



# Multi-Vehicle Tracking and Counting in Traffic Surveillance: A Survey

Bhaskar<sup>1</sup>, Preeti<sup>2</sup>

P.G. Student, Department of CSE, Sat Kabir Institute of Technology and Management, Haryana, India<sup>1</sup>

Assistant Professor, Sat Kabir Institute of Technology and Management, Haryana, India<sup>2</sup>

**ABSTRACT:** Rapid urbanization and the surge in vehicles have really bumped up the need for smart traffic systems. These systems help monitor road traffic, analyze data, and make traffic management easier. At the heart of Intelligent Transportation Systems (ITS) are vehicle tracking and counting abilities. They play a big role in analyzing traffic flow, addressing congestion, preventing accidents, and developing plans for smart cities. This survey explores methods for detecting, tracking, and counting vehicles in surveillance setups. It looks at a range of tech – background subtraction, optical flow, feature detection, object segmentation, and deep learning strategies. The review also delves into tracking algorithms such as Kalman and Particle Filters, Mean Shift and CamShift, and systems for tracking many objects simultaneously. Finally, it touches on recent gains in machine learning and deep nets and what these mean for vehicle tracking, both good and bad. This essay examines key challenges, including occlusion, changing lighting, shadows, scale changes, and real-time processing requirements. It also reviews how well these systems work by checking accuracy, precision, recall, vehicle counting, and efficiency. In the end, it highlights current research, major hurdles, and where future studies might lead in improving these traffic surveillance systems and making them more reliable.

**KEYWORDS:** Multi-Vehicle Tracking, Vehicle Counting, Traffic Surveillance

## I. INTRODUCTION

Rapid urbanization, accompanied by the increase in the number of vehicles, has led to many problems in traffic management, road safety, and transportation planning. To cope with such problems, intelligent transportation systems (ITS) can be used for monitoring and management of traffic. Multi-vehicle tracking and counting constitute important aspects within the ITS system due to their applications in traffic flow analysis, monitoring congestion, detection of traffic incidents, planning of infrastructure, and smart city development. Automated vehicle surveillance systems have attracted significant attention because they reduce human-related errors and improve cost and scalability.

The fields of computer vision and image processing have greatly advanced the development of automatic object detection and tracking systems. Shrivastava et al. [1] designed and analyzed the use of a Mean Shift-based object-tracking software on videos, demonstrating the usefulness of object-tracking algorithms of tracking object movement in video images. In another study, Shrivastava et al. [2] detailed the various methods employed in video-based object tracking technology and stressed the significance of object detection and tracking in such systems.

Methods for object detection and tracking have also been successfully applied of human tracking. The approach of Nalwa et al. [3] involved developing a computer vision-based people counting system using a webcam and proved the feasibility of computer vision-based systems in automated counting tasks. In addition to this, Nalwa et al. [4] performed an analysis of various people tracking systems under varying environmental settings. Occlusion, change in illumination levels, and object interaction were among the major challenges posed by such settings. Similar problems arise when developing vehicle tracking systems in an urban environment.

In the field of traffic surveillance, remarkable progress has been made regarding the issues of multi-vehicle detection, tracking, and counting. Specifically, Wang et al. [5] developed a multi-vehicle tracking and counting system for traffic intersections based on fisheye camera footage that operates in real time. Their system proves the viability of real-time multi-vehicle tracking and counting even in complicated traffic situations and emphasizes the significance of robust tracking algorithms for the preservation of vehicle identity. Moreover, in more recent work by Bihanda et al. [6], an approach to multi-vehicle tracking and counting based on RT-DETR and ByteTrack algorithms was presented for

average daily traffic surveys. Although many technologies have enabled vehicle tracking and counting, some challenges remain. Variables such as vehicle occlusions, shadow effects, changing weather, changing camera angles, lighting conditions, and dense traffic all impact performance. Moreover, since traffic analysis must be performed in real time, more efficient algorithms are required to process large volumes of data.

The present survey reviews various approaches to multi-vehicle tracking and counting methods used in traffic surveillance systems. The work will analyze existing approaches to multi-vehicle tracking and counting, including conventional image processing techniques, classical tracking algorithms, and modern deep learning methods. Moreover, the survey will explore the capabilities of current solutions, their strengths and weaknesses, and the performance of the existing technologies. The goal of this survey is to give researchers and professionals an insight into current trends in developing multi-vehicle tracking and counting systems.

## II. RESEARCH BACKGROUND

Vehicle tracking and counting have now become important elements of Intelligent Transportation Systems (ITS), facilitating effective traffic surveillance, traffic analysis, road safety management, and urban planning. With the rapid growth of vehicle traffic worldwide, conducting manual traffic surveys has become difficult. Therefore, efforts have been made to develop computer vision algorithms to detect, track, and count vehicles. Earlier studies in this field mainly focused on the detection and tracking of vehicles using image processing. Huang [10] has proposed an approach for real-time detection and sub-feature-based tracking of multiple vehicles for traffic surveillance. It was shown in the research that detection and tracking methods can be effectively utilized to keep track of moving vehicles irrespective of the traffic scenarios.

With the evolution of video surveillance techniques, scientists began implementing methods of detecting and tracking of vehicles to improve counting accuracy. Xiang et al. [7] developed a system for vehicle counting, which involved detecting and tracking vehicles in aerial videos. In their work, Xiang et al. used aerial images to estimate traffic and concluded that the combination of both techniques boosts the efficiency of vehicle counting. As camera technology and video analytics became increasingly sophisticated, new approaches to tracking emerged. A new methodology was presented by Meng et al. [12], which involved the use of a video-based counting method for vehicles in expressways based on vehicle detection and correlation-matched tracking via Pan-Tilt-Zoom (PTZ) cameras. Such a solution proved effective in handling issues related to vehicle movement, camera movement, and viewpoint differences. The findings confirmed that counting vehicles was possible when both detection and tracking were used.

The greater accessibility of machine learning/deep learning technology had another impact on vehicle tracking and counting methods. In their comparative study of vehicle counting using object detection and tracking, Mandal & Adu-Gyamfi [11] showed that both methods, independently or in combination, achieve good results for detecting cars. They have pointed out the criticality of the choice of a particular tracking method, especially considering the difficulties related to occlusions and the overlap of objects in motion. Current research efforts have focused on enhancing multi-object tracking by improving feature extraction and association techniques. Specifically, Nguyen et al. [9] developed an adaptive multi-vehicle motion-counting technique that can cope with various vehicle motions encountered in real traffic situations. The adaptive mechanism helped in achieving higher counts because of efficient management of vehicle interaction and motion behavior. Another contribution is made by Zhonglin et al. [8], who introduced the concept of SFFSORT (Shallow Feature Fusion Simple Online and Realtime Tracking). The method utilizes the idea of shallow feature fusion with multi-object tracking for better performance of object counting in crowds.

There is a definite trend of progression from conventional image processing approaches towards highly intelligent approaches involving deep learning and multi-object tracking. However, despite numerous technological advances in this area, researchers still need to address issues such as handling occlusions, varying lighting and weather conditions, scaling challenges, and the need for real-time processing. The studies discussed in this section provide the research foundation for understanding the development of vehicle detection, tracking, and counting technologies. They highlight the significance of integrating object detection, tracking algorithms, and intelligent data association techniques to achieve reliable traffic surveillance and support the growing demands of intelligent transportation systems.

### III. BACKGROUND SUBTRACTION METHODS FOR VEHICLE DETECTION AND TRACKING

The Background Subtraction technique is the most popular method for detecting vehicle movements in traffic surveillance systems. The basic concept involves the separation of vehicles (moving objects) from the static background using the difference between the current video scene and a background template. Over the years, many background-subtraction algorithms have been developed.

- 1. Frame Differencing Method:** Frame differencing is one of the most straightforward methods of subtracting backgrounds from an image sequence. Moving objects can be detected by comparing consecutive frames. If the difference exceeds a predetermined threshold, the pixel is considered part of a moving object. The algorithm is simple and effective for real-time processing. On the other hand, frame differencing is ineffective for detecting slowly moving objects such as vehicles. Moreover, it usually generates noisy images due to missing motion information.
- 2. Running Average Background Mode:** The running average method updates the background model continuously by averaging pixel values over time. The background image is updated according to the following relationship:
$$B_t = \alpha I_t + (1 - \alpha)B_t - 1 \quad (1)$$
where  $B_t$  represents the updated background, it is the current frame, and  $\alpha$  is the learning rate. This method can adapt to gradual changes in illumination and environmental variations. However, sudden changes in lighting or stationary vehicles may adversely affect its performance [12].
- 3. Median Filtering Background Model:** In this approach, the background is generated by calculating the median value of each pixel across a sequence of video frames. Since moving vehicles occupy a pixel location for only a short duration, the median operation effectively removes foreground objects and produces a clean background model. The method is robust against noise and temporary object movements but requires significant memory and computational resources when processing long video sequences [7].
- 4. Gaussian Background Model:** According to the Gaussian Background Model, the intensity values of each pixel can be assumed to have a Gaussian (or normal) distribution. A pixel is considered to belong to the foreground when the observed intensity differs significantly from the calculated mean. This approach offers more robust performance when illumination changes than frame differencing.
- 5. Gaussian Mixture Model (GMM):** The Gaussian mixture model (GMM) is among the popular background subtraction methods in the domain of vehicle surveillance systems. Unlike other methods that consider only one type of Gaussian distribution for a particular pixel, GMM considers various types of Gaussian distributions corresponding to the various conditions in the background. As a result, GMM is capable of managing dynamic backgrounds, including moving trees, shadows, rain, etc. Since the Gaussian mixture model works well in dynamic environments, it has been widely applied to vehicle detection and tracking.
- 6. Codebook-Based Background Subtraction:** This method uses codewords to model the background, which consists of various pixel intensity values over time. Pixels that cannot be described by any of the codewords in the system are considered the foreground pixels. This technique works very well for lighting variation, repeated motion, and shadows. Nevertheless, the training process might prove to be very challenging, especially in extensive surveillance networks [9].
- 7. ViBe (Visual Background Extractor):** The ViBe algorithm is a pixel-wise background subtraction technique that keeps track of a set of previously observed pixel values at each pixel. If a pixel value matches a set of saved pixel values, then it belongs to the background region. ViBe adapts its background estimation model at random, allowing it to adjust when the environment changes quickly. This algorithm performs well for vehicle detection due of its low computational complexity [8].
- 8. Deep Learning-Based Background Subtraction:** Emerging advances in deep learning have been responsible for the creation of background subtraction algorithms utilizing neural networks. The CNN algorithm can identify complex background structures and is capable of differentiating between vehicles in motion and shadows/reflections with greater accuracy compared to other conventional methods. Deep learning algorithms are effective in tough traffic conditions but require large labeled datasets [6].

Table 1: Comparison of Background Subtraction Methods

Method	Advantages	Limitations
Frame Differencing	Simple, fast, low computation	Poor detection of slow-moving vehicles
Running Average	Adapts to gradual changes	Sensitive to sudden illumination changes
Median Filtering	Robust to noise	High memory requirements
Gaussian Model	Handles moderate illumination variations	Less effective in dynamic backgrounds
Gaussian Mixture Model (GMM)	Robust to dynamic environments	Higher computational cost
Codebook Model	Handles shadows and illumination changes	Complex training process
ViBe	Fast adaptation, real-time performance	Sensitive to parameter selection
Deep Learning-Based Methods	High accuracy and robustness	Requires large datasets and powerful hardware

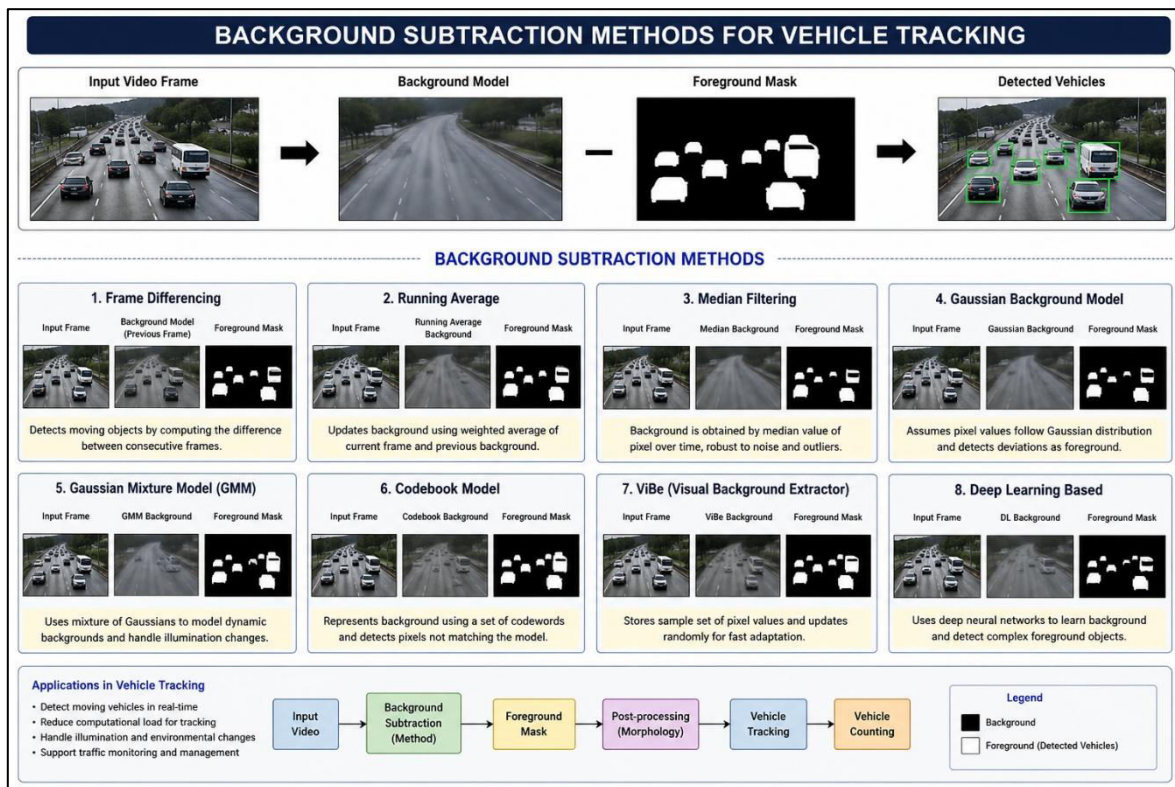


Figure 1: Background Subtraction Method

### Vehicle Tracking Algorithms

**1. Kalman Filter:** The Kalman Filter is one of the most popular model-based trackers for the surveillance of vehicles. It is an algorithm that estimates a moving vehicle's future position and velocity, as well as its trajectory of a moving vehicle based on past states and measurements. The Kalman Filter has two phases: first, the future state of the vehicle is predicted using the motion model. Second, prediction is refined by incorporating real data collected in the current frame. Due to their high processing efficiency, Kalman Filters are excellent for tracking vehicles in real time. However, Kalman Filters may perform poorly in tracking highly nonlinear vehicle movements or trajectory changes. According to Wang et al. [13], Kalman filters successfully tracked multiple vehicles at traffic junctions in real time.

2. **Particle Filter:** The Particle Filter is a probabilistic tracking technique developed to overcome the limitations of Kalman Filters in nonlinear and non-Gaussian environments. Instead of representing a vehicle's state using a single estimate, the Particle Filter maintains a collection of weighted particles that represent multiple possible states of the vehicle. As new observations become available, the particles are updated and resampled based on their likelihoods. This enables the algorithm to accurately track vehicles under complex traffic conditions involving occlusions, abrupt motion changes, and nonlinear trajectories. Although Particle Filters generally provide better tracking accuracy than Kalman Filters, they require greater computational resources due to the large number of particles involved in the estimation process. According to Mandal and Adu-Gyamfi [14], Particle Filters have shown superior performance in challenging traffic surveillance environments where conventional linear tracking models often fail.
3. **Mean Shift:** The Mean Shift algorithm is a non-parametric algorithm that iteratively estimates the location of the target object by determining the maximum probability density in the feature space. Color histograms or texture features can be used to characterize the target vehicle. During tracking, the search window is continuously moved towards regions with high feature distribution similarity of the target model. This approach is computationally inexpensive and effective for vehicle tracking in scenarios where appearance characteristics remain stable, and the vehicle undergoes moderate motion. However, the algorithm uses a static search window and is prone to tracking failures when the target vehicle exhibits significant variations in scale, orientation, and shape. Shrivastava et al. [3] effectively used mean-shift-based object tracking in their video surveillance applications.
4. **CamShift (Continuously Adaptive Mean Shift):** The CamShift algorithm is a more advanced approach to the Mean Shift algorithm that is developed to address the drawback of a static tracking window. In contrast to the Mean Shift method, CamShift tracks the object by adjusting the shape, location, and orientation of the search window in response of the target object's appearance. Therefore, this feature enables CamShift to apply to vehicle tracking tasks, since the scale of the object changes as it approaches or recedes from the camera. Thanks to these characteristics, CamShift can be effectively used in traffic-tracking applications. According to Shrivastava et al., CamShift was among the most efficient adaptive tracking methods [4].
5. **Multi-Object Tracking (MOT):** Multi-Object Tracking (MOT) is defined as the ability to simultaneously track several vehicles without losing unique identifiers of each object within the video stream. In general, MOT frameworks consist of the integration of object detection, motion prediction, appearance modeling, and data association procedures. Classical MOT algorithms relied on the use of Kalman Filters and the Hungarian matching algorithm; however, more advanced methods use of deep learning models to enhance tracking precision. The most common algorithms used in multi-object tracking include SORT (Simple Online Realtime Tracking), DeepSORT, ByteTrack, and SFFSORT. Specifically, DeepSORT uses an appearance model trained with deep learning algorithms, ByteTrack performs data association via confidence-based matching, and SFFSORT uses shallow feature fusion. These MOT algorithms can effectively address the challenges of dense traffic areas, characterized by frequent vehicle interactions, occlusions, and overlaps among moving objects. As shown in recent studies performed by Zhonglin et al. [8] and Bihanda et al. [6], the application of MOT frameworks has greatly enhanced the accuracy of vehicle counting and tracking tasks in practice.

Table 2: Comparison of Vehicle Tracking Methods

Method	Working Principle	Advantages	Limitations	Computational Complexity	Suitable Traffic Conditions	Example Applications
<b>Kalman Filter</b>	Predicts vehicle position using previous state information and updates the prediction using new measurements.	Fast, computationally efficient, suitable for real-time tracking, and handles noisy measurements well.	Assumes linear motion and Gaussian noise; performance decreases for abrupt or nonlinear movements.	Low	Moderate traffic, predictable vehicle motion	Traffic surveillance, vehicle trajectory prediction, and real-time tracking
<b>Particle Filter</b>	It represents the vehicle state using multiple weighted	Handles nonlinear motion and non-Gaussian noise,	High computational cost, requires many particles	High	Complex and crowded traffic environments	Autonomous vehicles, advanced surveillance

	particles and estimates the most probable state through resampling.	robust to occlusion and sudden motion changes.	for accurate estimation.			systems
<b>Mean Shift</b>	Tracks objects by iteratively shifting the search window toward the highest probability density region in feature space.	Simple implementation, low computational cost, effective for stable object appearance.	Fixed window size, sensitive to scale and shape variations, may lose the target during occlusion.	Low	Low to moderate traffic density	Basic object tracking, surveillance applications
<b>CamShift</b>	An extension of Mean Shift that dynamically adjusts the search window size and orientation during tracking.	Adapts to scale changes, efficient for real-time applications, better than Mean Shift for moving objects.	Performance is affected by changes in illumination and significant occlusions.	Low to Medium	Moderate traffic with varying vehicle sizes	Vehicle tracking, face tracking, video surveillance
<b>SORT (Simple Online Realtime Tracking)</b>	It combines the Kalman Filter prediction with the Hungarian Algorithm for data association.	Very fast, suitable for real-time applications, and easy to implement.	Relies heavily on detection quality, poor under long occlusions.	Low	Moderate traffic environments	Real-time traffic monitoring
<b>DeepSORT</b>	Enhances SORT by incorporating deep appearance features for object re-identification.	Better identity preservation, handles short-term occlusions effectively.	Higher computational requirements than SORT.	Medium	Dense urban traffic	Vehicle and pedestrian tracking
<b>ByteTrack</b>	Associates both high-confidence and low-confidence detections to improve object matching.	High tracking accuracy, robust under crowded conditions, and partial occlusions.	Requires a powerful object detector and parameter tuning.	Medium to High	Dense and complex traffic scenes	Multi-vehicle tracking and counting
<b>SFFSORT</b>	Uses shallow feature fusion with online tracking to improve object association and tracking robustness.	Improved tracking accuracy, effective in crowded environments, and enhanced counting performance.	More computationally intensive than traditional tracking methods.	Medium to High	Heavy traffic and large-scale surveillance systems	Traffic analysis, vehicle counting

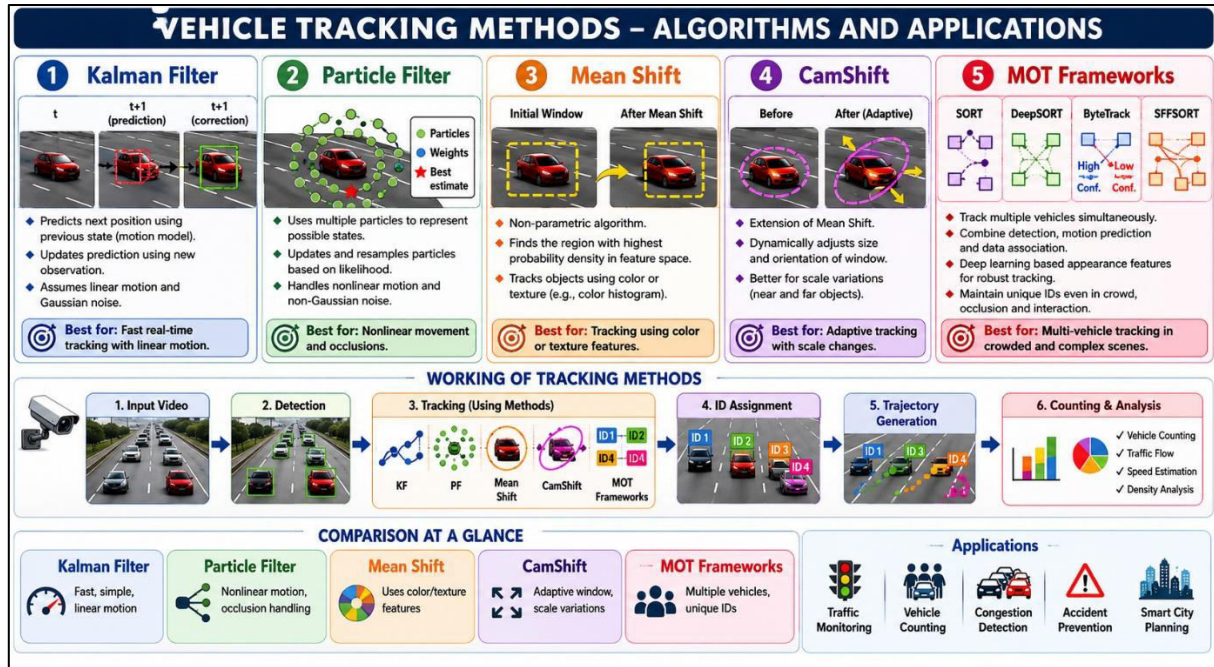


Figure 2: Vehicle Tracking Methods

#### IV. CONCLUSION

Vehicle tracking and vehicle counting have emerged as critical technologies in current ITS, enabling applications such as traffic management, congestion mitigation, accident prevention, and smart cities. In this review article, the main methods used for vehicle detection, tracking, and counting in traffic surveillance systems were thoroughly explored. Traditional methods relying on background subtraction, optical flow, feature extraction, and object segmentation laid down the base for the development of advanced traffic analysis techniques. Different tracking algorithms were described here, such as the Kalman Filter, the Particle Filter, the Mean Shift, and the CamShift. The survey also focused on the evolution of Multi-Object Tracking (MOT) algorithms, including SORT, DeepSORT, ByteTrack, and SFSORT. With motion prediction, appearance models, and deep learning approaches, the latest MOT algorithms are known to deliver accurate results in terms of robust object identity maintenance and vehicle counting in complex scenarios with numerous objects in the scene. It is also important to note that while conventional tracking algorithms are efficient and suitable for real-time tracking, deep learning algorithms outperform them in the presence of occlusions. Although notable progress has been made, several issues still affect the performance of vehicle tracking systems, including illumination differences, weather conditions, shadows, changes in camera viewpoint, occlusion, and real-time computations. The future work will involve implementing deep neural networks, transformer-based object detectors, sensor fusion, edge computing, and artificial intelligence-based tracking algorithms. In general, improvements in vehicle-tracking and counting technologies will be essential for the efficient operation of future smart traffic management systems.

#### REFERENCES

- [1] D. K. Shrivastava, K. Bhatia, Shivkant, R. Sharma, "Development and analysis of the mean shift-based Video object tracking tool," *International Journal of Innovative Research in Computer and Communication Engineering*. Vol-08, Issue-07, July 2020.
- [2] D. K. Shrivastava, K. Bhatia, Shivkant, R. Sharma, "An overview of video-based object Tracing Techniques," *International Research Journal of Modernization in Engineering Technology and Science*, Vol. 02, Issue 08, August 2020, pp. 316-321.

- [3] Nalwa, N., Bhadola, S., Bhatia, K. and Sharma, R., 2022. Development of a Vision-Based People Counting Framework using a Webcam. *International Journal Of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)*, 5(7), pp.1665-1671.
- [4] Nalwa, N., Bhadola, S., Bhatia, K. and Sharma, R., 2022. A Detailed Study on People Tracking Methodologies in Different Scenarios. *International Journal of Innovative Research in Computer and Communication Engineering*, 10(6), pp.5907-5913.
- [5] Wang, W., Gee, T., Price, J., and Qi, H., 2015, January. Real-time multi-vehicle tracking and counting at intersections from a fisheye camera. In 2015, *IEEE Winter Conference on Applications of Computer Vision* (pp. 17-24). IEEE.
- [6] Bihanda, Y.G., Faticah, C., and Yuniarti, A., 2024. Multi-vehicle tracking and counting framework in average daily traffic survey using RT-DETR and BYTETRACK. *IEEE Access*, 12, pp.121723-121737.
- [7] Xiang, X., Zhai, M., Lv, N. and El Saddik, A., 2018. Vehicle counting based on vehicle detection and tracking from aerial videos. *Sensors*, 18(8), p.2560.
- [8] Zhonglin, T., Ab Wahab, M.N., Akbar, M.F., Mohamed, A.S.A., Noor, M.H.M., and Rosdi, B.A., 2023. SFFSORT multi-object tracking by shallow feature fusion for vehicle counting. *IEEE Access*, 11, pp.76827-76841.
- [9] Nguyen, X.D., Vu, A.K.N., Nguyen, T.D., Phan, N., Dinh, B.D.D., Nguyen, N.D., Nguyen, T.V., Nguyen, V.T., and Le, D.D., 2022. Adaptive multi-vehicle motion counting. *Signal, Image and Video Processing*, 16(8), pp.2193-2201.
- [10] Huang, L., 2010, March. Real-time multi-vehicle detection and sub-feature-based tracking for traffic surveillance systems. In 2010, the 2nd International Asia Conference on Informatics in Control, Automation and Robotics (CAR 2010) (Vol. 2, pp. 324-328). IEEE.
- [11] Mandal, V. and Adu-Gyamfi, Y., 2020. Object detection and tracking algorithms for vehicle counting: a comparative analysis. *Journal of Big Data Analytics in Transportation*, 2(3), pp.251-261.
- [12] Meng, Q., Song, H., Zhang, Y.A., Zhang, X., Li, G. and Yang, Y., 2020. Video-Based Vehicle Counting for Expressway: A Novel Approach Based on Vehicle Detection and Correlation-Matched Tracking Using Image Data from PTZ Cameras. *Mathematical Problems in Engineering*, 2020(1), p.1969408.
- [13] Wang, W., Gee, T., Price, J., & Qi, H. (2015). Real-time multi-vehicle tracking and counting at intersections from a fisheye camera. *IEEE Winter Conference on Applications of Computer Vision*, pp. 17–24.
- [14] Mandal, V., & Adu-Gyamfi, Y. (2020). Object detection and tracking algorithms for vehicle counting: A comparative analysis. *Journal of Big Data Analytics in Transportation*, 2(3), 251–261.



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details