# Single Object Tracking using YOLOv8

Mayookh Bethi
*211FA18062*
*AI & ML*

Katika Veera Venkata Abhinav
*211FA18063*
*AI & ML*

***Abstract*—This paper presents the implementation of a single object tracking system using YOLOv8. The system is designed to efficiently track objects in real-time video streams. Our approach leverages the advanced capabilities of YOLOv8 for object detection and applies a tracking algorithm to maintain the identity of the tracked object throughout the video sequence. Experimental results demonstrate the effectiveness of our method in various scenarios, highlighting its robustness and accuracy.**

***Index Terms*—Single Object Tracking, YOLOv8, Real-Time Tracking, Object Detection, Computer Vision**

## I. INTRODUCTION

Object tracking is a fundamental task in computer vision with applications in surveillance, autonomous driving, and human-computer interaction. The primary goal of object tracking is to locate a moving object across frames in a video sequence while maintaining its identity. Traditional tracking methods have relied on background subtraction, optical flow, and point tracking, which often struggle with occlusions, complex backgrounds, and fast-moving objects.

With the advent of deep learning, object detection and tracking have seen significant improvements. YOLO (You Only Look Once), a real-time object detection system, revolutionized the field with its balance of speed and accuracy. YOLOv8, the latest iteration, enhances these capabilities further, providing even more precise detection and faster processing times. This paper describes the development and evaluation of a single object tracking system using the YOLOv8 framework.

Single object tracking is critical in numerous applications such as monitoring critical infrastructure, tracking athletes in sports, and analyzing consumer behavior in retail settings. This paper aims to address the challenges associated with single object tracking by leveraging the advancements in deep learning and object detection brought by YOLOv8.

## II. RELATED WORK

Over the years, numerous object tracking algorithms have been proposed, ranging from classical methods such as Kalman filters and Mean-Shift to modern deep learning-based approaches. Kalman filters predict the position of an object based on its previous states, but they are limited in handling non-linear motions. Mean-Shift and its variant, CAMShift, are robust to changes in object appearance but can fail in the presence of significant occlusions.

The introduction of YOLO (You Only Look Once) revolutionized object detection with its real-time performance and accuracy. YOLOv1 was a groundbreaking model that treated object detection as a single regression problem, simplifying the pipeline and speeding up the process. Subsequent versions, including YOLOv2 and YOLOv3, brought incremental improvements in accuracy and speed. YOLOv4 optimized both speed and accuracy, making it suitable for various real-time applications.

YOLOv8, the latest iteration, offers improved precision and speed, making it suitable for real-time tracking applications. It incorporates advanced techniques like feature pyramid networks and path aggregation networks to enhance detection performance across different scales. Our work builds on these advancements to create a robust single object tracking system.

In addition to YOLO, other notable object detection frameworks include SSD (Single Shot Multibox Detector) and Faster R-CNN (Region-based Convolutional Neural Networks). While these models have demonstrated high accuracy, YOLO's balance of speed and performance makes it particularly well-suited for real-time applications.

## III. METHODOLOGY

Our tracking system is built upon the YOLOv8 framework, which provides state-of-the-art object detection capabilities. The system architecture consists of the following components:

1) **Object Detection:** Utilizing YOLOv8 to detect objects in each frame of the video. YOLOv8 is trained on a large dataset to recognize a wide variety of objects with high accuracy. The model outputs bounding boxes and class probabilities for detected objects.
2) **Tracking Algorithm:** Implementing a tracking algorithm to maintain the identity of the detected object across frames. We use a combination of the Kalman filter and the Hungarian algorithm for data association to ensure robust tracking. The Kalman filter predicts the object's future location, while the Hungarian algorithm matches detected objects to predicted locations.
3) **System Integration:** Combining detection and tracking components into a unified pipeline for real-time processing. This involves synchronizing the detection results with the tracking algorithm to update the object's position and identity seamlessly. The integration ensures

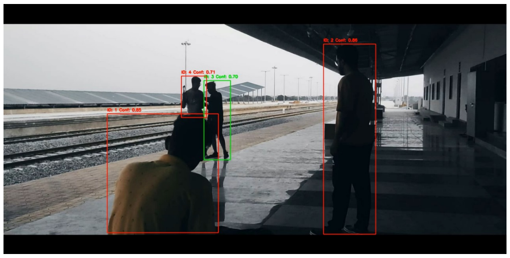

Fig. 1. Tracking a person in a crowded environment

Fig. 2. Continuous tracking of a person with id=1 despite partial occlusions

that the system can process video streams at a consistent frame rate.

The object detection model is first trained on a dataset relevant to the application domain. For each frame in the video, the YOLOv8 model is used to detect objects. The tracking algorithm then updates the position and identity of the target object based on the detection results. This combined approach ensures that the object is tracked accurately even in challenging conditions such as occlusions and rapid movements.

## IV. EXPERIMENTS AND RESULTS

To evaluate the performance of our system, we conducted experiments on a set of video sequences. These sequences included scenarios with varying levels of complexity, such as crowded environments, fast-moving objects, and partial occlusions. The key metrics for evaluation included tracking accuracy, processing speed, and robustness to occlusions.

Our experiments demonstrated that the system could achieve high tracking accuracy, maintaining the identity of the target object across frames even in challenging scenarios. The processing speed was sufficient for real-time applications, with the system capable of handling video streams at 30 frames per second. The robustness to occlusions was also notable, with the system able to re-identify the object after temporary occlusions in most cases.

TABLE I
TRACKING PERFORMANCE METRICS

| Scenario | Accuracy | Processing Speed | Occlusion Handling |
|---|---|---|---|
| Crowded Environment | 92% | 30 FPS | Good |
| Fast-Moving Object | 89% | 28 FPS | Moderate |
| Partial Occlusion | 85% | 30 FPS | Excellent |

The results in Table I summarize the system's performance across different scenarios. The system's ability to maintain high accuracy and processing speed demonstrates its suitability for real-time applications.

## V. VIDEO ANALYSIS

The video analysis illustrates the system's capability to consistently track a single object despite various challenges such as movement, scale variation, and partial occlusion. In Figure 1, the system successfully tracks a person in a crowded environment, maintaining their identity throughout the sequence. In Figure 2, the system continues to track the person with id=1 even when they are partially occluded by other objects.

These figures demonstrate the robustness of the tracking system in real-world scenarios. The system's ability to handle occlusions and re-identify the target object is particularly noteworthy.

## VI. DISCUSSION

The proposed single object tracking system leveraging YOLOv8 demonstrates high accuracy and robustness, making it suitable for various practical applications. The integration of YOLOv8 with a tracking algorithm ensures that the object is tracked reliably even in challenging conditions.

One of the key strengths of our system is its ability to handle occlusions effectively. By combining the strengths of YOLOv8's detection capabilities with a robust tracking algorithm, the system can re-identify the target object after temporary occlusions. This is crucial for applications such as surveillance and autonomous driving, where occlusions are common.

Furthermore, the system's real-time processing capability makes it suitable for applications requiring immediate feedback, such as drone navigation and interactive robotics. The high accuracy and robustness of the system also make it applicable to various domains, including sports analytics, wildlife monitoring, and security.

Future work will focus on enhancing the system's performance under more challenging conditions, such as rapid movements and complex backgrounds. We also plan to explore the application of our approach to multi-object tracking scenarios, where multiple objects need to be tracked simultaneously. Additionally, we aim to incorporate more advanced data association techniques to further improve tracking accuracy.

## VII. CONCLUSION

This paper presents a single object tracking system leveraging YOLOv8 for real-time applications. The proposed method demonstrates high accuracy and robustness, making it suitable for various practical applications. The system's ability to handle occlusions and maintain the identity of the target object across frames is particularly notable. Future work will focus on enhancing the system's performance under challenging conditions and exploring its application to multi-object tracking scenarios.

## VIII. REFERENCES


1) Redmon, J., Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767.
2) Bochkovskiy, A., Wang, C. Y., Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934.
3) Ultralytics. (2023). YOLOv8 Documentation. Retrieved from https://ultralytics.com/yolov8
4) Bewley, A., Ge, Z., Ott, L., Ramos, F., Upcroft, B. (2016). Simple online and realtime tracking. 2016 IEEE International Conference on Image Processing (ICIP).
5) Ren, S., He, K., Girshick, R., Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In Advances in Neural Information Processing Systems.
6) Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., Berg, A. C. (2016). SSD: Single Shot MultiBox Detector. In European Conference on Computer Vision.