International Journal of Computers and Informatics

Journal Homepage: https://www.ijci.zu.edu.eg

Int. j. Comp. Info. Vol. 11 (2026) 101–110

Paper Type: Original Article

A Brief Comparative Study of YOLOv7, YOLOv8, and YOLOv9 for Threat Object Detection

Toqa Yousri [1,*], Mustafa Abdul Salam [2,3], Mahmoud Ismail [1], and Ahmed R. Abas [1]

[1] Faculty of Computers and Informatics, Zagazig University, Zagazig, Egypt;
Emails: toqayousri890@gmail.com; mmsabe@zu.edu.eg; arabas@zu.edu.eg.

[2] Faculty of Computers and Artificial Intelligence, Benha University, Benha, Egypt; mustafa.abdo@fci.bu.edu.eg.

[3] Department of Computer Engineering and Information, College of Engineering in Wadi Addwasir, Prince Sattam bin Abdulaziz University, Saudi Arabia; m.aboelnour@psau.edu.sa.

Received: 26 Feb 2026 **Revised**: 15 Apr 2026 **Accepted**: 29 Jun 2026 **Published**: 30 Jun 2026

Abstract

Threat object detection has become as essential task in X-ray security screening systems due to the rapid growth in population and possible dangerous threats in public spaces. Consequently, advanced Machine Learning (ML) and Deep Learning (DL) models have been utilized for object detection to mitigate the limitations found in traditional detection models such as slow training time, complex backgrounds, multi-scale prohibited objects, random stacking and occlusion scenarios. This paper introduces a comprehensive comparative study of three prominent detection models YOLOv7, YOLOv8, and YOLOv9 which achieved remarkable insights in real-time detection with notable performance and robustness. Experimental results conducted on the GDXray dataset show the outstanding performance of the three models, surpassing traditional models by achieving higher performance and deduction of false positives and false negatives rates.

Keywords: Object Detection, X-ray, YOLOv7, YOLOv8, YOLOv9.

1 | Introduction

To guarantee the humans safety in public spaces such as airports, subways, and different areas, a vital inspection of their baggage is performed by X-ray scanners to detect possible threats and safety regulations caused by different threat items such as explosives, guns, knives, and radioactive items [1].

Currently, the detection of contraband in X-ray images primarily relies on manual methods, resulting in the accuracy of security inspections being significantly influenced by the unique circumstances and expertise of the security personnel as the inadequate and false inspection of dangerous objects frequently appears, leading to significant security risks [2].

To overcome the aforementioned security issues, advanced deep-learning-based models' detectors have been implemented in safety-critical contexts, including X-ray security inspections at public terminals aid inspectors in detecting both the existence and position of contraband items, thus minimizes human labor and protects the public from serious threats [3].

Over the past decades, numerous real-time detection models have been developed to address a variety of detection limitations such as overlapping and occlusion [4]. This study presents a comprehensive literature

Corresponding Author: toqayousri890@gmail.com

review on real-time detection models and their application to different public-spaces terminals to guarantee safety for humans. Multiple articles were published by various researchers focusing on the development of real-time object detection algorithms and their related limitations. Consequently, lots of studies have been identified on object detection in which advanced methods and related datasets have been considered. In [4], authors reviewed the security images datasets and label assignments which were developed over the recent years in deep learning detection models to enhance their generalization to detect hazardous and multi-scale objects found in complex backgrounds. The subject of X-ray image-based analysis has undergone a revolution because to recent advancements in computer vision techniques, which greatly simplify the processing and interpretation of X-ray images in real-time systems [5]. The literature [6], introduced a comprehensive study of You Only Look Once (YOLO) models, which primarily examines the remarkable development during ten years in YOLO models from YOLOv1 to YOLOv10, generalizability, detection mechanisms, label assignment, and their applications. Additionally, the literature [7] introduced in-depth analysis of YOLO models and how they achieved notable performance in object detection industry by balancing both speed and accuracy. Regarding real-time detection of various dangerous objects and prominent performance, with a focus on developed computer vision (CV) datasets, YOLO models offered a comprehensive assessment of real-time detection in different environments [8]. Moreover, authors in [9] mentioned the usage of YOLO models in video surveillance systems as a part of security screening tasks. The literature [10] reviews diverse deep learning detection methodologies in addition to YOLO series, identifies possible future solutions to enhance the overall performance associated with real-time detection by determining multi-scale objects, contextual modelling, spatial correlations, models' complexity, and dealing with different scenes from images and videos.

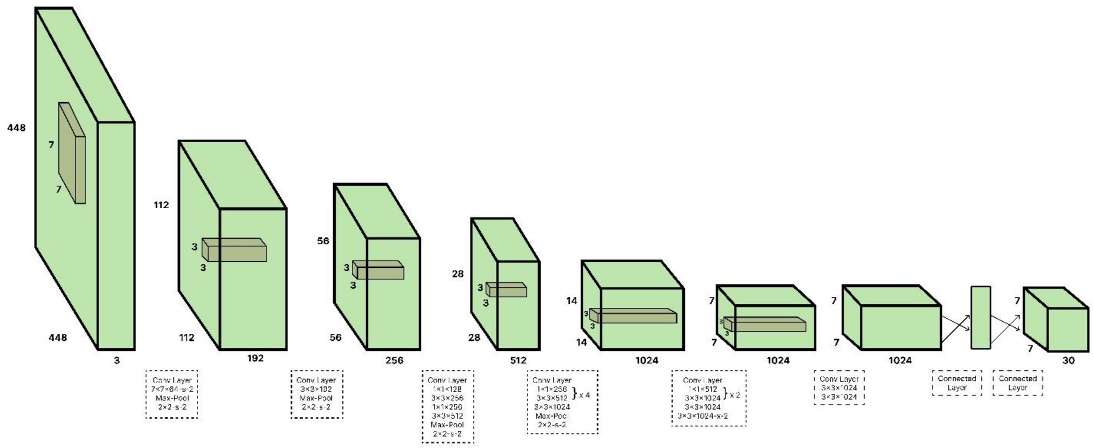

Figure 1. YOLO architecture [6].

The main benefit of this study is described as:

- The object detection task is defined and recent methods for precise recognition and localization are detailed.
- Reviews a group of YOLO models and their architectures.
- Represents a complete bibliography of related works specialized in threat object detection problem and X-ray datasets.
- Draws attention to important aspects of powerful YOLO models for real-time detection, including the description of these models, their mechanisms, and the evaluation metrics.
- Describes limitations noted in these detection models to improve detection performance.

Here is the structure of the paper: Object detection and real-time models' structures and applications are defined in Section II. In Section III, we provide the considerable literature on object detection by means of

YOLOv7, YOLOv8, and YOLOv9 models. Section IV presents the experimental results conducted on GDXray dataset in terms of mAP@50, Precision (P), Recall (R), F1-Score, and Accuracy (ACC). Section V reviews the limitations and future work. Finally, Section VI concludes the overall study.

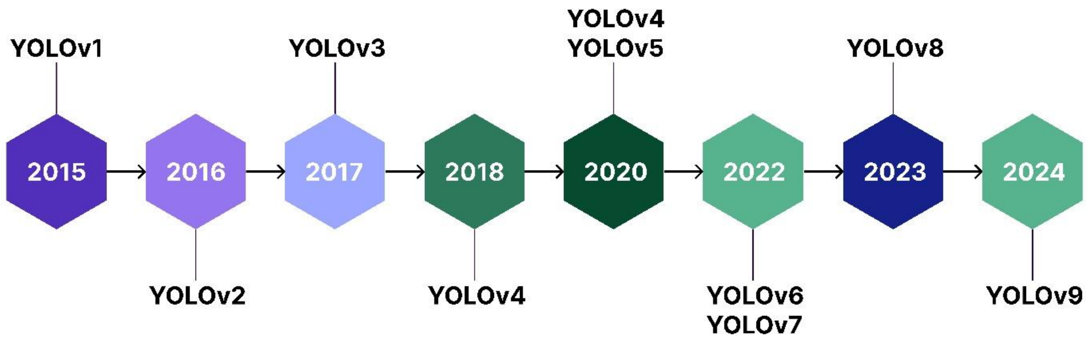

Figure 2. YOLO series timeline during last ten years.

2 | Background

In this section, the models' definition as well as the structures are used to describe the real-time object detection task in detail and their applications.

2.1 | YOLO Models

You Only Look Once (YOLO) was proposed in 2015 by Joseph Redmon et al. [11], it is regarded as one of the earliest techniques to present a real-time, one-stage approach to object detection. The model receives an input image and divides it into $S \times S$. Every grid cell is in-charge of identifying the items whose centers are located inside their own area. These predictions are given by Convolution Neural Network (CNN) architecture which consists of 24 layers followed by 2 fully connected layers which are responsible for feature extraction step from the input image. The fully connected layers receive these feature maps and use them to calculate the bounding box coordinates and class probabilities [12]. Figure 1 represents YOLO architecture. Despite its effectiveness, YOLOv1's structural design has some drawbacks. It is difficult to detect little objects that are very close to one another since each grid cell must be assigned a maximum of one class and two bounding box predictions as the model frequently overlooks small items cluttered natures or flocks of birds in complex backgrounds [13].

Furthermore, prominent development was performed to mitigate the limitations found in YOLOv1 to meet real-time detection requirements. The literature [14] mentioned diverse versions developed into YOLO series such as YOLOv2 which improved anchor boxes using K-means clustering and included the DarkNet-19 backbone, YOLOv3 extended this further by incorporating residual connections and multi-scale detection with a DarkNet-53 architecture, YOLOv4 uses PANet and CSPDarkNet-53 in addition to mosaic data augmentation, YOLOv5 and YOLOv6 employ CSPNet with dynamic anchor refinement and more improvements in PANet were introduced in, respectively, YOLOv7 had an EfficientRep backbone with dynamic label assignment while YOLOv8 introduced a Path Aggregation Network with dynamic label assignment, and YOLOv9 which relies on multi-level auxiliary for feature extraction. Figure 2 shows the timeline of YOLO series developments during the last ten years.

Due to the remarkable performance of different YOLO version, they are being utilized in diverse areas of real-time detection systems and application such as autonomous vehicles and traffic safety, where the technology improves safety and navigation through obstacle identification, pedestrian pose estimate for intention prediction, and traffic sign recognition [15], healthcare for identifying abnormalities in medical photos, supporting precise and effective diagnosis [16], monitoring to find anomalies and detect intrusions [17], and in agriculture, among other applications, it facilitates robotic fruit harvesting, crop stress detection, and monitoring [18]. When YOLOv7, YOLOv8, and YOLOv9 are compared, accuracy and efficiency show a consistent improvement. Figure 3 shows key differentiations between these models.

2.1.1 | YOLOv7

In 2022, Wang et al. [19] proposed YOLOv7, a real-time detection model with improved features to address the limitations of other traditional models that surpassed these models on pre-trained datasets regarding speed and accuracy. Extended efficient layer aggregation networks (E-ELAN) are a feature of YOLOv7. Expand, shuffle, and merge cardinality are used by E-ELAN to continuously improve the network's learning capacity without erasing the initial gradient route.

2.1.2 | YOLOv8

In 2023, Ultralytics developed YOLOv8 [20], incorporating state-of-the-art developments in neural network architecture and training techniques. YOLOv8 integrates object localization and classification tasks into a singular, end-to-end differentiable neural network framework, preserving the equilibrium between speed and precision by using three key components utilizing CNN for backbone for multi-scale feature extraction, Path Aggregation Network (PANet) for neck, and the head module that generates the final predictions, encompassing bounding box coordinates, object confidence scores, and class labels.

2.1.3 | YOLOv9

In 2024, Wang et al. [21] proposed YOLOv9, a real-time detection model which was developed through the integration of Programmable-Gradient Information (PGI) to leverage information bottleneck issue in deep learning models and Generalized Efficient Layer Aggregation Network (GELAN) as a lightweight backbone, it has demonstrated significant competitiveness. YOLOv9's superior architecture enables the deep model to decrease the parameter count by 49% and the computational load by 43% in comparison to YOLOv8.

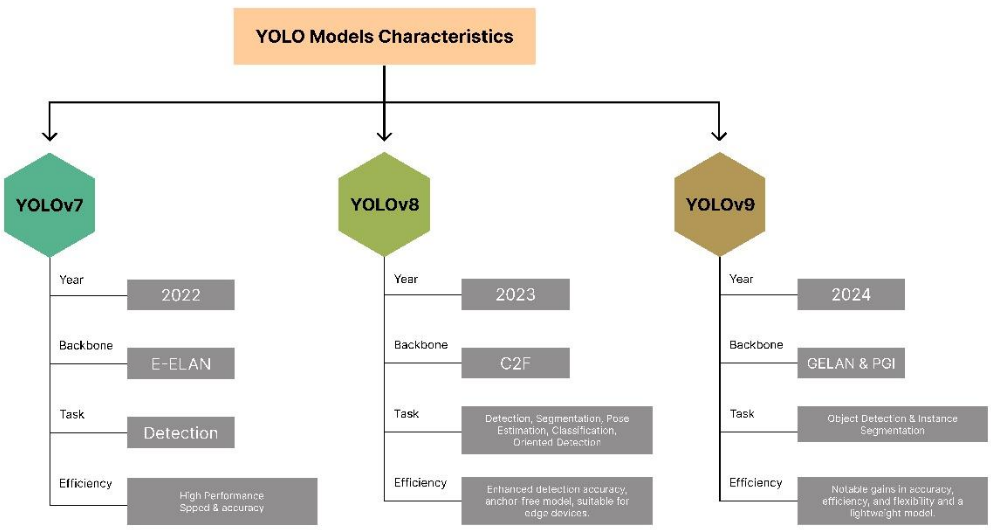

Figure 3. YOLOv7, YOLOv8, and YOLOv9 comparison.

3 | Literature Review

This section introduces the creation of a threat object detection models in X-ray imaging which is increasingly being employed to keep public space security screening and surveillance systems operational. This section introduces contemporary deep learning approaches and emphasizes the need of real-time detectors in this field such as YOLOv7, YOLOv8, and YOLOv9.

3.1 | YOLOv7 in Object Detection

Deena et al. [22] introduced a system for automatically detecting banned objects in X-ray images, with the goal of identifying possible hazards including firearms, explosives, and other prohibited items in real time using YOLOv7 network. The experimental results demonstrated the model's prominent accuracy and speed, therefore stating its effectiveness for real-world security applications.

Gan et al. [23] proposed an improved version of YOLOv7, named YOLO-CID to decrease resource waste, optimize the network's convolutional layer structure, improve the model's capacity to extract important information from complicated background images, and speed up feature extraction. Additionally, they employed a bidirectional weighted feature pyramid network to decrease the computational complexity. The experimental results have shown higher performance and speed compared with the baseline YOLOv7.

Jing et al. [24] introduced EM-YOLO, an improved version of YOLOv7 which consists of a material feature extractor (MFE) and an edge feature extractor (EFE) to solve the problems of overlapping and background clutter in X-ray image detection. The proposed model performs feature extraction based on features displayed in X-ray images. The model also uses CBAM attention mechanism for better feature extraction. EM-YOLO surpassed multiple detection models in terms of accuracy and can be used to help with manual inspection in real world scenarios.

Zhang et al. [25] proposed a novel detection model based on YOLOv7 called MFPIDet to detect forbidden items in complicated contexts with reliability using a backbone module for multi-scale attention to promote the representation of local features for overlapped items. Additionally, authors employed effective key components such as Multi-scale Feature Extraction module (MFEM), Squeeze-Excitation (SE) block, and AF-FPN that stands for adaptive fusion feature pyramid network. Consequently, they helped to enhance extracting objects' features in complex backgrounds in addition to contextual information. Experimental results show a higher detection performance compared to the original YOLOv7 model.

3.2 | YOLOv8 in Object Detection

Chen et al. [26] introduced YOLO-SRW, a real-time detection model based on YOLOv8 to boost the performance of identifying dangerous items in complex scenarios such as noisy backgrounds and variant objects scales. Using RFLSKA and SRFD modules have significantly improved comprehension and reaction to small objects and the capacity of the model to extract features from targets with poor resolution and sparse features. Experiments provide improved identification accuracy for small and multi-scale items in X-ray images, thus lowering false positives and false negatives.

Wang et al. [27] proposed LM-YOLOv8 to mitigate the common issues in real-time segmentation task such as limited capacity to simulate long-range spatial dependencies, making it challenging to precisely distinguish the borders of closely spaced objects and features redundancy or multi-scale feature parsing of hazardous items with different sizes and materials. Experimental results demonstrate the model's efficacy for X-ray hazardous object detection that is validated by its robustness on complex hard/hidden subsets.

Wang et al. [28] introduced YOLOv8n-GEMA, which is an improved detection model based on YOLOv8. This model provided many enhancements in the detection process by optimizing the feature capture and interdependencies among overlapping items, hence increasing the model's detection effectiveness in intricate cases. Moreover, the model significantly deals with multi-scale objects by employing auxiliary bounding boxes at different scale ratios for loss calculation ensuring precise predictions of bounding boxes. Experiments conducted show considerable enhancement in the recognition of overlapping and multi-scale targets.

Haq et al. [29] enhanced YOLOv8 detection model by introducing ESI-YOLO that incorporates Efficient Multi-Scale Attention (EMA) and Wise-IoU (WIoU) loss function for more accurate representation of multi-scale features and detection performance. Ablation results show that combining EMA and WIoU has a mutually beneficial impact that outperforms adding each component separately. Experimental results demonstrate improved accuracy without sacrificing computational efficiency, which makes it a viable option for realistic X-ray security inspection systems.

3.3 | YOLOv9 in Object Detection

Barr et al. [30] YOLOv9-TH, a real-time object detection model that incorporates an improved version of YOLOv9 along with a Transformer Head (TH). The proposed model incorporates an extra prediction head for identifying objects of diverse sizes and replaces the original prediction heads with transformer heads to utilize self-attention techniques. Experimental findings validate the model's resilience and appropriateness for practical applications, especially in the detection of minute objects in different applications.

Shah et al. [31] introduced a two-staged detector by integrating CNN for classifying object's type as threat or non-threat at the first stage and YOLOv9 for object tracking, accurate spatial detection and analysis of temporal persistence which are employed for the detection of prohibited items. The proposed methodology markedly enhances video-based threat identification and automated security monitoring systems, providing superior detection accuracy, resilient real-time performance, and considerable adaptability to intricate surveillance settings.

M et al. [32] introduced a robust detection model by combining the robustness of YOLOv9 with transformer-based and CNN-based structures for identifying firearms such as guns, knives, and rifles in real-time surveillance systems. The model has been trained on complex background cases like occlusion, human density, and limited illuminations. The findings suggest that YOLOv9 might be used to significantly improve crime prevention by stopping widespread, dangerous activity in surveillance systems. The study advances the field of smart surveillance by offering a high-performance, scalable approach that balances population safety with AI-driven object detection.

Narkhede et al. [33] enhanced YOOv9 model for real-time video surveillance systems to increase the safety of intersections. The proposed detection model makes use of carefully positioned cameras to record real-time traffic footage, which is then processed using YOLOv9 for precise detection and localization after undergoing picture enhancement procedures. Experimental results demonstrate the proposed model's robustness to handle various intersecting conditions while maintaining high performance, reliability, and speedy real-time processing.

Table 1. Summary of different YOLO model characteristics mentioned in the literature.

Ref.	Model	Methodology	Dataset	Results
[22]	YOLOv7	A system for automatically detecting banned objects in X-ray images	OPIXray	Achieved a precision of 73%, recall of 76%, mAP@0.5 of 49.1%, and mAP@0.5:0.9 of 39.7%.
[23]	YOLOv7	The detection model decreases resource waste, optimizes the convolutional layer structure, improves feature extraction.	CLCXray	Achieved 82.7% accuracy, recall 81.2%, and 80.2% mAP.
[24]	YOLOv7	The detection model combined multiple mechanisms to enhance detection.	SIXray	Achieved 4% higher accuracy than original YOLOv7.
[25]	YOLOv7	Enhances detection in complicated contexts with reliability for multi-scale attention.	PIDray	Achieved mAP of 68.7% which is 3.5% higher than baseline YOLOv7.
[26]	YOLOv8	Boosts the performance of identifying dangerous items in complex cases.	SIXray	Achieved mAP of 92.8% which is 2.4% higher than original YOLOv8.

[27]	YOLOv8	Mitigates common issues in real-time segmentation.	PIDray	Achieved mAP of 84.8% which is 4% higher than YOLOv8-seg.
[28]	YOLOv8	Increases the detection effectiveness in intricate cases by optimizing the feature capture and extraction.	SIXray, HiXray, CLCXray, and PIDray	Achieved 3.6%, 1.6%, 0.9%, and 3.4% higher mAP than YOLOv8 on datasets, respectively.
[29]	YOLOv8	The model performs accurate representation of multi-scale features and detection performance.	CLCXray and PIDray	Achieved 0.9% and 3.5% higher mAP than YOLOv8 on datasets, respectively.
[30]	YOLOv9	Identifies multi-scale objects and utilizes self-attention mechanism.	VisDrone2021 and DIOR	Achieved mAP@50 of 54.2% and 92.3% on datasets, respectively.
[31]	YOLOv9	Enhances video-based threat identification and automated security monitoring systems.	PETS2006 and ABODA	Achieved 99.81% accuracy, 99.67% precision, and 99.01% recall.
[32]	YOLOv9	Identifies dangerous activity in surveillance systems to prevent crimes.	N/A	Achieved mAP of 94.1%, 95.2% precision, and 92.8 recall.
[33]	YOLOv9	Makes use of carefully positioned cameras to record real-time traffic footage.	Real-world traffic data.	Achieved a 68% decrease in near-miss events and a 62.5% decrease in possible confrontations.

4 | Experimental Results

This section delves into the experimental results conducted on GDXray dataset [34] with YOLOv7, YOLOv8, and YOLOv9 independently to examine their performance in object detection task terms of mAP@50, Precision (P), Recall (R), F1-Score, and Accuracy (ACC).

4.1 | Dataset Setup

We have used the publicly available dataset GDXray, which is the only collection of temporally linked greyscale X-ray scans that show well-disguised contraband information at object detection tests, to evaluate the sustainability of specific YOLO models. The 19,407 scanned photos that make up the entire GDXray are divided into five categories: castings, welds, baggage, natural items, and settings. Nevertheless, we employed a subset of the baggage category, which included 2,909 photos with four categories of forbidden items: Shurekin (1232), Knife (77), Blade (237), and Gun (1350). 70% of the photos were used for training, 20% for testing, and 10% for validation in order to verify the detection capabilities of our suggested model. Additionally, we employed Focal Loss function to solve class imbalance problem found in the datasets using $\alpha = 0.75$ and $\gamma = 4$.

4.2 | Experimental Setup

We performed comparison studies on a subset of the GDXray dataset to illustrate the effectiveness of YOLO models. The Google Colabolatory framework, Intel iCore5 Tesla T4, 15095MiB, and Windows 10 64-bit operating system were used for all of our studies. The PyTorch torch-2.8.0+cu126 deep learning framework, Ultralytics 8.4.19, and the Python 3.12.12 programming language were utilized.

4.3 | Evaluation Metrics

We evaluated the results of the comparative study using common evaluation metrics for object detection task. Precision (P) refers to the ratio of correctly classified positive samples to all positively classified samples either correctly classified or not. Recall (R) refers to the measurement of positive samples identification by the model. F1-Score is the harmonic mean of precision and recall ranging from 0 to 1. Mean Average Precision (mAP) measures the average precision score across all classes by contrasting the detected box with the bounding box of the ground-truth box. at IoU threshold 0.5. Accuracy (ACC) refers to the model's ability to precisely identify and recognize the objects.

4.4 | Results and Discussion

Table 2 demonstrates the results obtained after conducting the comparative experiments conducted on YOLOv7, YOLOv8, and YOLOv9 regarding the aforementioned evaluation metrics in threat object detection task. Regarding mAP@50, it can be noted that YOLOv8 surpasses YOLOv7 and YOLOv9 by achieving 76.14% while YOLOv7 obtained 74.41% and YOLOv8 73.8%, demonstrating the model's ability to correctly and accurately identify things when compared to the actual objects. In terms of precision (P), YOLOv7 has recognized more true positives to the total number of detected items by obtaining 78.44% outperforming YOLOv8 with slight difference by 0.14% achieving 78.3% and YOLOv9 that achieved 76.4%. For the results of recall (R), YOLOv7 obtained the highest value by 78.3% compared to YOLOv8 with 74.2% and YOLOv9 that obtained a higher recall than YOLOv8 of 0.5% by achieving 74.7%. Regarding accuracy, YOLOv7 had the best performance of 80% compared with 78.6% and 76% for YOLOv8 and YOLOv9, respectively.

Table 2. Results of comparative models on GDXray dataset.

Model	mAP@50 (%)	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
YOLOv7	74.41	78.44	78.3	78.37	80
YOLOv8	76.14	78.3	74.2	76.25	78.6
YOLOv9	73.8	76.4	74.7	75.55	76

5 | Limitations and Future Work

Despite its widespread use and popularity, YOLO has a number of important drawbacks that restrict its effectiveness in some situations. Its inability to detect small things, especially in cluttered or complicated settings, is a significant obstacle. Because the resolution of feature maps tends to decrease with increasing model depth, small objects frequently lack adequate feature representation in YOLO's grid-based structure [7].

Despite the widespread adoption of YOLO models, its compatibility with varied real-world applications reveals a persistent challenge in handling highly dynamic scenes. For instance, in environments with rapid motion or in scenarios involving occlusions, they may still function poorly in situations with occlusions or in environments with fast motion. There is still room for improvement in the models' capacity to generalize across various kinds of blur and motion artefacts [14].

6 | Conclusion

This study provides a thorough examination of recent detection models of YOLO series and their adaptability in computer vision tasks, which have been employed in variant real-time detection models in many industries. A comprehensive review and comparative experiments were conducted on the widely used YOLOv7, YOLOv8, and YOLOv9 to demonstrate their capability in object detection process. Furthermore, the conducted experiments have shown the importance of using real-time detection models and the use of diverse evaluation terms for fair comparison. The versions of YOLO were compared regarding mAP@50, Precision (P), Recall (R), F1-Score, and Accuracy (ACC). From the experiments conducted on GDXray, YOLOv7 obtained the highest values in terms of precision and recall with 78.44% and 78.3%, respectively. Moreover, YOLOv8 achieved the highest mAP@50 of 76.14% demonstrating the model's capability to detect objects accurately compared with the ground truth ones. YOLOv9 's performance was slightly the lowest among other models but it has shown almost similar performance when compared to YOLOv8. Consequently, the experiment's outcomes will help researchers who want to utilize either model as a reference when choosing an experiment. However, further investigation into the mentioned models will reveal a distinct performance gap, as one model may outperform the other in many use cases and applications.

Author Contribution

All authors contributed equally to this work.

Funding

This research received no funding.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] C. Ma, L. Zhuo, J. Li, Y. Zhang, and J. Zhang, 'Occluded prohibited object detection in X-ray images with global Context-aware Multi-Scale feature Aggregation', Neurocomputing, vol. 519, pp. 1–16, Jan. 2023, doi: 10.1016/j.neucom.2022.11.034.

[2] C. Ma, L. Zhuo, J. Li, Y. Zhang, and J. Zhang, 'EAOD-Net: Effective anomaly object detection networks for X-ray images', IET Image Process., vol. 16, no. 10, pp. 2638–2651, Aug. 2022, doi: 10.1049/ipr2.12514.

[3] A. Liu et al., 'X-Adv: Physical Adversarial Object Attacks against X-ray Prohibited Item Detection', Feb. 2023.

[4] C. Zhao, L. Zhu, S. Dou, W. Deng, and L. Wang, 'Detecting Overlapped Objects in X-Ray Security Imagery by a Label-Aware Mechanism', IEEE Transactions on Information Forensics and Security, vol. 17, pp. 998–1009, 2022, doi: 10.1109/TIFS.2022.3154287.

[5] A. Singh and Dhiraj, 'Advancements in machine learning techniques for threat item detection in X-ray images: a comprehensive survey', Int. J. Multimed. Inf. Retr., vol. 13, no. 4, p. 40, Dec. 2024, doi: 10.1007/s13735-024-00348-2.

[6] C.-Y. Wang and H.-Y. M. Liao, 'YOLOv1 to YOLOv10: The fastest and most accurate real-time object detection systems', Aug. 2024.

[7] M. L. Ali and Z. Zhang, 'The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection', Computers, vol. 13, no. 12, p. 336, Dec. 2024, doi: 10.3390/computers13120336.

[8] Y. Y. Aung and K. Z. Oo, 'Detection of Guns and Knives Images Based on YOLO v7', in 2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT), IEEE, May 2024, pp. 1–6. doi: 10.1109/AIIoT58432.2024.10574710.

[9] M. Gueye, A. K. Diop, and A. D. Gueye, 'A survey of video sequence analysis in CCTV systems based on deep learning in the context of public safety', in 2025 International Conference on Artificial Intelligence, Computer, Data Sciences and Applications (ACDSA), IEEE, Aug. 2025, pp. 1–7. doi: 10.1109/ACDSA65407.2025.11166667.

[10] S. Zulfqar, S. Saeed, M. A. Zia, A. Ali, F. Mehmood, and A. Ali, 'A comprehensive overview of deep learning models for object detection from videos/images', Jan. 2026.

[11] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, 'You Only Look Once: Unified, Real-Time Object Detection', May 2016.

[12] M. Yurdakul, H. Sazak, M. Kotan, and Ş. Taşdemir, 'A Review of YOLO Family from YOLOv1 to YOLO26', Feb. 27, 2026. doi: 10.20944/preprints202602.1844.v1.

[13] M. Kotthapalli, D. Ravipati, and R. Bhatia, 'YOLOv1 to YOLOv11: A Comprehensive Survey of Real-Time Object Detection Innovations and Challenges', Aug. 2025.

[14] R. Sapkota et al., 'YOLO advances to its genesis: a decadal and comprehensive review of the You Only Look Once (YOLO) series', Jun. 2025, doi: 10.1007/s10462-025-11253-3.

[15] C. Gheorghe, M. Duguleana, R. G. Boboc, and C. C. Postelnicu, 'Analyzing Real-Time Object Detection with YOLO Algorithm in Automotive Applications: A Review', Computer Modeling in Engineering & Sciences, vol. 141, no. 3, pp. 1939–1981, 2024, doi: 10.32604/cmes.2024.054735.

[16] K. Das and A. Kr. Baruah, 'Object Detection in the Healthcare Industry Using Yolo: A Better Way Than CNN', 2025, pp. 29–41. doi: 10.1007/978-3-031-80813-5_3.

[17] H. Niu, D. Feng, and T. Hou, 'Research on foreign object intrusion detection in railway tracks based on MSL-YOLO', Journal of Engineering and Applied Science, vol. 72, no. 1, p. 136, Dec. 2025, doi: 10.1186/s44147-025-00708-7.

[18] T. Zoubek et al., 'Advancing precision agriculture with computer vision: A comparative study of YOLO models for weed and crop recognition', Crop Protection, vol. 190, p. 107076, Apr. 2025, doi: 10.1016/j.cropro.2024.107076.

[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', Jul. 2022.
[20] M. Yaseen, 'What is YOLOv8: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector', Aug. 2024.
[21] C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, 'YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information', Feb. 2024.
[22] S. Deena, J. L S, and S. S, 'Prohibited Object Detection using YOLOv7 on X-Ray Images of Airport Luggages (OPIXray)', in 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), IEEE, Jul. 2023, pp. 1–7. doi: 10.1109/ICCCNT56998.2023.10306920.
[23] N. Gan et al., 'YOLO-CID: Improved YOLOv7 for X-ray Contraband Image Detection', Electronics (Basel)., vol. 12, no. 17, p. 3636, Aug. 2023, doi: 10.3390/electronics12173636.
[24] B. Jing, P. Duan, L. Chen, and Y. Du, 'EM-YOLO: An X-ray Prohibited-Item-Detection Method Based on Edge and Material Information Fusion', Sensors, vol. 23, no. 20, p. 8555, Oct. 2023, doi: 10.3390/s23208555.
[25] L. Zhang, Z. A. Huang, C. Shi, H. Ma, X. Li, and X. Wu, 'MFPIDet: improved YOLOV7 architecture based on multi-scale feature fusion for prohibited item detection in complex environment', Complex & Intelligent Systems, vol. 10, no. 6, pp. 8095–8108, Dec. 2024, doi: 10.1007/s40747-024-01580-3.
[26] M. Chen, Z. Zhang, N. Jiang, X. Li, and X. Zhang, 'YOLO-SRW: An Enhanced YOLO Algorithm for Detecting Prohibited Items in X-Ray Security Images', IEEE Access, vol. 13, pp. 68323–68339, 2025, doi: 10.1109/ACCESS.2025.3560840.
[27] T. Wang, P. Yuan, and A. Wang, 'Dangerous Goods Detection in X-Ray Security Inspection Images Based on Improved YOLOv8-seg', Electronics (Basel)., vol. 15, no. 5, p. 1112, Mar. 2026, doi: 10.3390/electronics15051112.
[28] A. Wang, P. Yuan, H. Wu, Y. Iwahori, and Y. Liu, 'Improved YOLOv8 for Dangerous Goods Detection in X-ray Security Images', Electronics (Basel)., vol. 13, no. 16, p. 3238, Aug. 2024, doi: 10.3390/electronics13163238.
[29] A. Haq, N. Suciati, and N. D. Bui, 'ESI-YOLO: Enhancing YOLOv8 with Efficient Multi-Scale Attention and Wise-IoU for X-Ray Security Inspection', International Journal of Robotics and Control Systems, vol. 5, no. 3, pp. 1770–1789, Jul. 2025, doi: 10.31763/ijrcs.v5i3.1983.
[30] M. Barr, 'Coupling the Power of YOLOv9 with Transformer for Small Object Detection in Remote-Sensing Images', Computer Modeling in Engineering & Sciences, vol. 143, no. 1, pp. 593–616, 2025, doi: 10.32604/cmes.2025.062264.
[31] S. F. A. Shah, M. Q. Khan, A. A. Alsuwaylimi, and L. A. Alenezi, 'A Two-Stage Spatiotemporal CNN-YOLOv9 Framework for High-Precision Real-Time Abandoned Object Detection in Public Surveillance Videos', IEEE Access, vol. 14, pp. 35983–35997, 2026, doi: 10.1109/ACCESS.2026.3669588.
[32] A. S. M, A. Ashish, B. Hazela, A. P. Singh, P. Asthana, and S. Singh, 'Weapon Detection in Public Surveillance Footage Using YOLOv9 for Crime Prevention Systems', in 2025 3rd International Conference on IoT, Communication and Automation Technology (ICICAT), IEEE, Dec. 2025, pp. 1–6. doi: 10.1109/ICICAT68430.2025.11414477.
[33] M. M. Narkhede, R. S. Chaudhari, and N. B. Chopade, 'Real-time object detection and warning for road intersection safety: A YOLOv9-based approach', Sādhanā, vol. 51, no. 1, p. 53, Feb. 2026, doi: 10.1007/s12046-025-03057-6.
[34] D. Mery et al., 'GDXray: The Database of X-ray Images for Nondestructive Testing', J. Nondestr. Eval., vol. 34, no. 4, p. 42, Dec. 2015, doi: 10.1007/s10921-015-0315-7.