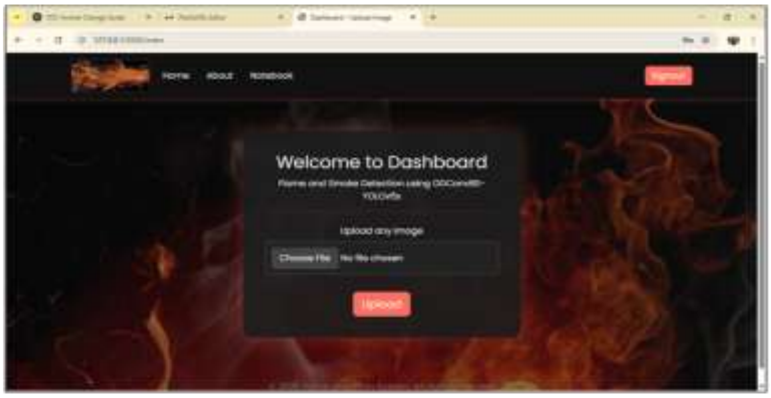

Fig. 5: Upload Image Form

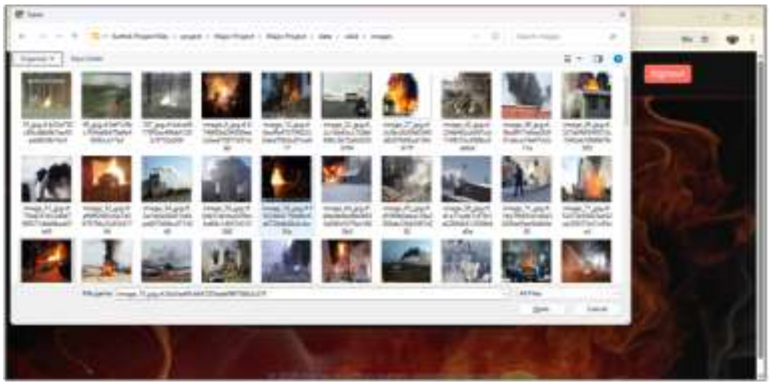

Fig.6 :Dataset Folder Interface

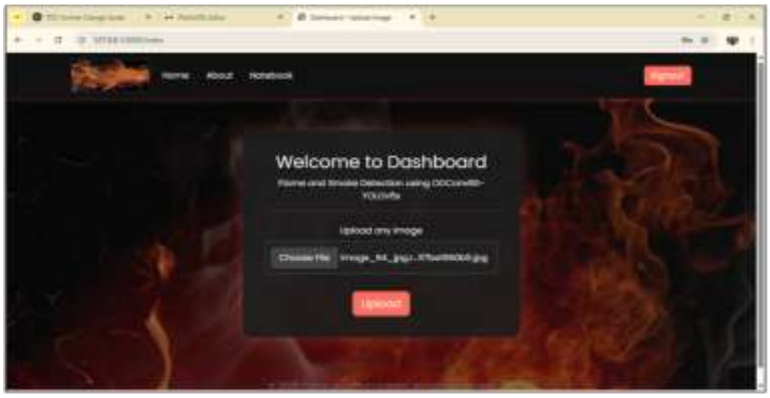

Fig. 7: Image Uploaded Confirmation

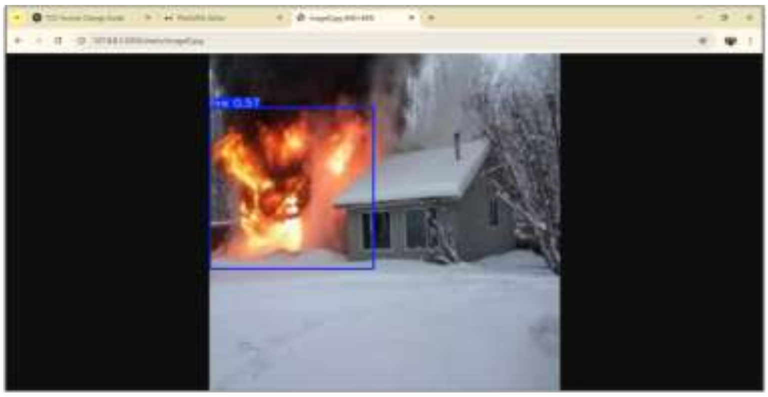

Fig. 8: Output Showing Detected Flame/Smoke

Fig. 9: Alteranate Output with Bounding Boxes

CONCLUSION

This study comprehensively evaluated several state-of-the-art object detection frameworks—such as SSD, Faster R-CNN, YOLOv5 (including its variants), and YOLOv8—for the purpose of real-time flame and smoke identification, particularly in aerial and satellite imagery. Among these, the enhanced YOLOv5s model embedded with ODConvBS emerged as the most effective, offering superior precision, recall, and mean Average Precision (mAP).

The introduction of architectural improvements, including the ODConvBS module for dynamic feature learning, Gnconv-FPN for refined spatial encoding, Shuffle Attention for cross-scale feature fusion, and the SIOU loss function for more accurate bounding box regression, collectively contributed to a substantial uplift in model performance. The modified ODConvBS-YOLOv5s achieved an mAP of 87.6%, outperforming its predecessors and other benchmark algorithms.

Additionally, the development of a web-based interface using Flask and SQLite provides a user-friendly platform for uploading and analyzing images, thereby expanding the model's usability in practical monitoring systems.

Looking forward, the proposed framework sets a foundation for future research on more efficient architectures. Incorporating lightweight backbone networks and novel attention mechanisms can further reduce computational complexity while sustaining high accuracy. These advancements aim to support real-time deployment in industrial safety, surveillance, and disaster management environments, marking a significant step toward intelligent fire detection systems.

REFERENCES

[1] Khan, S., Muhammad, K., & Hussain, T. (2021). *DeepSmoke: Deep learning model for smoke detection and segmentation in outdoor environments*. *Expert Systems with Applications*, **182**, 115125. https://doi.org/10.1016/j.eswa.2021.115125

[2] Anwar, M. W., & Bhatti, A. (2020). *Vision-based fire detection using machine learning and deep learning techniques: A review*. *Fire Technology*, **56**(5), 1883–1902. https://doi.org/10.1007/s10694-019-00946-1

[3] Redmon, J., & Farhadi, A. (2018). *YOLOv3: An incremental improvement*. arXiv preprint arXiv:1804.02767. https://arxiv.org/abs/1804.02767

[4] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). *YOLOv4: Optimal speed and accuracy of object detection*. arXiv preprint arXiv:2004.10934. https://arxiv.org/abs/2004.10934

[5] Jocher, G., Chaurasia, A., Qiu, J., & Stoken, A. (2021). *YOLOv5 by Ultralytics*. GitHub. https://github.com/ultralytics/yolov5

[6] Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollar, P. (2017). *Focal Loss for Dense Object Detection*. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2980–2988). https://doi.org/10.1109/ICCV.2017.324

[7] Ren, S., He, K., Girshick, R., & Sun, J. (2015). *Faster R-CNN: Towards real-time object detection with region proposal networks*. *Advances in Neural Information Processing Systems*, **28**. https://papers.nips.cc/paper_files/paper/2015/hash/14bfa6bb14875e45bba028a21ed38046-Abstract.html

[8] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). *SSD: Single shot multibox detector*. In *European Conference on Computer Vision* (pp. 21–37). Springer, Cham.

[9] Ge, Z., Liu, S., Wang, F., Li, Z., & Sun, J. (2021). *YOLOX: Exceeding YOLO series in 2021*. arXiv preprint arXiv:2107.08430. https://arxiv.org/abs/2107.08430

[10] Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2022). *YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors*. arXiv preprint arXiv:2207.02696. https://arxiv.org/abs/2207.02696

[11] Guo, M. H., Lu, C. Z., Liu, Z. N., Peng, J. X., & Cheng, M. M. (2022). *Attention mechanisms in computer vision: A survey*. *Computing Research Repository*, arXiv:2111.07624. https://arxiv.org/abs/2111.07624

[12] Su, J., Zhu, H., Li, Z., & Jiang, Y. (2023). *Omni-dimensional dynamic convolution for visual recognition*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. https://doi.org/10.1109/TPAMI.2023.3245303

[13] Liu, Z., Ning, J., Cao, Y., Wei, Y., Zhang, Z., Lin, S., & Hu, H. (2021). *Video Swin Transformer*. arXiv preprint arXiv:2106.13230. https://arxiv.org/abs/2106.13230

[14] Ge, Z., Liu, S., Wang, F., Li, Z., & Sun, J. (2021). *YOLOX: Exceeding YOLO series in 2021*. arXiv preprint arXiv:2107.08430. https://arxiv.org/abs/2107.08430

[15] Jocher, G., Chaurasia, A., & Qiu, J. (2023). *YOLOv8 by Ultralytics*. GitHub. https://github.com/ultralytics/ultralytics