EMITTER International Journal of Engineering Technology
Vol. 12, No. 2, December 2024, pp. 167~181
DOI : 10.24003/emitter.v12i2.882

Evaluating YOLOv5s and YOLOv8s for Kitchen Fire Detection: A Comparative Analysis

Norisza Dalila Ismail[1], Rizauddin Ramli[1], Mohd Nizam Ab Rahman[1]

[1]Department of Mechanical and Manufacturing Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, Malaysia
Corresponding Author: rizauddin@ukm.edu.my

Received September 10, 2024; Revised October 17, 2024; Accepted December 3, 2024

Abstract

Accurate and timely detection of kitchen fires is crucial for enhancing safety and reducing potential damage. This paper discusses comparative analysis of two cutting-edge object detection models, YOLOv5s and YOLOv8s, focusing on each performance in the critical application of kitchen fire detection. The performance of these models is evaluated using five main key metrics including precision, F1 score, recall, mean Average Precision across various thresholds (mAP50-95) and mean Average Precision at 50 percent threshold (mAP50). Results indicate that YOLOv8s significantly outperforms YOLOv5s in several metrics. YOLOv8s achieves a recall of 0.814 and an mAP50 of 0.897, compared to YOLOv5s' recall of 0.704 and mAP50 of 0.783. Additionally, YOLOv8s attains an F1 score of 0.861 and an mAP50-95 of 0.465, whereas YOLOv5s records an F1 score of 0.826 and mAP50-95 of 0.342. However, YOLOv5s shows a higher precision of 0.952 compared to YOLOv8s' 0.914. This detailed evaluation underscores YOLOv8s as a more effective model for precise fire detection in kitchen settings, highlighting its potential for enhancing real-time fire safety systems. Additionally, by offering the future work of integration of sensors with latest YOLO involvement can further optimize efficiency and fast detection rate.

Keywords: Convolutional neural network, Deep learning, Kitchen fire detection, Performance metrics, YOLO

1. INTRODUCTION

To protect materials and maintain safety, fire detection is essential in indoor environments especially kitchens. Cooking appliances [1][2], flammable materials [3] and heat sources [4] usage are particularly vulnerable to kitchen fire occurrences [5]. The National Fire Protection Association (NFPA) reported that one of the main causes of house fires and injuries in the United States is the cooking appliances [3]. Rapid and precise fire detection in kitchens can save fatalities, minimize property damage, and prevent injuries. However, since smoke and steam are frequent byproducts of cooking activites [6][7], conventional fire detection devices such as smoke detectors have difficulties in accurately detecting fires. Recent developments focusing on vision-based approaches that use Convulational Neural Networks (CNN) for real-time fire detection situations in order to overcome these difficulties. The You Only Look Once (YOLO) architecture has demostrated great result in object identification [8] due to its high accuracy and speed [9][10], making it ideal for real-time applications [11] [12].

This article aims to address this need by comparing the performance of two widely used architectures, the YOLO model using YOLOv5s and YOLOv8s, particularly in kitchen settings. The objective is to evaluate each model's feasibility for practical implementation in fire-prone indoor environment such as kitchens.

2. RELATED WORKS

In deep learning models, the YOLO based one-stage detection technique bypasses the region extraction step and instead solves the detection issue as a regression problem. This allows for the immediate extraction of the target's location and class information from the images, as opposed to the two-stage algorithm. The first YOLO model was introduced by Redmon et al. [13]; however, despite this, it has a very low recall value and precision of detection. In order to address the problems with the earlier version and enhance the speed and efficiency of the detection, the YOLOv2 [14] and YOLOv3 [15] models were introduced. Bochkovskiy [16] introduced YOLOv4, which has greatly enhanced the accuracy of the detection model. Later that year, Jocher [17] presented the YOLOv5 model, a more lightweight network with a minimum size of 2.62 MB built on the Pytorch framework. YOLOv5 surpasses YOLOv4 in both inference speed and detection accuracy, achieving rapid detection at 140 frames per second on Tesla P100. YOLOv6 [18] was then released in June 2022 to achieve the objective of creating an object detector used in the industry scale. In the following month of 2022, the YOLOv7 by Wang et al. [19] was released which suggested a number of architectural reforms to maintain high detection speed while enhancing accuracy. Next version of YOLOv9 [20] was released in early of year 2024 adding new approaches such as Programmable Gradient Information (PGI) and the Generalized Efficient Layer Aggregation Network (GELAN).

Improvised of YOLO architecture has been made in numerous of research fire imaging detection. Fire-YOLO [21] used an improvement of YOLOv3 with added network of hollow convolution and DenseNet to enhance early detection of small-scale flames. A lightweight CNN network model designed specifically for ship fire detection [22] is made from modified YOLOv4 algorithm despite the constrained computational resources availbility in maritime settings. Similarly to a new lightweight model ES-YOLO [23] produced from improvement of YOLOv5s with replacement of EfficientNetv2 network and SioU loss function able to reduce the computational complexity and improves the speed and detection acuracy. Few studies present a fire detection model utilizing traditional machine learning method rather than deep learning approaches such as [24] and [25]. Summary of related works shown in Table 1.

Table 1. Summary of Related Works

Author	Models used	Dataset	Performance Metrics	Findings
[21]	Fire-YOLO	Public websites	Precision: 91.50% Recall: 59.62% mAP: 80.23%	Enhances feature propagation of fire small targets identification, improve network performance and reduce model parameters
[22]	Modified YOLOv4-tiny algorithm	Homemade ship fire dataset	Precision: 0.928 Recall: 0.875 mAP: 0.906	Ship fire detection accuracy and detection efficiency
[23]	ES-YOLO	PASCAL VOC2007 dataset	mAP:20% improvement, 15% recall improvement of YOLOv5s	Lightweight and real-time fire detection with improved accuracy
[24]	Gaussian process classification	Open sources	Prediction above 90%	Achieve a high correct detection probability when the training images are either adequate or inadequate
[25]	Support Vector Machine (SVM)	Bilkent University, Signal Processing Lab	Accuracy: 93.33%	Precision and detect faster real-time fire detection.

3. ORIGINALITY

Despite the promising capabilities of object detection models for fire detection, there remains a deficiency of comparative research among different YOLO architectures specifically for kitchen fire detection. The majority of research has concentrated on assessing a single model separately. Previous comparison-based studies have tended to focus on a limited number of techniques or less sophisticated models. Moreover, a thorough evaluation of the performance indicators and detection capabilities unique to fires has been absent. This paper presents a comprehensive comparison research focused on the systematic evaluation of YOLOv5s and YOLOv8s for kitchen fire detection. The paper highlights the potential benefits and practical implications of new image processing techniques and Internet of Thing (IoT) integrations, providing the most comprehensive comparative review of contemporary YOLO algorithms for fire detection applications.

4. SYSTEM DESIGN

The YOLOv5s and YOLOv8s architectures were used in this research. Both models were trained using Google Colab, utilizing the NVIDIA Tesla T4 GPU with 16-GB of RAM and CUDA version 12.2. The Tesla T4's outstanding performance and economical power consumption is appropriate for deep learning task. Low GPU memory usage (3 MiB) during startup with nvidia-smi monitoring both GPU utilization and memory usage, suggesting a cost effective method for training. The efficiency of cloud-based resources for this application was demonstrated by running 50 epochs with a batch size of 16.

4.1 Data Source

The custom dataset of kitchen fire images used in this study being uploaded in a website Roboflow, a platform for managing and augmenting image datasets. The link to dataset can be viewed from https://universe.roboflow.com/fire-detection-t00kz/indoor-fire-toktd. The custom dataset comprises images specifically curated for kitchen indoor fire detection, ensuring relevance to the study's objective. A total of 137 images are utilized for training both YOLOv5 and YOLOv8 architectures, with an additional 39 images reserved for model validation. All image sizes were adjusted to 640x640 pixels to meet the input requirement during dataset preprocessing. The selection of dataset and partitioning strategy is to offer a varied and representative sample for assessing the performance of the object detection models.

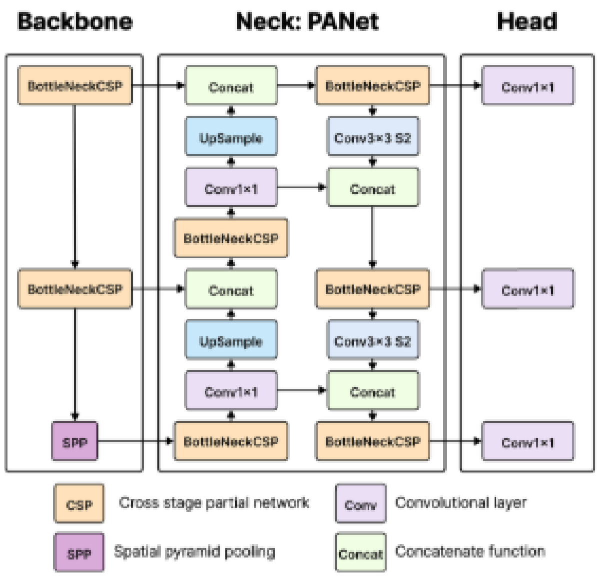


Figure 1. YOLOv5 architecture [22]

4.2 YOLOv5

One object detection algorithm technique that is popular for its reliability, simplicity and accuracy [8] is called You Only Look Once version 5, or YOLOv5. The YOLOv5 [18] was released by Ultralytics in mid of the year 2020 and consists of three foundations: backbone, neck and head, as shown in Figure 1.

The backbone depends on the CSP-Darknet53 convolutional network and employs the Cross Stage Partial (CSP) strategy to expedite information flow while alleviating issues linked to redundant and vanishing gradients [24][25]. A version of the Spatial Pyramid Pooling (SPP) is incorporated into the neck of the YOLOv5 model, and the Bottle-NeckCSP is integrated into the Path Aggregation Network (PANet) [26]. Combination of these techniques improves the receptive field in order to retain the network speed by separating important context features. The CSPNet strategy [25] improved the PANet in YOLOv4 as a feature pyramid network, to provide a better pixel localization accuracy in YOLOv5. The neck of the architecture is crucial in order to manage object scaling and allow the model to perform remarkably well on unobserved data.

The head of YOLOv5 is made up of three convolution layers similar to its predecessors [25]. These layers, which differ slightly from the preceding versions in the computation of target coordinates for bounding boxes, predict coordinates of bounding box, score and object classification [25] [27].

According to Johnston et al. [8], there are four different levels of neural network models in YOLOv5, from simplest to most complex including YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x. With the help of an effective inference strategy and a lightweight network architecture, YOLOv5s built on a single-stage detector that enables quick and precise object recognition [28]. Based on the demand for quick, small model and high precision, YOLOv5s is selected as the model to be used in this paper.

4.3 YOLOv8

In January 2023, Ultralytics introduced YOLOv8 [29] which represents a remarkable breakthrough in the object detection area, image recognition as well as instance segmentation. To accomplish these features, YOLOv8 essentially upgrades YOLOv5 [30]. Similar to YOLOv5, YOLOv8 is made up of three primary architectural components: the head, neck and backbone. Figure 2 illustrates the structure of the YOLOv8: the head performs object identification and classification prediction, the neck combines image frames featured by the backbones. The adoption of an anchor-free model by YOLOv8, which differs from the anchor-box technique used in previous YOLO models, is one of its most notable aspects [31]. This change enables the model to accurately predict an object's centre without the hustle of anchor boxes such as inability to handle irregularities and lack of generalization. Reducing the quantity of box predictions enables the YOLOv8 model to expedite the speed of the Non-Maximum Supression (NMS) process, the critical post-processing step in charge of sifting through candidate detections following interference [30] [32].

YOLOv8s has a more complex architecture with 168 layers and 11,126,425 parameters, while YOLOv5s has 182 layers and 7,251,912 parameters. High parameters in YOLOv8s indicate that it may require more computational resources during training. This also include more memory (RAM) to load and execute which require powerful processors [30].

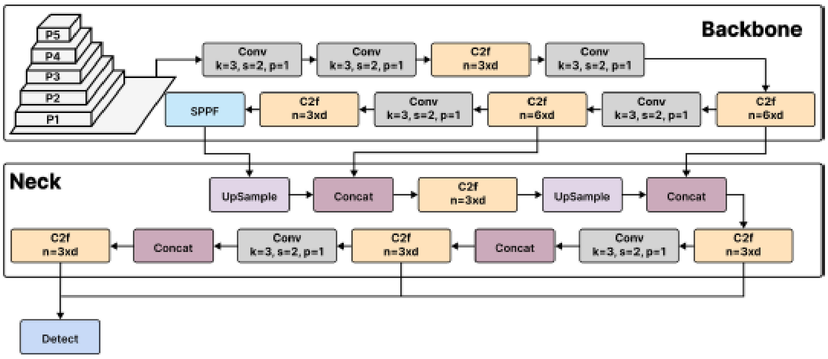


Figure 2. YOLOv8 architecture [31]

4.4 Evaluation Metrics

The confusion matrix depicted in Figure 3 is employed for evaluating the performance of a model. This matrix, described by Zeng [33], displays the relationship between actual and predicted classifications. The confusion matrix comprises four classifications such as True Negative (TN), True Positive (TP), False Negative (FN), and False Positive (FP) based on actual and predicted values. The meaning of each classification such as TP represents the total number correctly identified as positive samples; TN represents the total number of correctly classified as negative samples; FP denotes the number of incorrectly classified positive samples while FN represents incorrectly predicts the negative class [34].

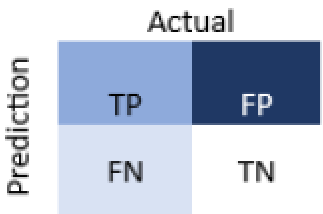


Figure 3. Confusion matrix

4.4.1 Precision

Fundamental evaluation metric in object detection is precision. The accuracy of the model's positive prediction was measured using precision [35][36][37]. Equation 1 represents the precision level, where TP is the total of correctly predicted positives and FP is the number of false positives. Less false positives indicate the model is high precision or high accuracy.

$$Precision = \frac{TP}{(TP + FP)} \tag{1}$$

4.4.2 Recall

Recall was used to measure the ability of selected models for accuracy detection for all instances of fire [38]. A higher recall value means the model is less likely to miss any fires. It is also defined as the ratio of TP to the total number of actually positive instances [36][37]. Recall is determined as in Equation (2) which represents the number of correctly predicted positive instances, where FN refers to negative instances.

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

4.4.3 F1 Score

The F1-score is a combination of the value of precision and recall extending an analysis of the model's accuracy [38]. The F1-score shown in Equation (3)[39] serves as the ratio of both product of precision and recall with multiplication of 2. The best value of F1-score is 1.0 while the worst is 0.0.

$$F1\,Score = \frac{2\,X\,Precision\,X\,Recall}{Precision + Recall} \tag{3}$$

4.4.4 Mean Average Precision

The Mean Average Precision (mAP) measures the balance between precision and recall values. To achieve this, firstly the Average Precision (AP) is computed for each class and finalized with averaging across all classes [35]. By calculating the area under the precision-recall curve, AP measures precision at different levels of recall values [40]. Equation (4) shows precision at specific recall level. Considering both precision and recall, a higher mAP denotes better object detection performance [35][36].

$$AP = \int_1^0 Precision\,(Recall)d\,(Recall) = a \int precision\,(r)dr \tag{4}$$

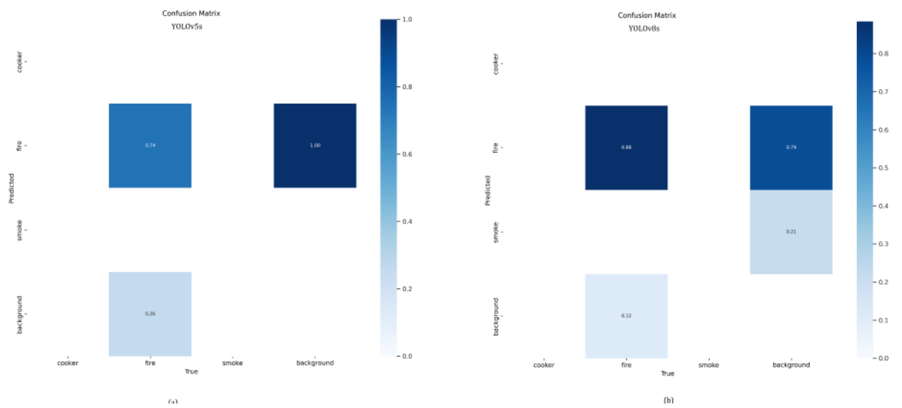
5. EXPERIMENT AND ANALYSIS

5.1 Inference Speed and Computational Performance

Both model runs at different inference speed. YOLOv8s processes images at 4.8 ms per image while YOLOv5s processed slightly slower with 5.5 ms per image. The improvement shown in YOLOv8s model despite its complex architecture demonstrates faster inference speed resulting in optimized layers and operations. However, the YOLOv8s with high processing power requirements will required a powerful and expensive hardware setup. Despite being a little more computationally intensive, YOLOv8s can provide real-time applications a significant benefit due to its higher inference speed.

5.2 Performance evaluation

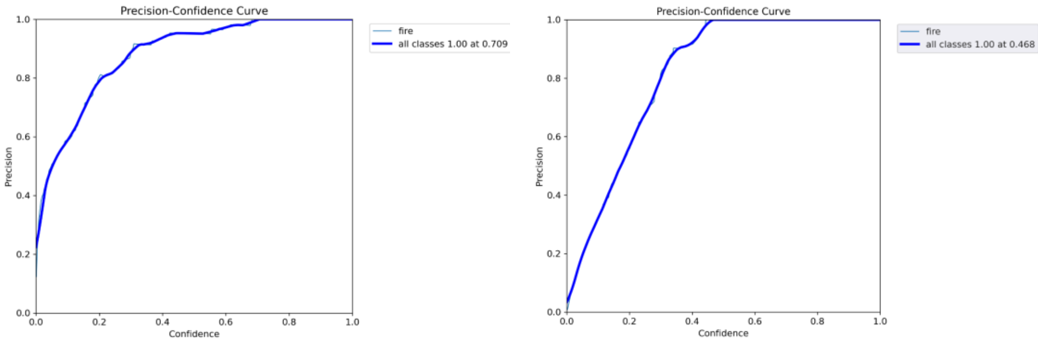
Further evaluation was made for both types of YOLO test. The identification of results in terms of the TP, TN, FP and FN for both fire and smoke was displayed in the confusion matrix in Figure 4. Based on our analysis, there is a slightly large variance that performed from each model. For instance, in YOLOv5s and YOLOv8s, the percentage of true positive value is 0.74 and 0.88 respectively. Therefore, an average true positive value for fire was 0.81. For smoke in YOLOv5s, the positive value is 0 while for YOLOv8s, true positive rate of 0.71 with 0.29 being incorrectly classified as background. The result is lower for smoke classification because both models unable to detect correctly between smoke and background images. Despite this, the results indicate that additional development is necessary and provide satisfactory performance outcomes first several model approaches towards both database and models to raise high real-time true positive values for application in kitchen and indoor environment.



(a) (b)

Figure 4. Confusion matrix of (a) YOLOv5s and (b) YOLOv8s based on testing images

The precision-confidence curves for both YOLOv5s and YOLOv8s in Figure 5 provide critical insights into their performance in detecting kitchen fires. YOLOv8s displays a smooth, steadily rising indicates a consistent improvement in precision as confidence increases. This model achieved perfect precision (1.00) at a confidence level value of 0.709 denoting it can accurately classify fire instances with lower false positives at lower confidence threshold. Meanwhile, the model YOLOv5s precision reaches 1.00 at a lower confidence threshold of 0.468 suggesting that the model at lower confidence level enable to achieve high precision but may faces challenges with precision consistency as confidence increases. The curve initially shows a steeper rise but it levels off, determine that YOLOv5 may generate more false positives at moderate confidence levels.



(a) (b)

Figure 5. Precision-confidence curves of (a) YOLOv8s and (b) YOLOv5s

Results in Figure 6 shows that YOLOv5s has significantly higher precision value of 0.952 compared to YOLOv8s with value of 0.914, indicating that YOLOv5s is more accurate in correctly identifying the objects it detects. mAP50 values for both models show a significant difference indicating a substantial improvement in detection accuracy of YOLOv8s over YOLOv5s. Meanwhile for mean Average precision (mAP50-95) to indicate how well the models are performing across different thresholds, YOLOv5s has a lower mAP50-95 of 0.342 compared to YOLOv8s. The higher mAP50-95 score indicates better overall performance compared to YOLOv5s, suggesting improvements in model architecture or training.

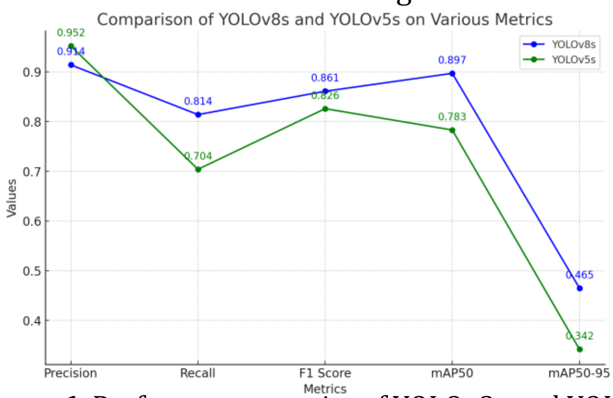
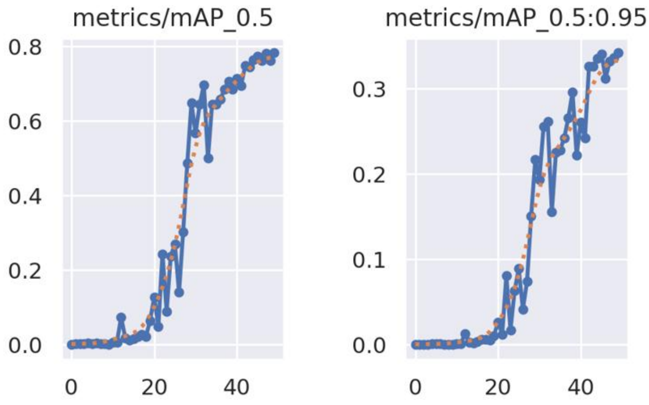


Figure 6. Performance matrics of YOLOv8s and YOLOv5s

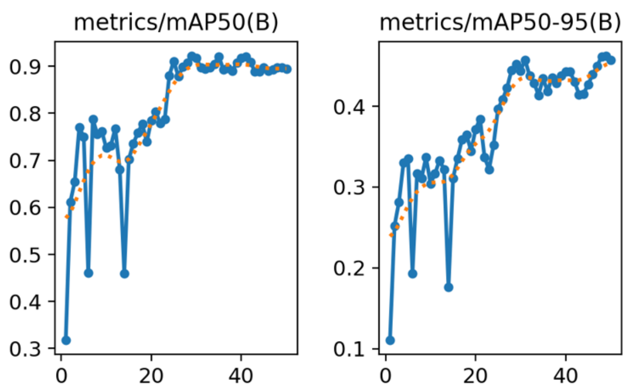
To explain further about mAP performance based on Figure 7 for YOLOv5s model using two different metrics of mAP50 in Figure 7(a) and mAP50-95 in Figure 7(b). The mAP50 metric measures precision when the overlap between predicted and actual boxes is at 50 percent. It starts at 0.1 and increases to 0.78 over 50 training epochs. This means better detection as training progresses. Meanwhile the mAP50-95 metric measures precision over a range of overlaps from 50 to 95 percents, Figure 7(b) shows precision starts near zero and increases to around 0.34. There is an improvement but at a lower rate due to limitation of evaluation criteria.



(a) **(b)**

Figure 7. YOLOv5s graphs: (a) mAP50 (b) mAP50-95

Figure 8 illustrates the performance of YOLOv8s model using the same metrics. The mAP50 in Figure 8(a) starts at 0.3 and increases sharply to the value of 0.80 within 10 epochs then gradually reaches 0.9 by 50 epochs. These results indicate fast and sustained improvement in detection accuracy. The mAP50-95 in Figure 8(b) starts at above 0 and increases to about 0.45 by 50 epochs. There is a significant improvement across different overlap levels surpassing YOLOv5s model.



(a) **(b)**

Figure 8. YOLOv8s graphs: (a) mAP50 (b) mAP50-95

Recall value in Figure 6 for YOLOv5s model is 0.704 to indicate that this model is able to detect about 70.40% of all actual objects in the images while for YOLOv8s is much better at detecting objects, finding 81.40% of all actual objects. Figure 9 and Figure 10 show the results of applying the kitchen fire images training detection using YOLOv5s and YOLOv8s models.

Figure 9. Example snapshot of detection using YOLOv8s

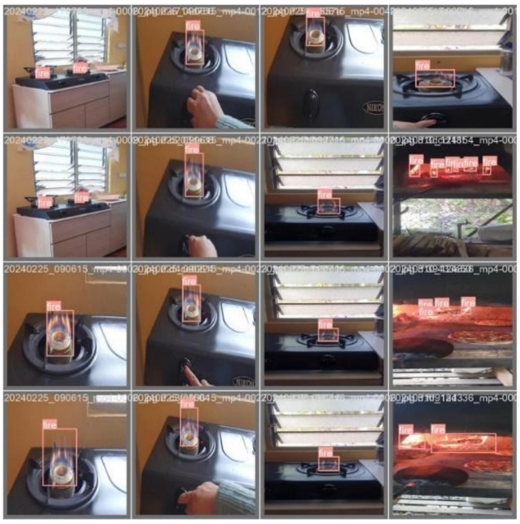


Figure 10. Example of snapshot of detection using YOLOv5s

6. CONCLUSION

This study carried out a comparative analysis of YOLOv5 and YOLOv8s for fire detection in kitchen environments to evaluate their performance. The results displayed according to the five important performance metrics;

mAP50, mAP50-95, F1 Score, Precision and Recall. The result shows that model YOLOv5s achieves high precision than YOLOv8s, proving fewer false positives in detection. Both models have successfully detected fire and smoke in the kitchen environment. YOLOv8s integrates several architectural improvements than YOLOv5s, which include more effective feature extraction layers, improved backbone networks, and improved object detection head designs. These innovations increase the accuracy of object localization and classification. Based on the findings of this comparative study between YOLOv5s and YOLOv8s, our future research is to further enhance the performance and applicability of object detection models such as focus on refining the algorithms used for kitchen fire detection with challenging environment with varying lighting and reflection for confusing detection. Besides, exploring the integration of additional sensors to increase the accuracy and fast detection tasks. Combining sensor data with visual data from cameras can provide a richer set of information in making more reliable detections and decrease false alarms. Future work could explore the underlying architectural differences contributing to these performance gains and assess the models' effectiveness across diverse datasets and real-world applications.

Acknowledgments

The authors would like to thank the Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia for supporting this project.

REFERENCES

[1] S. Hall, and T. Mcgree, **Home cooking fires: supporting tables** September 2023 (NFPA ®)

[2] M. Ahrens, **Home fires involving cooking equipment.** National Fire Protection Association Fire Analysis and Research Division, 2009

[3] A. Dinesh, T. Polanco, and R. Engdahl, **Burns from ignited household aerosols in the kitchen: a case series.** *Scars Burn Heal*, 3, 205951311772820, 2017

[4] M. Spearpoint, C. Hopkin, and D. Hopkin, **Modelling the thermal radiation from kitchen hob fires.** *Journal Fire Science*, 38 (4), 377–394, 2020

[5] M.B. Hamida, and M.A. Hassanain, **Fire safety in the built-environment: a case study in a residential facility,** *Architecture, Civil Engineering, Environment,* 12 (2), 2019

[6] H.B. Choi, E.H. Hwang, and D.M. Choi, **Indoor air quality sensor utilization for unwanted fire alarm improvement in studio-type apartments**. *Fire*, 6 (7), 2023

[7] J. Milke, and R. Zevotek, **Analysis of the response of smoke detectors to smoldering fires and nuisance sources.** *Fire Technology*, 52 (5), 1235–1253, 2016

[8] J. Johnston, K. Zeng, and N. Wu, **An evaluation and embedded hardware implementation of yolo for real-time wildfire detection.** *2022 IEEE World AI IoT Congress,* AIIoT 2022, 138–144, 2022

[9] L. An, L. Chen, and X. Hao, **Indoor fire detection algorithm based on second-order exponential smoothing and information fusion.** *Information*, 14 (5), 258. 2023

[10] E. Casas, L. Ramos, E. Bendek, and F. Rivas-Echeverria**, YOLOv5 vs. YOLOv8: Performance benchmarking in wildfire and smoke detection scenarios.** *Journal of Image and Graphics,* 12 (2), 127–136, 2024

[11] A. Rehman, D. Kim, and A. Paul, **Convolutional neural network model for fire detection in real-time environment.** *Computers, Materials & Continua,* 77 (2), 2289–2307, 2023

[12] Y. Li, J. Shang, M. Yan, B. Ding, and J. Zhong, **Real-time early indoor fire detection and localization on embedded platforms with fully convolutional one-stage object detection.** *Sustainability*, 15 (3), 1794, 2023

[13] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, **You Only Look Once: Unified, real-time object detection.** *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788, 2016

[14] J. Redmon, and A. Farhadi, **YOLO9000: Better, faster, stronger.** *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 6517–6525, 2017

[15] J. Redmon, and A. Farhadi, **YOLOv3: An incremental improvement**, 2018

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

[17] G. Jocher, **YOLOv5 by Ultralytics**, 2020

[18] C. Li, L. Li, H. Jiang, K. Weng, Y. Geng, L. Li, Z. Ke, Q. Li, M. Cheng, W.; Li, Y. Nie, B. Zhang, Y. Liang, L. Zhou, X. Xu, X. Chu, X. Wei, and X. Wei, **YOLOv6: A single-stage object detection framework for industrial applications**. arXiv preprint arXiv:2209.02976, 2022

[19] C.-Y. Wang, A. Bochkovskiy, and H.-Y.M. Liao, **YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors**. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 7464-7475),* 2022

[20] C.-Y. Wang, I.-H. Yeh, and H.-Y.M. Liao, **YOLOv9: Learning what you want to learn using programmable gradient information.** arXiv preprint arXiv:2402.13616, 2024

[21] L. Zhao, L. Zhi, C. Zhao, W. Zheng, **Fire-YOLO: A Small Target Object Detection Method for Fire Inspection**. *Sustainability*, 14, 4930, 2022

[22] H. Wu, Y. Hu, X. Mei, and J. Xian, **Ship fire detection based on an improved yolo algorithm with a lightweight convolutional neural network model.** *Sensors*, 22(19):7420. 2022

[23] S. Wang, and X. Wang, **ES-YOLO: A new lightweight fire detection model.** *Third International Conference on Computer Vision and Data Mining (ICCVDM 2022)*, 125111F, 2023

[24] X. Wan, J. Cai, S. Luo, Z. Tian, L. Zhang and X. Xia, **Gaussian Process for the Machine Learning-based Smart fire Detection System**. *2022 IEEE 6th Information Technology and Mechatronics Engineering Conference (ITOEC)*, Chongqing, China, 100-104, 2022

[25] M. A. Rahman, S. T. Hasan and M. A. Kader, **Computer Vision Based Industrial and Forest Fire Detection Using Support Vector Machine (SVM).** *International Conference on Innovations in Science, Engineering and Technology (ICISET)*, Chittagong, Bangladesh, 233-238, 2022

[26] J. Zhang, J. Zhang, K. Zhou, Y. Zhang, H. Chen, and X. Yan, **An improved yolov5-based underwater object-detection framework.** *Sensors*, 23 (7), 3693, 2023

[27] M.L. Mekhalfi, C. Nicolo, Y. Bazi, M.M. Al. Rahhal, N.A. Alsharif, and E. Al. Maghayreh, **Contrasting Yolov5, transformer, and efficientdet detectors for crop circle detection in desert**. *IEEE Geoscience and Remote Sensing Letters*, 19, 2022

[28] Z. Liu, X. Gao, Y. Wan, J. Wang, and H. Lyu, **An improved Yolov5 method for small object detection in UAV capture scenes.** *IEEE Access,* 11, 14365–14374, 2023

[29] W. Liu, K. Quijano, and M.M. Crawford, **YOLOv5-Tassel: Detecting tassels in RGB UAV imagery with improved yolov5 based on transfer learning.** *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 15, 8085–8094, 2022

[30] S. Wu, Z. Li, S. Li, Q. Liu, and W. Wu, **Static gesture recognition algorithm based on improved yolov5s.** *Electronics (Basel),* 12 (3), 596, 2023

[31] G. Jocher, A. Chaurasia, and J. Qiu, **YOLO by Ultralytics**, 2023

[32] J.H. Kim, N. Kim, and C.S. Won, **High-speed drone detection based on yolo-v8.** *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2023-June, 2023

[33] A. Vats, and D.C. Anastasiu, **Enhancing retail checkout through video inpainting, Yolov8 detection, and deepsort tracking.** *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW),* 5530–5537, 2023

[34] M. Hussain, H. Al-Aqrabi, M. Munawar, R. Hill, and T. Alsboui, (). **Domain feature mapping with yolov7 for automated edge-based pallet racking inspections.** *Sensors*, 22 (18), 6927, 2022

[35] G. Zeng, **On the confusion matrix in credit scoring and its analytical properties**. *Communications in Statistics- Theory and Methods*, 49 (9), 2080–2093, 2020

[36] Y. Zhang, T. Zuo, L. Fang, J. Li, and Z. Xing, **An improved MAHAKIL oversampling method for imbalanced dataset classification.** *IEEE Access*, 9, 16030–16040, 2021

[37] E. Casas, L. Ramos, E. Bendek, and F. Rivas-Echeverria, **Assessing the effectiveness of yolo architectures for smoke and wildfire detection.** *IEEE Access*, 11, 96554–96583, 2023

[38] M. Vergara, L. Ramos, N.D. Rivera-Campoverde, and F. Rivas-Echeverria, **EngineFaultDB: A novel dataset for automotive engine fault classification and baseline results.** *IEEE Access*, 11, 126155–126171, 2023

[39] R. Padilla, S.L. Netto, and E.A.B. da Silva, **A survey on performance metrics for object-detection algorithms.** *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 237–242, 2020

[40] R. Padilla, W.L. Passos, T.L.B. Dias, S.L. Netto, and E.A.B. da Silva, **A comparative analysis of object detection metrics with a companion open-source toolkit.** *Electronics (Basel)*, 10 (3), 279, 2021

[41] Z. Ning, X. Wu, J. Yang, and Y. Yang, **MT-YOLOv5: Mobile terminal table detection model based on YOLOv5.** *Journal of Physics: Conference Series*, 1978 (1), 012010, 2021

[42] H. Zhu, H. Wei, B. Li, X. Yuan, and N. Kehtarnavaz, **A review of video object detection: datasets, metrics and methods.** *Applied Sciences*, 10 (21), 7834, 2020