**Journal of Information Systems Engineering and Management**
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

**Research Article**

# YOLO Accuracy Enhancement for Dense and Dynamic Scenes

Chahreddine Medjahed[1], Narimane Wafaa Krolkral[2], Freha Mezzoudj[3], Ahmed Slimani[4]
[1]*Computer Science Department, Hassiba Benbouali Chlef University,c.medjahed@univ-chlef.dz*
[2]*EEDIS Laboratory, Djillali Liabes University,wafaa.krolkral@univ-sba.dz*
[3]*The National Polytechnic School of Oran Algeria, freha.mezzoudj@enp-oran.dz*
[4]*LabRI-SBA Lab Ahmed Draia University – ADRAR, ah.slimani@esi-sba.dz*

**ARTICLE INFO**

Received: 24 Dec 2024

Revised: 12 Feb 2025

Accepted: 26 Feb 2025

**ABSTRACT**

Computer vision relies on object detection and supports different applications, including self-driving cars, intelligent monitoring, and so on. The deep neural network "You Only Look Once" with his version (YOLOv8) has greatly helped the object detection tasks by improving the accuracy and making the inference process faster. Even so, it is still difficult to achieve excellent real-time performance when conditions include clutter, obstructions, and objects that are small. This article introduces some interesting new improvements to the YOLOv8 architecture that involve two modules Cross Stage Partial Transformer block (C3TR) and Adaptive Downsampling (Adown), as well as the use of multiple sizes of receptive fields to better capture multi-scale features.

Experimental results show that the improved YOLOv8 outperforms the baseline in both accuracy and across various datasets. The outcomes of this research confirm the effectiveness of these enhancements. The modified models showed consistent improvements in accuracy, with mAP scores increasing by +2.3% to +5.8% over the original YOLOv8.

**Keywords:** Deep Learning; YOLOv8; Small Object Detection; Scene Understanding.

## INTRODUCTION

In today's modern world, where technology is evolving faster than ever, the ability of intelligent machines to see and understand their surroundings is no longer science fiction; it is a necessity. Advanced artificial intelligence models are focusing on deep learning. Those models have been trained on real data with heterogeneous sources—different sensors, cameras, and so on to be able to recognize unseen data. In general, those models will be used successfully in such related domains as analyzing current situations, predicting future scenarios, proposing solutions, and developing intelligent systems.

Object detection methods play an important role in endowing intelligent systems with the ability to see and interact with the physical world. Actually, YOLOv8, the most widely used version, has become a top choice for both cutting-edge research and real-world applications due to its remarkable performance. Many practical applications use YOLOv8-based object detection [10-13], making it a compelling area for ongoing research and development. In autonomous driving, it enables real-time detection of pedestrians, vehicles, and traffic signs, which contribute to ensuring safer navigation. In the medical field, where speed and precision are essential, YOLO models assist in identifying abnormalities in medical images, such as tumors in MRI scans. Also, in intelligent agriculture, they are used to monitor livestock or detect plant diseases from drone imagery. These real-world uses make it clear that fast and dependable models like YOLO play a key role in today's intelligent systems.

However, even with the remarkable YOLO variants' abilities, and especially the version eight, many challenges and weaknesses still stand. Among the reasons for those critical situations are the unpredictable and uncontrolled quality of real data, which contribute to increasing the challenges with even advanced object detection models [2]. This situation provides strong motivation to keep improving these models so they can perform even better in real-world conditions. Motivated by these challenges, the objective of this paper is to make YOLOv8 more reliable in real-life

**Journal of Information Systems Engineering and Management**
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

**Research Article**

situations. Some new improvements and updates at many levels on many modules among its whole architecture are proposed and tested on different types of data to see how well they perform in more complicated settings.

Therefore, the proposed methodology applied in this paper involved a multi-step strategy. We began by evaluating the baseline performance of YOLOv8 under complex visual conditions to identify its limitations in real-world scenarios. Based on what we learned, we made some improvements to the design, such as adding the Cross-Stage Partial Network with Transformer (C3TR) attention module [4] to help focus on important features, using dual-scale spatial pooling to capture both small and large details, and applying asymmetric downsampling to keep spatial information while not making the model much bigger.

We implemented the proposed modifications and compared them to three variants of YOLOv8. We have validated our results in terms of mean Average Precision (mAP), Frames Per Second (FPS), and parameter size to make visual systems more flexible and intelligent. The targeted goal is that the model not only will function well in clean and controlled conditions but also be able to flourish in the highly complex and uncontrollable conditions of the real world. Therefore, the main objectives of this research are as follows:
- Analyze the baseline performance of YOLOv8 across several complex datasets to understand its real-world limitations.
- Suggest changes to the design, like adding attention modules (e.g., C3TR), using dual-scale pooling, and applying asymmetric downsampling, to improve how well features are represented and to increase detection accuracy.
- Validate the baselines and the obtained models, referred to as YOLOv10, YOLOv11, and our custom variant, against the original YOLOv8 using standard metrics like mean Average Precision (mAP), Frames per Second (FPS), and model size.
- Finally, to ensure the proposed improvements are generalizable across domains, test their performance in multiple real-world applications.

## RELATED WORK

The object detection techniques can be useful in daily applications, such as in intelligent agriculture, in the food industry, in other industry fields, in medical surgery, in car 'traffic, in earth observation, and so on [10-13]. Those systems are trained or pretrained and then fine-tuned on specific images to be able to tackle their tasks. Among the tools driving this progress are deep learning models, particularly the advanced convolutional neural network for object detection You Only Look Once (YOLO) family [1, 2].

In general, YOLO deep learning models are known for their real-time speed and impressive accuracy. Each new version focuses on improving the speed, accuracy, and flexibility of the underlying neural network. The first version, YOLOv1, is fast and able to detect objects in just one pass. However, it had difficulties recognizing small or overlapping objects. YOLOv2 introduces the use of anchor boxes and uses the Darknet-19 neural network as a backbone. Then YOLOv3 used even Darknet-53, a deeper neural network, as a backbone and added multi-scale detection [2]. YOLOv4 introduced two novel models, Cross Stage Partial Networks (CSPNet) and Path Aggregation Network (PANet), to improve feature extraction and path aggregation, respectively [5]. The YOLOv5 added automatic anchor box learning and scalable model sizes [6]. YOLOv7 improved the design by using re-parameterized convolutions and Extended Efficient Layer Aggregation Network (E-ELAN) structures, which made the neural network quicker and better for real-time use.

Recently, the novel YOLOv8 is considered a powerful and recommended model to be used. YOLOv8, developed by the open-source AI research team Ultralytics, uses a detection head and a module called Concatenate-to-Fuse (C2f). which helps reuse important features. While this version 8 architecture is powerful, it still struggles with detecting some small or densely packed objects. Some recent research focused on attention mechanisms such as the Convolutional Block Attention Module (CBAM) [8] and the transformer-based layers Cross-Stage Partial Transformer (C3TR) [9] to help models focus on the most relevant features.The figure 1 illustrates the rich architecture of YOLOv8, is inspired from the available online architecture at https://github.com/ultralytics/ultralytics/issues/189.

**Journal of Information Systems Engineering and Management**
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

**Research Article**

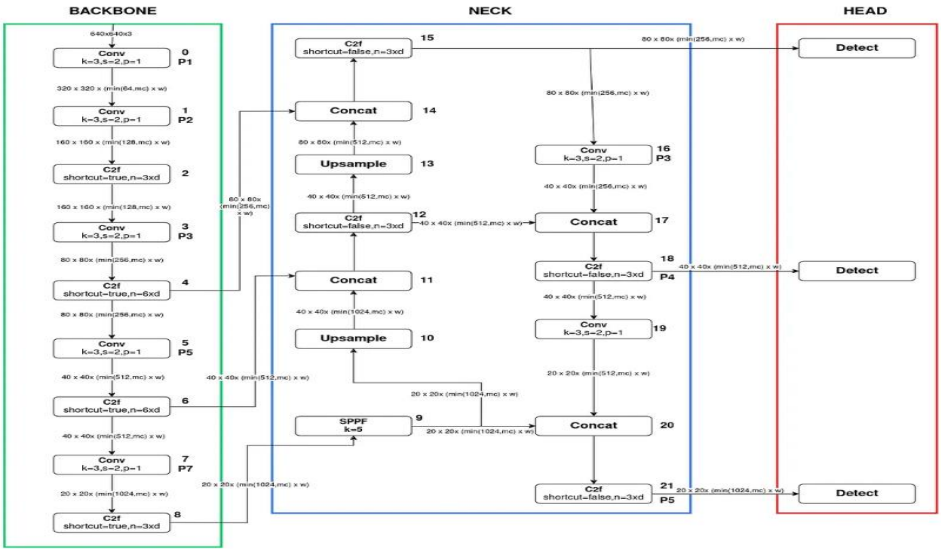


Figure 1: Illustration of the YOLOv8 architecture_ inspired from ultralytics github.

In [10], to recognize many crops, the authors proposed a hybrid approach by combining the YOLOv8 model with the threshold-based "Density-Based Spatial Clustering of Applications with Noise" (DBSCAN) model, least squares, and B-splines. They also created the DCGA-YOLOv8 by adding deformable convolutions to help it better adapt to crops that come in different sizes and shapes, along with a Global Attention Mechanism (GAM). This proposed model achieved detection performance, with F1 scores of 96.4%, 97.1%, and 95.9%, and mAP values of 98.9%, 99.2%, and 99.1% for cabbage, kohlrabi, and rice, respectively. For crop row clustering, the DBSCAN algorithm showed performance, reaching clustering rates of 98.9%, 97.9%, and 100% for the same crops.

In [11], the authors address the problem of detecting maize leaf diseases in dense field conditions using standard YOLO-based models. They propose an approach called GhostNet_Triplet_YOLOv8s by replacing parts of the network, specifically the Coarse-to-Fine (C2f) and Conv (Convolutional) modules, with GhostNet and C3 Ghost blocks. In addition, they add an attention mechanism. The obtained results indicate a precision rate of 87.50%, a recall rate of 87.70%, and an mAP@0.5 of 91.40%. This approach demonstrates a 0.3% improvement in mean precision (mAP), a 50.2% decrease in model size, and a notable 43.1% reduction in floating-point operations (FLOPs) when compared to YOLOv8.

The study in [12] presents a method that uses the YOLOv8 architecture to automatically identify major leaf diseases in onion crops, including anthracnose, stemphylium blight, purple blotch (PB), and Twister disease. The experimental results show that the improved YOLOv8 model, called YOLO-ODD and upgraded with the CABM and DTAH attention mechanisms, performs better than the YOLOv5 and YOLOv8 models in most disease categories, especially in finding anthracnose, purple blotch, and Twister disease. The suggested model attained notable metrics, with an overall accuracy of 77.30%, precision of 81.50%, and recall of 72.10%. This deep learning approach provides a reliable solution for the classification and detection of onion leaf diseases, integrating high accuracy and real-time processing.

The authors in [13] offer a streamlined vehicle identification model with an enhanced YOLOv8 architecture, designed to enhance performance in intricate traffic scenarios. They incorporate an adaptive subsampling module to enhance feature extraction efficiency, particularly for small or partially occluded objects, while preserving a lightweight architecture. A convolution detection head was also improved to minimize parameters. Tests conducted on the KITTI 2D and UA-DETRAC datasets give enhancements, including a 2% rise in mean average precision (mAP), a 12% augmentation in processing rate (FPS), a 33% decrease in parameters, and a 28% decline in floating point operations.

The study conducted by [14] presents an enhanced iteration of the YOLOv8 model dedicated to the identification of dynamic objects in videos. The application of pre-processing, along with structural modifications, allows the model

**Journal of Information Systems Engineering and Management**
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

**Research Article**

to improve sensitivity to motion. Tests conducted on KITTI and LASIESTA databases exhibit exceptional performance, achieving an accuracy of 90% and a mean Average Precision (mAP) score of 90%.

## METHODOLOGY

To increase the accuracy of YOLOv8, we suggest new architectures for the backbone, the neck, and the detection head (see figure 2). With these updates, we suppose that the model can recognize close details and also plan the scan of remote details in the image. A central innovation in our approach is the introduction of the C3TR block. This hybrid module integrates Transformer layers into the conventional C3 block originally used in YOLOv8.The motivation behind this design is to leverage the complementary strengths of Convolutional Neural Networks (CNNs) and Transformers. CNNs excel at extracting local features such as edges, textures, and fine details by focusing on small, localized image patches. However, they struggle to model relationships between distant parts of an image. In contrast, transformers provide a mechanism for capturing long-range dependencies and global context through self-attention mechanisms. The C3TR block, as illustrated in figure 2, achieves this by splitting input features into two parallel branches: Regular Vision Branch (Convolutions): Focuses on local feature extraction by processing small image regions.

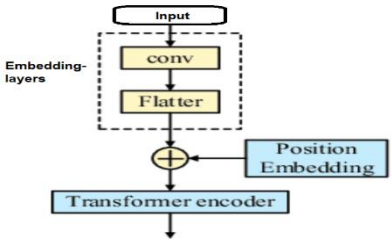


Figure 2: Global idea of C3TR used mechanism.

This branch detects fine-grained textures and edges essential for object localization and classification in complex scenes with occlusions or hidden objects. The transformer branch captures global relationships by attending to all regions of the image simultaneously. This helps the model to connect spatially distant but semantically related features.

Together, these two branches act as a complementary team: the convolutional path serves as a “Detail Detective,” while the Transformer path functions as a “Big-Picture Detective.” This synergy significantly enhances the model's ability to recognize challenging objects and improves overall detection robustness. In addition to integrating the C3TR block, the backbone architecture is deepened to extract richer hierarchical features.

In YOLO architecture, the number of C2f blocks in the early layers is increased from 3–5 to 5–8 in the middle layers, allowing the model to capture more complex patterns or features (see figure 2). Residual connections within these blocks ensure stable gradient flow during training, avoiding problems like vanishing gradients. The C2F blocks work as illustrated in figure 3. This adjustment is motivated by the philosophy of “quality over quantity,” which means instead of increasing the number of feature maps through the addition of extra convolutional layers or C2f (coarse-to-fine) modules. The proposed approach prioritizes improving the representational quality of the existing feature maps. The transformer-based C3TR block replaces the standard C2f block at layer 6 (P4/16 scale).

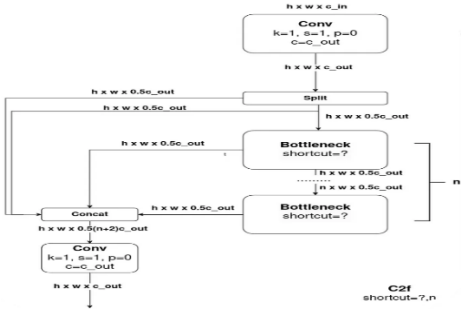


Figure 3: Illustration of global C2F block.

**Journal of Information Systems Engineering and Management**
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

**Research Article**

Effective object detection requires the model to recognize objects at varying scales. To address this, we propose a dual SPPF (Spatial Pyramid Pooling-Fast) architecture, as illustrated in figure 4. Our design concatenates the original SPPF layer, which uses a 5 × 5 kernel, with an additional SPPF employing a larger 7 × 7 kernel. The dual kernels simulate the fluctuating receptive field sizes observed in vision systems, enabling the network to capture complementary features at different spatial resolutions.

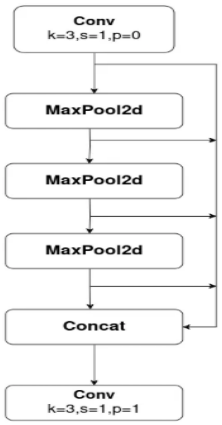


Figure 4: Illustration of a global SPPF block.

The concatenation of these pooled features maintains unique information from both receptive fields, avoiding information loss that could arise from early fusion methods. This choice improves the model's feature representation capability for objects of diffrent sizes. Another crucial enhancement is possible with the use of ADown (Asymmetric Downsampling) [3] to maintain rich feature semantics. A Down processes the input feature map through two branches:

- Branch 1: Applies a 3 × 3 convolution with stride 2, reducing spatial dimensions and preserving texture-based features.
- Branch 2: Uses MaxPooling to shrink the feature map, followed by a 1 × 1 convolution to compress the information.

Then, these two asymmetric processes are concatenated [1], which helps that the backbone becomes deeper and more sophisticated, and the detection head is simplified.

The original YOLOv8 head consists of three repetitions of multiple C2f blocks, incorporating complex concatenation and upsampling layers. The proposed modification reduces the total number of blocks but increases the depth of the remaining layers by using five repetitions of C2f blocks. A backbone, equipped with the C3TR block and dual SPPF, outputs rich features, which contribute to reprocessing fundamental patterns.

Additionally, having fewer parameters in the head helps mitigate the risk of overfitting to noisy training data, such as confusing background textures with actual objects. This approach aligns with the principle of Occam's Razor, which favors simpler models when sufficient representational power is already present earlier in the network. The addition of hybrid C3TR blocks, deeper backbone layers, expanded SPPF multiple receptive fields, and a streamlined detection head means YOLOv8 detects objects more accurately. As a result, they increase the ability of the model to extract complex features from multiple views and scales while focusing on controlling how much the model can both learn and generalize. Our Our enhancements make use of attention functions and multi-scale learning, which are widely used now in computer vision research to meet the issues of complex object detection [2]. All these improvements are illustrated in figure 5.

**Journal of Information Systems Engineering and Management**
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

**Research Article**

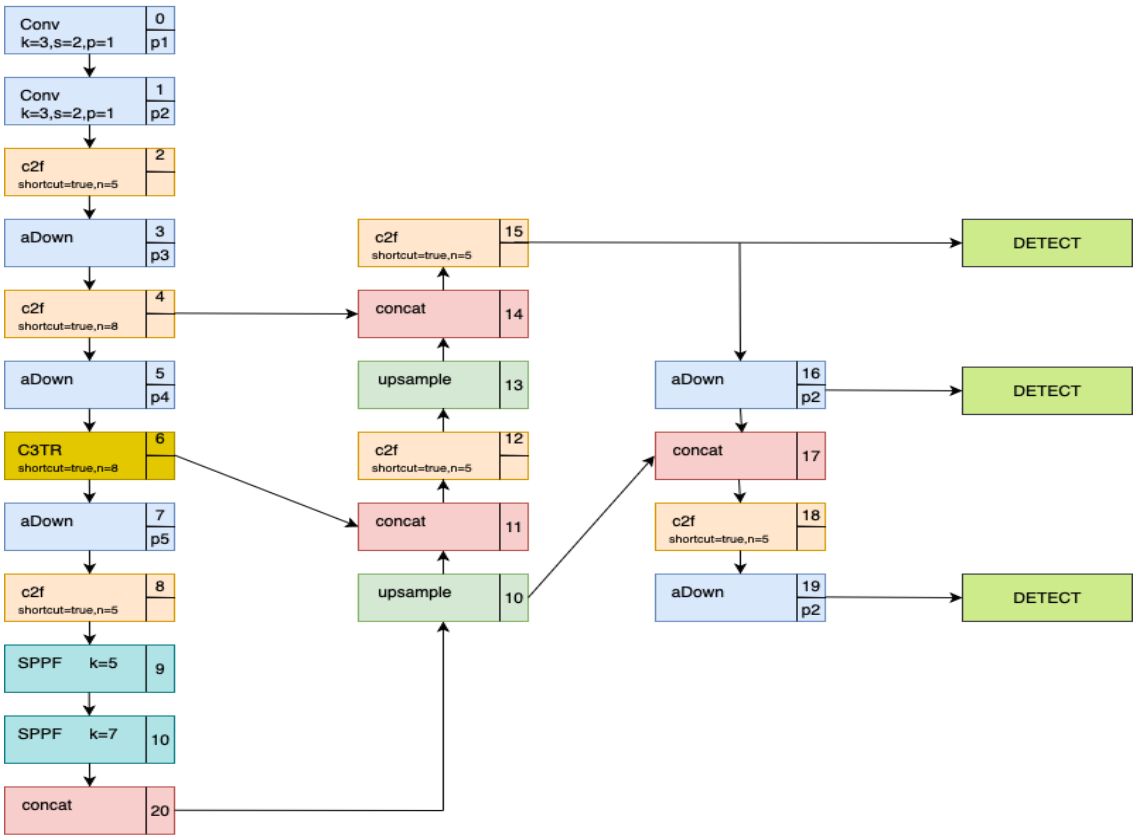


Figure 5: Architecture of the Improved YOLOv8 Model

## RESULTS

This section shows the results of the improved model, with detection accuracy and inference speed being the points of interest. We will start with checking the baseline performance and then point out the improvements that we made with our changes. The objective is to show the effectiveness of the architectural and optimization modifications in the system, particularly in complex and changing environments. To determine a performance baseline, we used YOLOv8n without changing any settings on 2 datasets, the Facemask Detection and Brain Tumor Detection datasets. These datasets presented some challenges, such as small object sizes, frequent occlusions, and class imbalance.

The results in Table 1 demonstrate the enhanced efficacy of our proposed model compared to YOLOv8n, YOLOv10n, and YOLOv11n across two datasets: Facemask Detection and Brain Tumor Detection. Our model attains the maximum mAP@0.5 and mAP@0.5:0.95 across both challenges. It achieves a precision of 0.855 on the facemask dataset and 0.914 on the medical dataset. These enhancements illustrate the model's robust generalization and dependability, particularly in essential applications where accuracy and precision are crucial.

To test the proposed model, we trained and tested it on another dataset, including on the Football Player Detection, AFFECTNET (Facial Expression), PPE Compliance, and Cattle. We compared the obtained results with the results of YOLOv8 under the same training conditions to ensure a consistent comparison, as shown in Table 2. In every instance, our model surpassed YOLOv8 regarding mean Average Precision (mAP) at both Intersection over Union (IoU) thresholds (0.5 and 0.5:0.95), demonstrating significant improvements in precision. With the Football Player Detection dataset, our model attained a mAP 0.5:0.95 of 0.524 compared to 0.503 for YOLOv8n and exhibited marginal enhancements in both precision and recall.

**Journal of Information Systems Engineering and Management**
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

**Research Article**

| Dataset | Model | mAP 0.5 | mAP 0.5:0.95 | Precision | Recall |
|---|---|---|---|---|---|
| Facemask Detection | Yolov8n | 0.738 | 0.486 | 0.743 | 0.738 |
| | Our model | **0.779** | **0.497** | **0.855** | 0.704 |
| | Yolov10n | 0.656 | 0.436 | 0.698 | 0.573 |
| | Yolov11n | 0.716 | 0.472 | 0.836 | 0.648 |
| Brain Tumor Detection | Yolov8n | 0.917 | 0.694 | 0.907 | 0.847 |
| | Our model | **0.933** | **0.706** | 0.914 | **0.882** |
| | Yolov10n | 0.869 | 0.638 | 0.821 | 0.806 |
| | Yolov11n | 0.92 | 0.69 | 0.917 | 0.855 |

Table 1: Comparison of Object Detection Performance Across YOLO Variants on Two Benchmark Datasets.

| Dataset | Model | mAP 0.5 | mAP 0.5:0.95 | Precision | Recall |
|---|---|---|---|---|---|
| Football Player Detection | YOLOv8n | 0.812 | 0.503 | 0.861 | 0.753 |
| | Our model | **0.819** | **0.524** | **0.867** | **0.77** |
| PPE Compliance | YOLOv8n | 0.738 | 0.464 | 0.689 | 0.753 |
| | Our model | **0.752** | **0.474** | 0.686 | **0.793** |
| Cattle Detection | YOLOv8n | 0.757 | 0.670 | 0.753 | 0.726 |
| | Our model | **0.762** | **0.677** | **0.959** | 0.724 |

Table 2: Benchmarking Our Model Against YOLOv8n on Diverse Real-World Datasets.

The best enhancement occurred in the Cattle Detection dataset, where our model exhibited comparable mAP and recall while achieving a substantially higher precision (0.959 vs. 0.753). These results affirm that our model exhibits superior generalization across varied domains in comparison to the baseline.

At this stage, we proceed with the testing phase to validate the performance of our model on unseen data used to demonstrate its robustness and effectiveness compared to YOLOv8n and other recent YOLO variants.

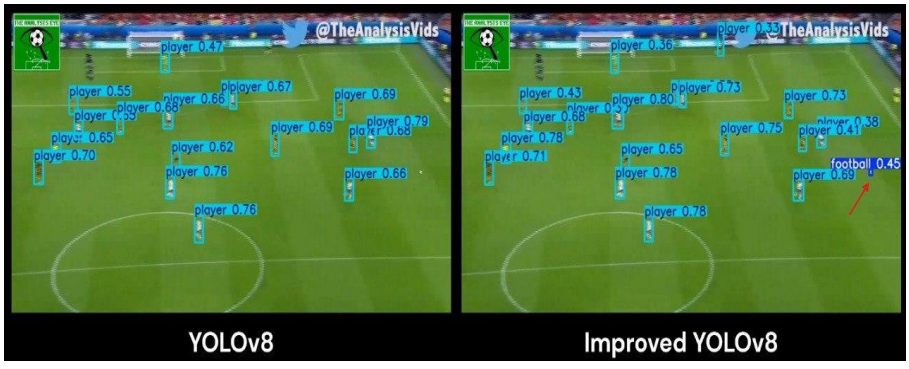


Figure 6: Football detection: Comparison Between YOLOv8 and Improved YOLOv8.

**Journal of Information Systems Engineering and Management**
2025, 10(56s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

**Research Article**

Figure 6 presents a comparative analysis of detection outcomes on a football match frame utilizing YOLOv8 (left) and our enhanced version (right). Although both algorithms properly identify the majority of players, our approach exhibits superior localization precision and confidence metrics. Our model effectively identifies and labels the football (confidence 0.45), whereas YOLOv8 entirely overlooks it. This underscores an augmented sensitivity to diminutive or low-contrast items, a vital enhancement for activities necessitating meticulous identification. The enhanced model consistently exhibits high confidence for the majority of identified players, signifying superior feature extraction and classification capabilities.

Figure 7 shows the comparison of personal protective equipment (PPE) detection by YOLOv8 and the new model. It can be observed that YOLOv8 is able to detect only a few object classes at very low confidence. 'Safety Vest 0.28 ' and 'Hardhat 0.38' are the objects discovered with a low confidence of just 0.28 and 0.38 respectively. In contrast, our improved model is capable of providing a more accurate and comprehensive detection. It can not only give high-confidence classifications of protective equipment like vests and hardhats but also find the ones that are falsely identified such as the violations of mandatory requirements. Specifically, it rightfully identifies two persons who have no hardhats, named as "NO-Hardhat 0.25" and "NO-Hardhat 0.24", which are practically hidden from the baseline YOLOv8 model. This breakthrough in technology shows how a safety compliance system can be developed for the building industry and factory safety that is able to identify fine-grain social norms and be highly sensitive to context variations of the real world.


Figure 7: Safety Gear Detection: Comparison Between YOLOv8 and Improved YOLOv8.

The results of this study demonstrate that the improved YOLOv8 model significantly outperforms the default YOLOv8 in various object detection tasks. The improved model has gotten better accuracy, precision, and recall, particularly on complex scenes, small objects, and occlusion. These improvements can be credited to the architectural changes that enhanced the generality of the model in a wide variety of conditions. A few drawbacks were noted in very noisy inputs, where the model accuracy was reduced a bit. Still, the results demonstrate the flexibility of the suggested improvements and their potential suitability to be implemented in practice in the domains where both speed and high detection accuracy are needed. Ongoing or future efforts can be aimed at improving the model architecture further and testing the model robustness on real-time data.

## CONCLUSION

The object detection tasks are essential in this period. This article tackles the challenges faced by YOLOv8 in delivering real-time object detection by proposing targeted improvements. By incorporating attention-based methods such as C3TR, introducing dual-scale pooling of spatial information, and applying asymmetric downsampling, the model achieves higher detection accuracy across different scenarios. The results clearly show that these enhancements introduce only a minor computational overhead, making them practical for real-world applications such as medical imaging, surveillance, agriculture, and autonomous systems. The model's adaptability across various domains further reinforces its reliability.

According to these research results, we confirm the effectiveness of the proposed approaches. The modified models showed promising improvements in accuracy, with mAP scores increasing by +2.3% to +5.8% over the initial YOLOv8. Moreover, the use of inference optimization tools led to up to 60% faster processing speeds, making the models more suitable for time-sensitive and resource-limited environments. Thanks to these promising findings, future studies can be conducted on better, more reliable, and more precise object detection methods. As future work, we plan to explore other deeper transformer architectures, particularly for detecting objects in crowded or unoccluded scenes.

## REFRENCES

[1] Wang, C. Y., Liao, H. Y. M., & Wu, J. C. (2023). YOLOv9: Learning what you want to learn using programmable gradient information (arXiv:2303.11564). arXiv. https://arxiv.org/abs/2303.11564.
[2] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 779–788). https://doi.org/10.1109/CVPR.2016.91.
[3] Li, Y., & Zhang, X. (2023). Structure diagram showing a comparison between ADown downsampling and regular downsampling. ResearchGate. https://www.researchgate.net/figure/Structure-diagram-showing-a-comparison-between-ADown-downsampling-and-regular_fig3_381496805.
[4] Vasanthi, P., & Mohan, L. (2023). A reliable anchor regenerative-based transformer model for x-small and dense objects recognition. Neural Networks, 165, 809–829. https://doi.org/10.1016/j.neunet.2023.06.014.
[5] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection (arXiv:2004.10934). arXiv. https://arxiv.org/abs/2004.10934.
[6] Jocher, G. (2020). YOLOv5 by Ultralytics. GitHub. https://github.com/ultralytics/yolov5.
[7] 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:2207.02696). arXiv. https://arxiv.org/abs/2207.02696.
[8] Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). CBAM: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 3–19). https://doi.org/10.1007/978-3-030-01234-2_1.
[9] Huang, T., Liu, J., &Deng, Z. (2023). C3TR: Efficient transformer-based attention module for object detection (arXiv:2301.07536). arXiv. https://arxiv.org/abs/2301.07536.
[10] Shi, J., Bai, Y., Zhou, J., & Zhang, B. (2023). Multi-crop navigation line extraction based on improved YOLO-v8 and threshold-DBSCAN under complex agricultural environments. Agriculture, 14(1), 45.
[11] Li, R., Li, Y., Qin, W., Abbas, A., Li, S., Ji, R., ... & Yang, J. (2024). Lightweight network for corn leaf disease identification based on improved YOLO v8s. Agriculture, 14(2), 220.
[12] Raj, A., Dawale, M., Wayal, S., Khandagale, K., Bhangare, I., Banerjee, S., ... & Gawande, S. (2025). YOLO-ODD: an improved YOLOv8s model for onion foliar disease detection. Frontiers in Plant Science, 16, 1551794.
[13] Zhang, M., & Zhang, Z. (2025). Research on Vehicle Target Detection Method Based on Improved YOLOv8. Applied Sciences, 15(10), 5546.
[14] Safaldin, M., Zaghden, N., & Mejdoub, M. (2024). An improved YOLOv8 to detect moving objects. IEEE Access.