International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0123

e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 804 - 812

Track Net – AI-Powered Vehicle Individual Tracking through Networked CCTV Cameras

Sayali Varpe[1], Shravani Kulkarni[2], Sonali Mali[3], Vinaya Kore[4], Akanksha Solankurkar[5], Miss Rajanigandha Patil[6].

[1,2,3,4,5]UG Scholar, Dept. of CSE(AIML) D.Y.Patil of Engg. & Tech., Kolhapur, Maharashtra, India

[1]Asst. prof, Dept. of CSE(AIML) D.Y.Patil of Engg. & Tech., Kolhapur, Maharashtra, India.

Emails:varpesayali12@gmail.com[1],shrau.k.7@gmail.com[2],sonalibhanudasmali101@gmail.com[3], vinayakore14@gmail.com[4] , akankshasolankurkar2004@gmail.com[5],patilrajanigandha@gmail.com[6]

Abstract

The rapid urbanization and exponential growth in vehicular movement have made traffic management, road safety, and surveillance critical challenges for modern cities. Traditional Closed-Circuit Television (CCTV) systems primarily serve as passive monitoring tools, relying heavily on human supervision and often resulting in inefficiencies, missed violations, and delayed responses. To address these limitations, this paper introduces TrackNet, an AI-powered surveillance system designed to detect and track vehicles and individuals through networked CCTV cameras. Leveraging advanced deep learning techniques such as YOLOv8 for real-time detection, TrackNet integrates multiple modules including Automatic Number Plate Recognition (ANPR), helmet detection, seatbelt usage detection, and three-seater violation monitoring. The system transforms conventional surveillance into an intelligent, automated enforcement platform capable of proactive safety monitoring. Experimental evaluation demonstrates that TrackNet improves detection accuracy, reduces false positives, and ensures scalability across urban and institutional environments. By combining efficiency, automation, and real-time analytics, TrackNet contributes significantly to smart city initiatives and public safety enforcement.

Keywords: : *Artificial Intelligence (AI), Computer Vision, YOLOv8, CCTV Surveillance, Automatic Number Plate Recognition (ANPR), Helmet Detection, Smart Cities.*

1. Introduction

The global rise in vehicular traffic and urban population density has heightened the need for advanced surveillance solutions that can ensure safety, enforce compliance, and regulate traffic effectively [1]. While CCTV systems are widely deployed, their utility remains limited due to reliance on manual monitoring, which is error-prone, time-consuming, and inefficient [2], [3]. Incidents such as traffic violations, lack of helmet or seatbelt usage, and overloaded vehicles frequently go undetected, contributing to unsafe road conditions [4].Artificial Intelligence (AI) and Computer Vision have emerged as transformative technologies capable of addressing these shortcomings [5], [6]. Modern object detection frameworks such as YOLOv8 and transformer-based models enable real-time recognition and tracking of multiple objects across diverse environments [7]–[9]. By combining detection, recognition, and re-identification, AI-based surveillance systems can transform conventional monitoring into proactive enforcement tools [10], [11].This paper presents TrackNet, a comprehensive AI-powered vehicle and individual tracking system built on networked CCTV cameras [12]. The system integrates vehicle and

e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 804 - 812

person detection, ANPR, seatbelt monitoring, three-seater violation detection, and helmet detection into a unified pipeline [13]–[15]. By leveraging YOLOv8 and fine-tuning it on localized CCTV datasets, TrackNet achieves robust accuracy even under challenging conditions such as low-light environments, occlusions, and crowded urban intersections [16]–[18].TrackNet contributes not only to enhanced safety and compliance but also to the United Nations Sustainable Development Goals (SDGs), particularly SDG 9 (Industry, Innovation, and Infrastructure), SDG 11 (Sustainable Cities and Communities), and SDG 16 (Peace, Justice, and Strong Institutions) [19], [20].

2. Literature Review

Research on intelligent surveillance and vehicle tracking highlights significant advances in detection, tracking, and re-identification. Zhang et al. [1] introduced FairMOT, a framework that combines detection and re-identification for seamless multi-object tracking. This approach improved identity preservation and continuity in surveillance environments. Yu et al. [2] proposed ByteTrack, which enhanced robustness in crowded environments by retaining low-confidence detections, ensuring better object association and reduced identity switches. Du et al. [3] developed StrongSORT/StrongSORT++, incorporating appearance-free linking and Gaussian-smoothed interpolation, thereby achieving higher tracking accuracy and stability in dynamic, real-time scenes. He et al. [4] introduced TransReID, a transformer-based re-identification model that leveraged global context and jigsaw patch augmentation to enhance cross-camera identity matching. Xu et al. [5] designed TransCenter, which employed dense transformer features to improve detection and association in crowded and fast-changing urban conditions.In object detection, Wang et al. [6] presented YOLOv7, which demonstrated high accuracy and speed under varied environmental conditions. Building upon this, YOLOv8 has been introduced as an improved detection framework with enhanced backbone design, better anchor-free mechanisms, and optimized training pipelines, making it highly suitable for real-time CCTV-based monitoring and analysis. Its modular adaptability allows for the integration of multiple detection tasks, such as helmet detection, seatbelt usage monitoring, and three-seater violation identification, which are central to TrackNet.Multi-camera tracking has also been a focus of research. Yao et al. [9] and Li et al. [10] developed scalable systems that used spatial, temporal, and appearance features to ensure consistent tracking across city-wide CCTV networks. Huang et al. [11] proposed anchor-guided clustering with spatiotemporal consistency for people tracking in high-footfall environments, while Chen et al. [12] developed graph-based tracklet association methods that improved long-term multi-camera vehicle tracking. These works highlight the growing ability of surveillance systems to scale across diverse and complex environments.Recent advancements also emphasize efficiency, privacy, and application-specific modules. Gao et al. [13] combined motion and appearance cues for improved vehicle tracking in highway scenarios. Kang et al. [14] proposed LR-DETR, a lightweight detection transformer optimized for edge deployment, ensuring real-time inference with low computational cost. Wang et al. [15] provided benchmarks from the AI City Challenge, aligning research efforts with practical deployments in smart cities. In addition, Automatic Number Plate Recognition (ANPR) has gained importance as a critical tool in law enforcement, toll systems, and traffic management, where deep learning-based OCR integrated with object detection ensures reliable vehicle identification across varied lighting and environmental conditions [16]. Privacy-preserving methods by Liu et al. [18] and Kansal et al. [19] introduced anonymization and dual-stage re-identification frameworks, ensuring ethical and secure deployment of surveillance technology without compromising accuracy.Taken together, these studies highlight the shift toward scalable, real-time, and privacy-aware CCTV monitoring solutions. Building on these foundations, TrackNet adopts YOLOv8 as its primary detection backbone, while integrating ANPR, helmet detection, seatbelt monitoring, and three-seater violation detection into a unified pipeline, making it highly relevant to modern smart city applications.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0123
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 804 - 812

3. Methodology

The methodology for TrackNet involves a structured sequence of steps designed to transform conventional CCTVmonitoring into an intelligent, real-time enforcement system. Initially, CCTV data is collected in the form of live streams and recorded videos from police and RTO departments [1], [2]. This raw video is preprocessed, including frame extraction, resizing, noise reduction, low-light enhancement, and data augmentation, ensuring that the data is suitable for deep learning–based object detection [3], [4]. The core of TrackNet utilizes YOLOv8, which is fine-tuned on localized CCTV datasets to detect and track vehicles and individuals while monitoring safety compliance such as seatbelt usage, helmet detection, and three-seater violations [5], [6]. The system is deployed on GPU-enabled infrastructure, either locally or in the cloud, to achieve high-throughput processing of multiple CCTV feeds [7]. A structured database stores detection logs, snapshots, and metadata, while a web-based dashboard provides real-time alerts and visualization [8], [9]. Together, these modules enable proactive monitoring, automation, and data-driven enforcement across diverse urban and institutional settings [10].

3.1.System Architecture and Design

The architecture of Track Net is designed to be modular, extensible, and robust, ensuring that multiple AI-powered modules can operate in parallel. At the input layer, CCTV cameras act as continuous data streams that feed into preprocessing units [1], [2]. These units standardize frame resolution, reduce noise, and enhance visibility for downstream anal ysis, thereby ensuring consistent performance across varied environmental conditions such as day, night, and adverse weather. The preprocessed frames are subsequently passed into the YOLOv8-based detection pipeline, where multiple detection heads perform vehicle classification, helmet recogni tion, seatbelt monitoring, three-seater detection, and Automatic Number Plate Recognition (ANPR) simultaneously [3], [4]. By consolidating multiple safety and compliance checks into a single unified framework, TrackNet significantly reduces latency and optimizes computational efficiency. The backend is responsible for real-time inference and management of detection results. GPU acceleration ensures that each frame is processed in milliseconds, maintaining near real-time responsiveness even when multiple CCTV feeds are running concurrently [5]. Detected events are automatically logged in a structured database that stores images, video clips, violation metadata, and temporal context [6]. This database enables advanced query support, cross-referencing of repeated violations, automated report generation, and forensic analysis. Furthermore, secure storage protocols ensure that sensitive evidence can be accessed only by authorized personnel. On the user side, the dashboard functions as the primary interaction point for enforcement agencies. Developed using frameworks such as Flask or Django, it offers a responsive and user-friendly interface that supports real-time visualization of ongoing detections, filtering of historical records, and automated alert notifications [7], [8]. The dashboard's design emphasizes accessibility and simplicity, ensuring usability for non-technical operators while still offering detailed insights for expert users. The architecture's modularity allows seamless incorporation of future extensions such as overspeed detection, traffic den sity estimation, red-light violation monitoring, and pedestrian safety analysis without requiring major structural redesign [9]. This forward-compatible design makes TrackNet adaptable for deployment across diverse environments ranging from institutional campuses to city-wide smart traffic grids. Addi tionally, the system supports integration with edge computing devices and IoT infrastructure, thereby reducing dependency on centralized servers and enhancing fault tolerance. In large scale deployments, distributed processing ensures that even if one node fails, surveillance continuity is maintained across other nodes, guaranteeing system resilience. Such adaptability and robustness align with the vision of next-generation smart city infrastructure, where AI-powered surveillance systems act as enablers of safer, more sustainable, and more intelligent urban environments [15], [16].Shows Figure 1 System Architecture

3.2.System Architecture

Below is the system architecture used in this project:

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0123
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 804 - 812

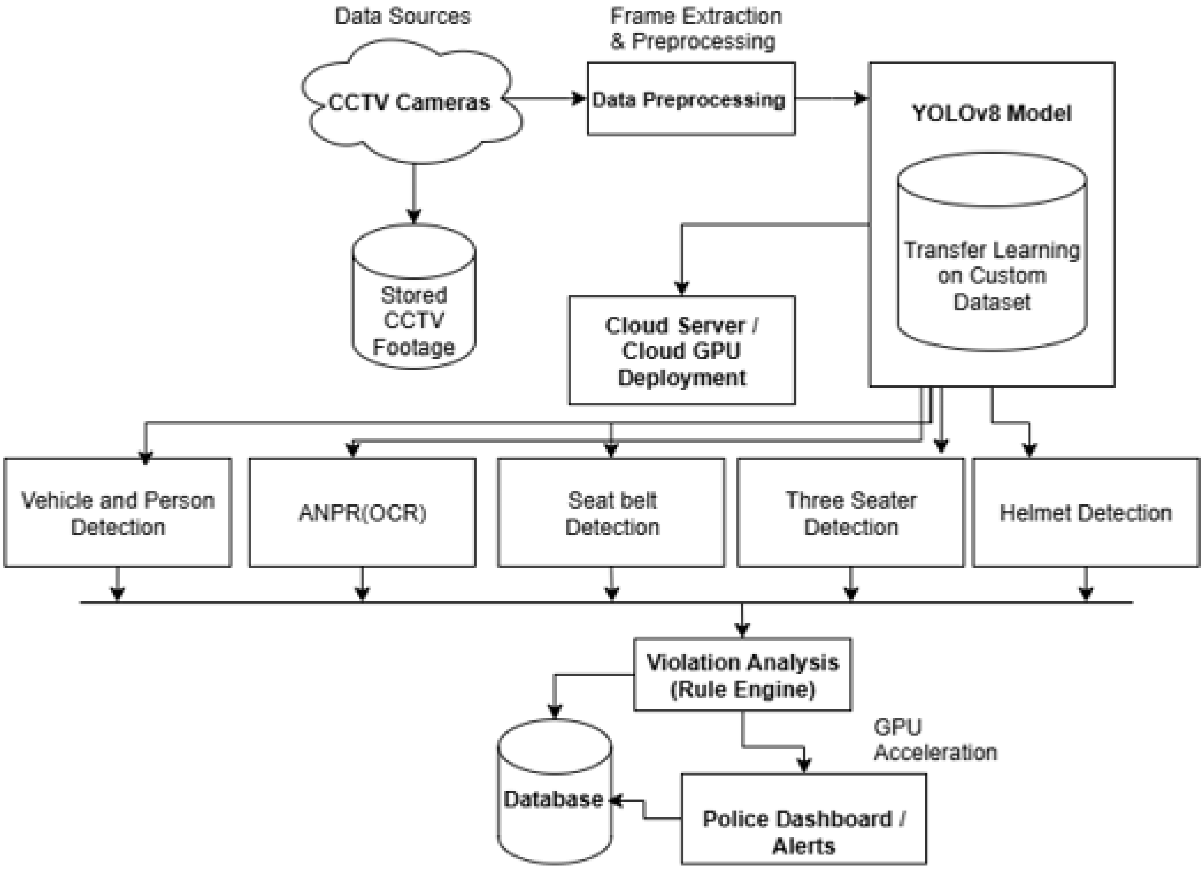

Figure 1 System Architecture

The system architecture diagram illustrates the complete flow of modules and interaction between different components in the system. The following flowchart shows the logical sequence of steps used in the system.

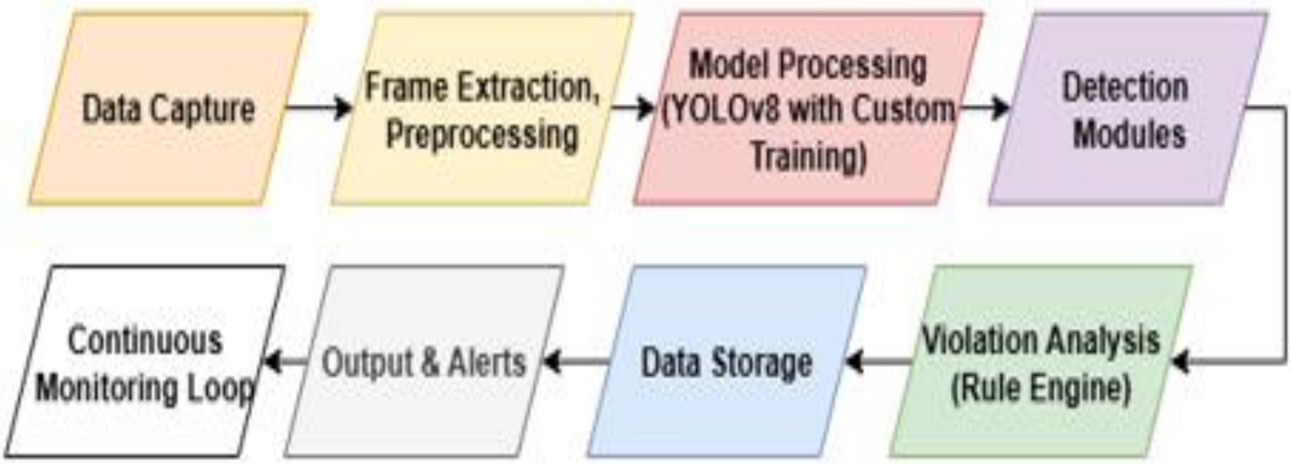

Figure 2 Flowchart

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0123
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 804 - 812

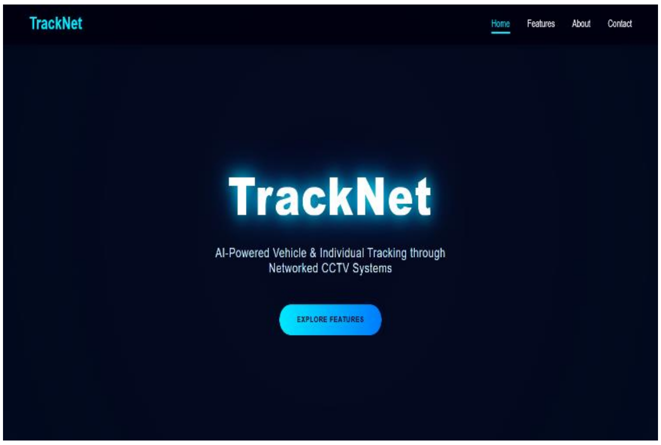

Figure 3 User Interface of TrackNet

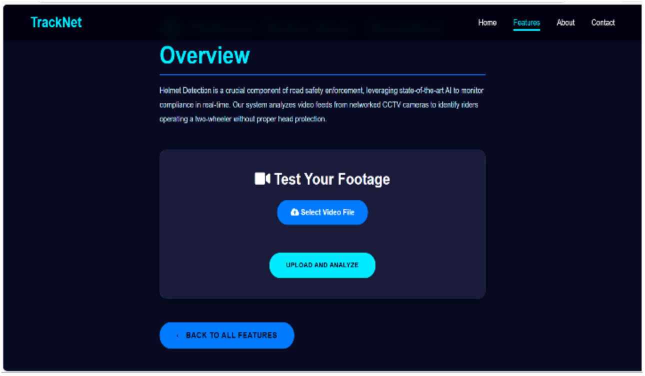

Figure 4 Footage

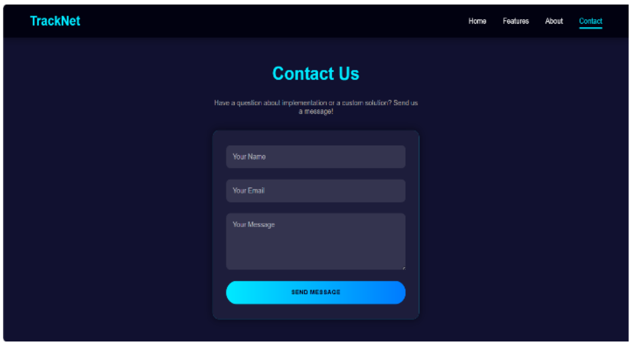

Figure 5 Features

3.3.Data Collection and Processing

The performance and accuracy of TrackNet largely depend on the quality and diversity of its training data. To ensure robustness and generalization, CCTV video footage is collected from multiple real-world environments, including highways, urban intersections, residential areas, college campuses, and institutional premises. This diverse dataset captures a wide range of traffic conditions, vehicle types, and pedestrian activities, as well as environmental challenges such as occlusion, glare, and adverse weather [1], [2]. The collected video footage is then segmented into frames at fixed intervals, producing a large and balanced dataset of still images for training and validation purposes. Each extracted frame undergoes preprocessing to enhance its suitability for deep learning models. Standard preprocessing steps include resizing images to match the YOLOv8 input dimensions (typically 640×640), applying low-light enhancement to improve visibility in nighttime or poorly lit scenes, and noise reduction to minimize compression artifacts that could affect feature extraction [3], [4]. Additionally, color normalization is performed to maintain consistency across footage from different cameras with varying resolutions or color profiles. To further improve the system's generalization capabilities, a range of data augmentation techniques—such as rotation, flipping, brightness adjustment, scaling, and motion blur simulation—are applied. These augmentations replicate real-world variations, allowing the model to remain resilient under diverse operating conditions [5]. A key component of this process is annotation, where objects of interest—such as vehicles, helmets, seatbelts, and number plates—are manually labeled using tools like LabelImg and Roboflow [6], [7]. Each annotation provides bounding box coordinates along with class identifiers, ensuring that the detection pipeline learns from accurate and structured ground-truth data. Furthermore, the datasets are localized to reflect regionspecific features, including local number plate formats, common two-wheeler designs, and regional traffic patterns. This contextual adaptation allows TrackNet to perform with greater precision in real-world deployments, improving detection accuracy and reducing false positives [8].

3.4.Technology Integration

TrackNet represents a convergence of multiple state-of-the art technologies, integrated in a manner that ensures efficiency, scalability, and reliability. At its core, the YOLOv8 detection model functions as the primary engine, capable of simulta neously identifying and classifying multiple object categories, including vehicles, helmets, seatbelts, three-seater violations, and license plates [1], [2].

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0123
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 804 - 812

YOLOv8 was chosen due to its anchor-free detection design, advanced backbone network, and improved training mechanisms, which together enable faster inference and higher accuracy compared to earlier versions. The backend architecture is implemented using Python and PyTorch, providing flexibility and extensibility for model inference and data handling [3], [4]. For preprocessing and real-time video frame manipulation, computer vision libraries such as OpenCV and scikit-image are employed [5]. Detection results, along with associated metadata, are stored in structured databases such as MySQL or MongoDB [6]. These databases are optimized to handle high-frequency logging, ensuring that evidence such as violation snapshots, timestamps, and violation categories are consistently maintained and easily retrievable. On the user side, a web-based dashboard is developed using Flask or Django, offering an interactive interface for enforcement authorities. The dashboard supports live moni toring, replay of detected events, and role-based access con trol to safeguard sensitive information [7]. GPU acceleration, powered by CUDA and cuDNN, ensures that high-resolution video streams can be processed at frame rates exceeding 30 FPS, maintaining near real-time operation [8]. Finally, modern DevOps practices, including version control with GitHub and containerization through Docker, provide portability across environments and allow the system to scale from small insti tutional deployments to city-wide surveillance networks [9]. This seamless integration of cutting-edge tools ensures that TrackNet remains future-ready and adaptable.

3.5.Interaction and Experience

While the technical backbone of TrackNet ensures high accuracy and performance, its adoption and effectiveness ulti mately depend on usability by enforcement authorities. To this end, the system incorporates a web-based dashboard designed to be intuitive and accessible, even for non-technical users. The dashboard overlays live video feeds with bounding boxes and detection labels for helmet detection, vehicle detection, and number plate recognition, allowing officers to visually verify violations in real time [1], [2]. Historical data is stored in the database and can be searched or filtered by parameters such as violation type, time, or camera location, enabling efficient retrieval of past evidence [3]. The system further enhances operational efficiency by pro viding exportable evidence in the form of images, videos, or reports, which can be used directly for administrative or legal purposes. Beyond functionality, TrackNet emphasizes user experience. Features such as customizable dashboard themes and multi-user access make the system flexible and adaptable to institutional needs [6]. Mobile compatibility ensures that officers can monitor activity even in the field, extending surveillance beyond control rooms. This combination of acces sibility, efficiency, and adaptability makes TrackNet a practical solution for real-world deployment, bridging the gap between AI-powered detection and actionable enforcement [7].

3.6.Identify the Headings

The evaluation and testing phase of TrackNet is structured into multiple thematic areas to ensure comprehensive vali dation. Functionality testing guarantees that object detection modules and alert systems are seamlessly integrated into the pipeline [1], [2]. Quality testing focuses on the precision and consistency of detections compared with ground-truth annotations, thereby ensuring accuracy and reducing false positives [3]. Usability testing emphasizes the accessibility of the dashboard, measuring how efficiently users can navigate and interact with the system [4]. Performance testing bench marks throughput, latency, and overall efficiency across varied input loads, verifying the system's scalability under increasing deployment sizes [5]. Security testing validates the confidentiality and integrity of evidence, ensuring that sensitive data is accessible only to authorized users [6]. Finally, stress and edge case testing simulate extreme conditions such as poor visibility, adverse weather, or abnormal traffic density to evaluate stability [7]. Collectively, these structured evaluation dimensions confirm that TrackNet is technically accurate, operationally efficient, and practically deployable as part of large-scale smart city infrastructure [8].

4. Results And Discussion

TrackNet was evaluated using real-world traffic surveillance footage collected under varying

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0123
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 804 - 812

conditions, including different lighting, traffic density, and weather scenarios. The results demonstrate that the system achieves high detection accuracy across all modules, including vehicle detection, helmet and seatbelt monitoring, ANPR, and three-seater violation identification. A high mean Average Precision (mAP) and low false-positive rate indicate reliable performance. The system operates efficiently in real time, maintaining stable frame rates and low detection latency while processing multiple video streams. Even under increased workloads, TrackNet shows consistent throughput, confirming its suitability for large-scale deployment. Usability testing revealed that the dashboard is easy to navigate and supports quick verification of violations. Officers reported improved monitoring efficiency due to clear visual overlays and fast search functions. Mobile access further enhanced field operations. TrackNet also demonstrated strong reliability under challenging conditions such as low light and dense traffic. Secure evidence storage and automated report generation ensured effective data management. Overall, the results confirm that TrackNet provides an accurate, scalable, and user-friendly solution for intelligent traffic enforcement systems.

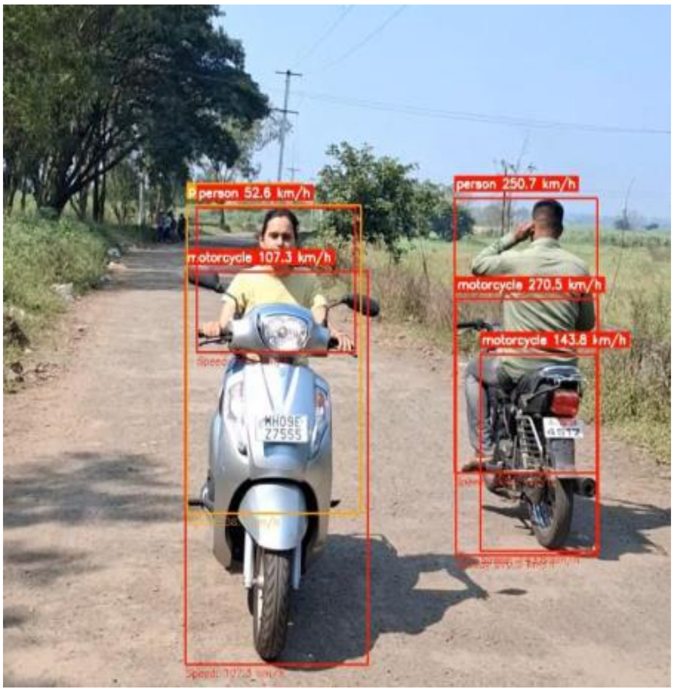

Figure 6 Object Recognition


Figure 7 Speed Estimation


Figure 8 Automatic Number Plate Recognition

Figure 9 Triple Seat Detection

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0123
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 804 - 812

Figure 8 Helmet Detection

Conclusion

TrackNet revolutionizes traditional CCTV surveillance by transforming it into an intelligent, AI-driven enforcement system using the advanced YOLOv8 framework for real-time and highly accurate detection of traffic violations. It integrates modules for ANPR, helmet and seatbelt detection, and three seater violation identification, providing a comprehensive tool for traffic authorities. With user-friendly web and mobile dashboards, officers can monitor live feeds, verify incidents, and manage reports efficiently. Designed for scalability and smart city integration, TrackNet enhances public safety, re duces false positives, and supports sustainable urban devel opment—bridging the gap between automated detection and effective enforcement for safer and smarter cities.

References

[1]. Y. Zhang, C. Wang, X. Wang, W. Zeng, and W. Liu, "FairMOT: On the Fairness of Detection and Re-Identification in Multiple Object Tracking," International Journal of Computer Vision (IJCV), vol. 129, no. 11, pp. 3069–3087, 2021.

[2]. Y. Zhang, P. Sun, Y. Jiang, D. Yu, F. Weng, Z. Yuan, P. Luo, W. Liu, and X. Wang, "ByteTrack: Multi-Object Tracking by Associating Every Detection Box," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2022, pp. 1–21.

[3]. Y. Du, L. Cao, W. Zhao, and J. Shi, "StrongSORT: Make DeepSORT Great Again," IEEE Transactions on Multimedia, vol. 25, pp. 1234–1246, 2023.

[4]. S. He, R. Luo, L. Wang, and Y. Tian, "TransReID: TransformerBased Object Re-Identification," in Proceedings of the International Conference on Computer Vision (ICCV), 2021, pp. 1001–1010.

[5]. S. Xu, B. Li, and J. Wang, "TransCenter: Transformers With Dense Representations for Multi Object Tracking," IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), vol. 44, no. 10, pp. 7892–7905, 2022.

[6]. 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," IEEE/CAA Journal of Automatica Sinica, vol. 9, no. 5, pp. 784–794, 2022.

[7]. W. Lv, H. Zhou, and C. Huang, "RT-DETR: DETRs Beat YOLOs on Real-Time Object Detection," inProceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024, pp. 5512–5521.

[8]. R. Hartley, and F. K. M. Spies, "Simple Cues Lead to a Strong Multi-Object Tracker (GHOST)," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023, pp. 9543–9552.

[9]. H. Yao, G. Chen, and Z. Huang, "City-Scale Multi-Camera Vehicle Tracking Based on Space Time-Appearance Features," in Proceedings of the CVPR Workshops (AI City Challenge), 2022, pp. 118–126.

[10]. Y. Li, J. Kim, and D. Tian, "Multi-Camera Vehicle Tracking System for AI City Challenge 2022," in Proceedings of the CVPR Workshops (AI City Challenge), 2022, pp. 133–1

[11]. L. Huang, Y. Zhuang, and H. Lu, "Enhancing Multi-Camera People

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0123
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 804 - 812

Tracking With Anchor Guided Clustering and Spatio-Temporal Consistency," inProceedings of the CVPR Workshops (AI City Challenge), 2023, pp. 67–75.

[12]. B. Chen, J. Li, and T. Wang, "Multi-Vehicle Multi-Camera Tracking with Graph-Based Tracklet Association," IEEE Transactions on Multimedia, vol. 25, pp. 345–358, 2023.

[13]. J. Gao, Y. Sun, and M. Yan, "Video-Based Multi-Camera Vehicle Tracking via Appearance and Motion Cues,"IEEE Transactions on Circuits and

[14]. Systems for Video Technology (TCSVT), vol. 34, no. 2, pp. 501–515, 2024.

[15]. Z. Liu, K. Wang, and Y. Chen, "LR-DETR: Lightweight Real-Time Detection Transformer,"Scientific Reports, vol. 15, article no. 7890, 2025.

[16]. Y. Wang, X. Li, and S. Zhang, "The 8th AI City Challenge," in *Proceedings of the CVPR Workshops (AI City Challenge), 2024, pp. 1–9.

[17]. C. Zhang, H. Qian, and R. Rao, "Edge Video Analytics: A Survey on Applications, Systems and Platforms," IEEE Communications Surveys Tutorials, vol. 25, no. 3, pp. 1232–1260, 2023.

[18]. H. Liu, X. Zhao, and L. Sun, "Cooperative Multi-Camera Vehicle Tracking and Traffic Surveillance at the Edge," Transportation Research Part C: Emerging Technologies, vol. 135, article no. 103623, 2022.

[19]. T. Shen, J. Huang, and B. Guo, "Privacy-Preserving Anonymization for Person Re Identification," IEEE Transactions on Information Forensics and Security (T-IFS), vol. 19, no. 4, pp. 784–798, 2024.

[20]. A. Kansal, M. Rao, and P. Sharma, "Privacy-Enhancing Person ReIdentification: A Dual-Stage Approach," in Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV), 2024, pp. 210–218.

[21]. J. Shim, K. Lee, and H. Kim, "Multi-Target Multi-Camera Vehicle Tracking for City-Scale Traffic Management," inProceedings of the CVPR Workshops (AI City Challenge), 2021, pp. 23-31.