**Journal of Intelligent Systems and Internet of Things**
**Vol. 13, No. 02, PP. 303-323, 2024**
DOI: https://doi.org/10.54216/JISIoT.130224

# A Comprehensive Review of Real-Time Vehicle Tracking for Smart Navigation Systems

**Veena R S[1], Seema Rani[2,*], Ch Madhava Rao[3], Piyush Kumar Pareek[4], Sandeep Dalal[2], Shweta Bansal[5]**

[1]Department of Information Science and Engineering, Dayananda Sagar Academy of Technology and Management, Bengaluru -560082, Karnataka, Orchid - 0000-0003-1368-0634, India
[2]Department of Computer Science & Applications, Maharshi Dayanand University, Rohtak, Haryana, India
[3]Associate Professor, Dept. of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India
[4]Professor and Head Department of AIML and IPR Cell Nitte Meenakshi Institute of Technology Bengaluru , Karnataka, India
[5]K R Managalam University, Gurugram, Haryana-122103, India

Emails: veena-ise@dsatm.edu.in; seema.rs.dcsa@mdurohtak.ac.in; cmadhavarao@kluniversity.in; piyush.kumar@nmit.ac.in sandeepdalal.80@gmail.com; S.bansal6281@gmail.com

**Abstract**

Vehicle tracking is one of computer vision's most important applications, with applications ranging from robotics and traffic monitoring to autonomous vehicle navigation and many more. Even with the significant advancements in recent research, issues like occlusion, fluctuating illumination, and fast motion still need to be addressed, calling for more investigation and creativity in this field. This study performs a thorough examination of various vehicle-tracking approaches and suggests a thorough classification scheme that divides them into four main categories: strategies that rely on features, segmentation, estimate, or learning. Two well-known methods are highlighted specifically in the estimation-based category: particle filters and Kalman filters.

**Received: November 03, 2023 Revised: March 21, 2024 Accepted: July 14, 2024**

**Keywords:** Vehicle Tracking; Kalman Filter; Particle filter; Computer vision; smart navigation system

## 1. Introduction

Object tracking in vehicles, which searches for and tracks objects in visual sequences, is a primary responsibility in computer vision. When object tracking, the target must be located and followed in the following frames after it is identified in the initial frame of the video. The application of object tracking is diverse. Systems for medical diagnosis (such monitoring the ventricular wall and managing medical equipment) [8], robotics (such as the ASIMO humanoid robot), traffic monitoring (such as traffic flow monitoring and traffic accident detection), autonomous vehicle tracking (such as path tracking), and activity recognition (such as Applications for object tracking include the identification of human activity and the learning of activity patterns [11].

- Robustness: Robust tracking systems can follow the target in difficult situations such shifting light, occlusion, and background clutter.
- Flexibility: Furthermore, to variations in the immediate surroundings, the target is also susceptible to changes in its intricate and sudden movements. To fix this issue, the tracking mechanism needs to possess the ability to recognize and follow the targets obvious at the moment properties.
- Information processing in real time a framework that works Using picture sequences must possess fast information handling. Consequently, it is necessary to put a high-performing algorithm into practice.

### 1.1 Problem statement

Accurately following objects (cars) across a video series is a challenge in object tracking for vehicles, despite real-world obstacles such shifting light, items that are hidden by other objects, and crowded backdrops. Effective algorithms must be used to accomplish all of this in real-time for applications like traffic monitoring and autonomous driving.

### 1.2 Motivation

Applications ranging from robotics and medical diagnostics to autonomous driving and traffic management depend on vehicles' capacity to follow things with reliability. Real-world situations, however, provide a number of difficulties. By creating real-time algorithms that can modify and recognize different features of objects, we can overcome obstacles such as varying light, occlusions, and irregular motions and ultimately create more intelligent machines, safer roads, and better medical procedures. This is possible by the development of robust object tracking systems.

This format is going to use for the rest of the article. Section 1 Discusses the Introduction. Section 2 contains the Literature Review on various vehicle-tracking methods. Section 3 Object Tracking Datasets. Section 4 Results and discussion. The conclusion and future scope covered in Section 5.

### 1.3 Classification of the method for object tracking in vehicle

There are various categories for object tracking techniques, for instance. Many author describe the different types of object tracking method. In this paper we focus on estimation-based tracking. The final grouping is employed in this work, and as Figure 1 illustrates, a thorough categorization of object tracking techniques is offered. Specifics about each technique are supplied below.

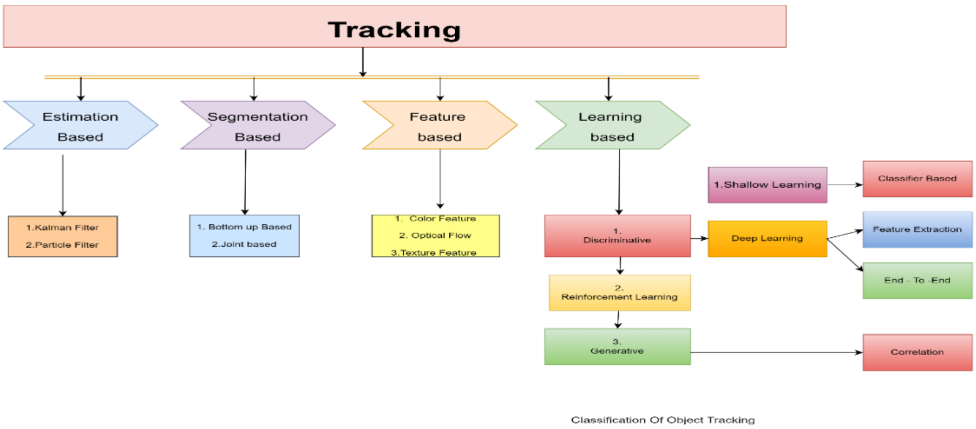


**Figure 1.** Object tracking

### A. Estimation Based Tracking

A state vector represents an object in the tracking problem, which can be state as an estimating problem. The state vector defines the location and speed of a thing as well as other characteristics of the way a system behaves dynamically. The dynamic mode estimation problem has a general foundation provided by Bayesian techniques [2]. The Bayesian filters, which are based on the most recent sensor data, allow the target to constantly update its location on its coordinates. There are two phases to this recursive technique: updating and predicting. While the target is updated in the updating phase using the latest observation utilizing the observation model, the prediction step employs the state model to estimate the target's new location in the subsequent step [3].

- Kalman Filter to make use of the Kalman filter [4] for object tracking, an ever-changing target model motion must be created. A linear system's Kalman filter is utilized to estimate position under the presumption that the Gaussian faults are present. Since Models that are dynamic are often nonlinear, in these cases other suitable methods are used instead of the Kalman filter. Among these algorithms is the expanded Kalman Filter. The structure See Figure 2 for utilizing the filter of Kalman. Some works, such [5] track using the Kalman filter.

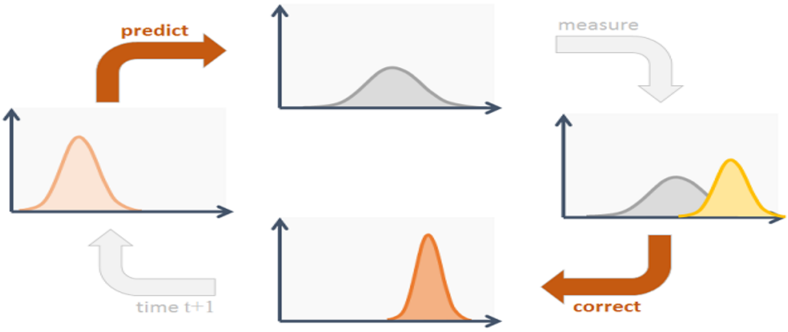


Figure 2. kalman filter framework [10]

- Particle Filter Problems with non-linear tracking occur. Thus, particle filters have been studied as possible remedies for these problems. Non-Gaussian noise quantification models often make use within a particle filter, a recursive Monte Carlo statistical processing approach.

The particle filter's primary concept illustrates the dispersion of a group of constituents. The probability that the weight of a particle indicates how likely it is to be sampled based on the likelihood density function. Among the issues with this approach is that particles having a higher likelihood of being selected more than once; this is addressed by resampling. The framework for applying Figure 3 depicts particle filters. Particulate filters find application in [1], [6].

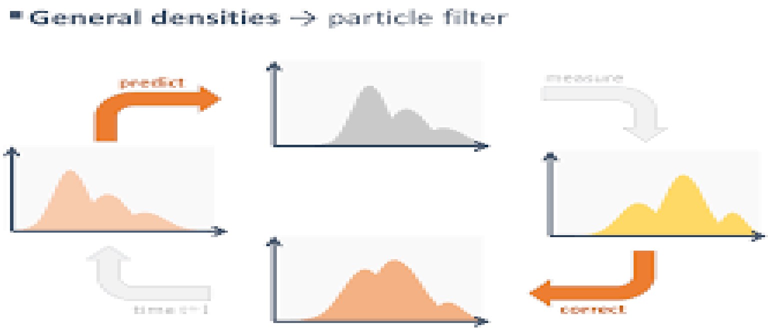


**Figure 3.** Particle filter Framework [10]

### B. Feature Based Method

One of the simplest approaches to object tracking is this one. Features like color, texture, optical flow, etc. are initially extract to track objects. In order to make objects in the feature space easily recognizable, these extracted characteristics need to be unique. After the characteristics have been retrieve, the next stage is to use those features and a similarity criterion to locate the object that is most similar in the following frame. One of the issues with these approaches is that in order to differentiate the target from other things, the distinctive, precise, and dependable characteristics of the object must be retrieve. The following are a few features that are employ in object tracking.

- Color; the objects can show via its color feature. Although there are many applications for this characteristic, the color histogram is one of the more popular ones. The color histogram, which counts the number of pixels in every color, shows how the colors are distributed throughout an image. The drawback of color histograms is that they only take into account an object's color, ignoring its texture or shape. As a result, two distinct objects may have the same histogram. A few articles, such [16]–[18], use color histograms for tracking.
- Texture: Texture is an information pattern that is repeat or a structure that is arrange at regular intervals. Texture characteristics are not acquiring directly. Image preprocessing techniques produce these features. An image's texture characteristic is crucial since it can used to define an image's contents or a specific area of it, in addition to the color feature. Because separate photos can occasionally see to have the same histogram and because color features alone are insufficient to distinguish identical items. The Gabor wavelet [19] is a highly

researched aspect of texture. The key characteristic of Gabor filters that makes them ideal for object tracking is their invariance to light, scale, rotation, and translation. In [20], a method for utilizing a Gabor filter to identify the position and body forms of moving animals is providing. Another textual descriptor is the local binary pattern [21]. Zhao et al. utilized LBP to define moving objects in addition to using a Kalman filter for target tracking [22].

- Optical flow: The apparent movement of the image's brightness patterns known as optical flow. Even in the absence of real motion, changes in the lighting can produce the illusion of motion. The algorithm for optical flow. Determines the brightness pattern displacement between two frames. Dense optical streaming techniques are those that compute displacement for every pixel in an image, whereas light-flow algorithms only estimate displacement tension for a subset of pixels [23]. Several works, such [24], and [25], use optical flow for tracking.

### C. Segmentation method

The most important and basic stage in visual tracking is segmenting foreground objects from a video frame. To distinguish foreground objects from the background scene, foreground segmentation is employed. Usually, the moving items in a scene are in the foreground. To be monitored, these objects must be kept apart from their surroundings [15]. The following looks at a few segmentation-based object tracking strategies.

- Bottom up based Method Foreground segmentation and object tracking are two distinct activities that must be perform in this kind of tracking. After extracting regions from each frame using a low-level segmentation technique, the foreground segmentation extracts certain features from the foreground regions and tracks them in accordance with those features [26]. Certain studies, like [27]–[31], employ a bottom-up tracking strategy.
- Joint based method Foreground segmentation and tracking are two distinct jobs in the bottom-up approach; one drawback of this approach is that segmentation errors propagate forward and result in erroneous tracking. The researchers combined the tracking and foreground segmentation methods to overcome this issue, which enhanced tracking performance.

### D. Learning based method

In learning-based approaches, various targets' attributes and appearances, together with their predicted locations, are learn in subsequent frames. During testing, the subjects can then identify and follow the item in subsequent frames by applying their newly acquired knowledge. The three categories of learning-based approaches are typically discriminative, generative, and reinforcement learning.

- Discriminative Method Discriminative trackers typically view tracking as a classification issue that separates the target from the background. The two categories of discriminative learning comprise Deep learning as opposed to shallow learning. Certain works, like [32]–[34], employ a joint-based tracking technique.
- Reinforcement learning we come across an agent in a reinforcement learning [68] situation that learns to choose the best course of action to accomplish the objective by interacting with the environment through trial and error. Some works, like [69]–[71], use a tracking technique based on reinforcement learning.
- Generative Method these techniques focus on simulating the final object's appearance. They take up on the properties of the object and utilize that knowledge to guess where it will be in the next frame. They do not specifically consider the background, instead depending mainly on the information of the target object. These techniques focuses on looking in locations that are closer to the target. Among these trackers are techniques based on correlation filters.

### E. Object Detection

An object detection technique is need for every tracking approach, either at the beginning of the movie or in every frame. One popular method for object detection is to leverage single-frame information. To lower the amount of false detections, some object detection techniques do, employ the temporal information derived from a series of frames. Typically, this temporal information takes the form of frame differencing, which draws attention to areas that change across successive frames. The tracker's job is then to construct the tracks by performing object correspondence from one frame to the next given the object regions in the image. (2006)

- Frame differencing it includes identifying regions with significant modifications, which most likely imply motion, by subtracting the pixel intensities between successive video frames. By concentrating their processing on these dynamic areas, object detection algorithms benefit from this and may become more efficient. For background subtraction, frame-differencing may also employ to isolate moving foreground objects that may include objects of interest for additional detection.

- Optical flow because it reflects other image information and may be utilized to derive object motion information, the optical flow technique is a crucial method for motion image sequence analysis. Along with the motion data of the item in the image, it also provides a plethora of data on the 3-D physical structure. By using the optical flow, we may accomplish noteworthy uses in a variety of fields, such as medical, information science, traffic management, military space, and meteorology [7].
- Background subtraction Creating a background model, or representation of the scene, and then identifying departures from the model for every incoming frame are the steps involved in achieving object detection. A moving item indicated by any appreciable deviation of an image region from the background model. The pixels that make up the areas that are changing designated for additional processing. To find linked regions that match the objects, a connected component approach is typically used. We call this procedure the background subtraction. Foreground detection, sometimes referred to as background subtraction, is an image processing technique that extracts an image's foreground for further processing. Background reduction is the most widely used technique for recognizing moving objects in movies captured by fixed cameras. It is typically completed if the image is a part of a video stream [5]. The following techniques are used in Background Subtraction to vary the intensity values of the background pixel: adaptive multi-cue Background Subtraction, mixture of Gaussians unmodal distributions, and non-parametric Kernel Density Estimation. The foreground cars create a foreground mask when they are isolated from the background using the Background Subtraction technique. In addition to being utilized in traffic analysis, target tracking, object tracking, and video appliances, background subtraction can also be used to identify foreground objects by comparing many frames. In the event that the threshold value is smaller than the difference image, the image is interpreted as either a background or a moving item. The video sequences are watched in a way that shows each moving object is comprised of a color and that I is composed of a static background B. In this case, the Background Subtraction encapsulates the formula like thus:

$$(s) = lif(IS, Bs) > t, \qquad ()$$

else 0 Where d is the distance between Is,t, the color at time t, and pixel s, and t is a threshold. In this case, $Xt$ is the motion label field at time t. The basis of this method is the calculation of the error between the background and current frames. In figure 4. shows vehicle analysis system in typical video.

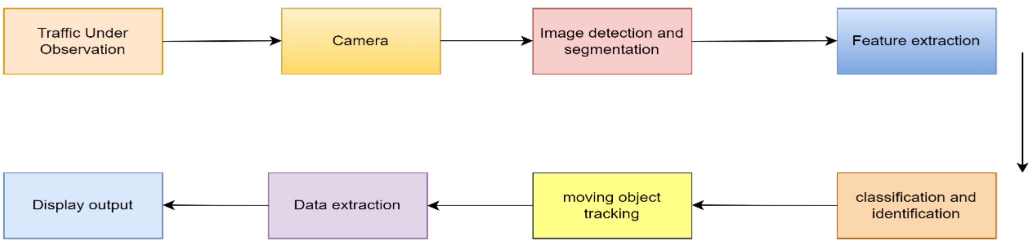


**Figure 4.** Vehicle analysis system in typical video [9]

## F. Object Tracking

An object tracker aims to pinpoint an object's location in every video frame, effectively creating its path over time. This process can also reveal the entire area the object occupies within each frame. Object tracking can be achieved in two ways: independent detection and correspondence, or joint estimation. In the first approach, object detection first identifies potential object regions in each frame, and then the tracker correlates these regions between frames. Conversely, the second approach iteratively updates object location and region data from previous frames to simultaneously estimate the object's region and correspondence in the current frame. The chosen model to depict an object's shape influences the kind of motion or deformation it can undergo. For example, a point representation only allows for translational movement, while an ellipse can handle parametric motions like affine or projective transformations, suitable for approximating rigid objects. Lastly, silhouettes or contours are the most descriptive representations for non-rigid objects, and their motion can be described using both parametric and nonparametric models.

- Point Based: Points are used to represent objects that are identified in successive frames. The association between the points is determined by the object's position and velocity from the previous frame. With this method, the objects in each frame must be detected by an external mechanism.
- Kernel Based: Objects appear and shape make up its kernel. The kernel might be a rectangular template or an elliptical shape with a corresponding histogram. Calculating the kernel's motion over a sequence of frames is necessary for object tracking. This motion often manifests as an affine, translational, or rotational parametric transformation.

- Silhouette Based: Estimating the object region in every frame is how tracking is done. Methods for tracking silhouettes make advantage of the data embedded inside the object region. Shape models, which typically represented as edge maps, and appearance densities are two possible formats for this data. Tracing silhouettes requires either shape matching or contour evolution given the object models. When applied in the temporal domain, both of these techniques essentially thought of as object segmentation using priors created from earlier frames.  There are different steps used for object Tracking.

### G. Object Classification

- Shape based: In computer vision and image processing there is a technique known as the 'shape-based object classification' which is used to categorize objects based on the geometric shapes of the objects. Shape-based classification is centered on an object's features as an object, while motion-based classification emphasizes the characteristics of its movement.
- Motion Based: It is a method of categorizing objects in videos or image sequences based on the patterns of their motion and is popularly referred to as motion-based object classification.
- Color Based: By understanding an objects' color characteristics, objects in a picture or a video can be recognized using a method known as color-base object categorization. It enlarges the object's visibility and make use of the color information's solidity.
- Texture Based: Texture-based item categorization is the method of identifying things in photos by analyzing the patterns of repetition or deviation in the surface features of objects. It is explained below in figure 5 how basic steps for object tracking.

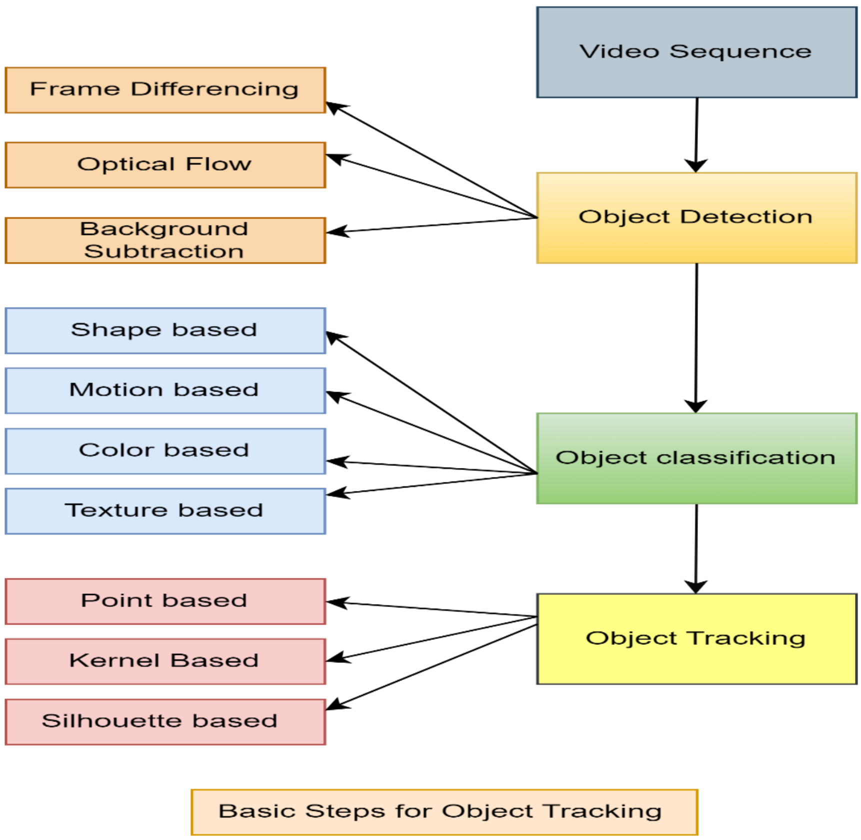


**Figure 5.** Basic steps for object tracking

## 2. Literature Review

Since the movement of cars is an important aspect to be analyzed, the study discusses the compatibility of the UAV technology and the video processing in traffic surveillance and proposes an approach. This method is demonstrating by analyzing the vehicle-tracking data from the Unmanned Aerial Vehicles (UAVs) and subsequently, the paper has proposed a model-based object tracking mechanism [9].

This method is specifically design to detect and track vehicles in traffic situations that are record by fixed cameras. It is necessary to employ a parameterized vehicle model for the intrafirm pairing as well as a recursive approximation on top of a model of motion when estimating the motion, which is also the significant component of this method. Besides, the study employs the clustering of features from the moving images to determine the initial position and orientation of the models [3].

Moreover, the study provides a car tracking system with the detection-based tracking framework and the network model for the efficient extraction of features. This system uses the intersection over union (IOU) and information on object details to change the detection box and enhance the detection accuracy. It employs You Only Look Once (YOLO) method for the identification of vehicles. The tracking model then employs a combination of geographical limitation and filter template matching to enable continuous tracking based on the object's spatial position, motion vector, and feature history. In comparison with the Staple algorithm in the OTB2015 dataset, the suggested approach outperforms the current approaches and stands first in terms of average accuracy and success rate, thus proving the efficiency of the suggested approach in the vehicle tracking system [10].

Nevertheless, the current study focuses on the use of CNNs for car recognition, but admits that feature selection is absent and hampers the accuracy of classification. To address this difficulty, this paper proposes the Enhanced Convolutional Support Vector Machine-Integrated Neural Network (ECNN-SVM) approach for detecting vehicles. This method employs Haar-like features which is famous for its ability to compactly represent an image and to collect data in many sizes for feature extraction. Based on the sample's class name and value of features, the improvement of bat optimization is to select features from the Haar-like feature pool. When integrated with a local binary pattern and Support Vector Machine (SVM) blend, the ECNN classifier effectively eliminates interference zones between moving objects and cars during classification [11].

Furthermore, the study suggested a real-time tracking system for several vehicles to address the issue of gridlock in the roads detection. The writers draw attention to the fact that video-based traffic monitoring systems are becoming more and more popular because of their extensive information sources, cost, and flexibility. The proposed method for tracking multiple vehicles includes improved vehicle detection, data association algorithms, and the Generalized Multi-Dimensional Probability Hypothesis Density (GM-PHD) filter. The paper also addresses the use of a unique vehicle recognition method based on the constant velocity model to improve track initialization, hence advancing real-time multi-vehicle tracking techniques for the identification of traffic congestion [12].

Object tracking is a well-known problem in the field of video processing that is frequently solve with Kalman filters. Certain research concentrate on evaluating basic conditions and using environmental data to direct tracking protocols. Specifically, Situation Assessment (SA) seeks to improve information quality and identify items from an operational perspective. In the field of public transportation, vehicle tracking is a particularly interesting problem, which has led to several research projects using Kalman filters. The objective of this study is to improve vehicle-tracking efficiency by including online situation assessment (SA) into the Kalman filter architecture. It explores the use of multi-object tracking (MOT) when it comes to traffic monitoring and CCTV, with a focus on tracking cars [2].

The authors suggest extending the Deep SORT tracking algorithm with low confidence in order to reduce false positive tracks. They also introduced the UA-DETRAC re-identification of vehicles self-generated data set to educate the Deep SORT neural network with convolution data integration. Using the UA-DETRAC test dataset and experimental results show performance that is both better than the most sophisticated online trackers and on par Using trackers in batch mode [13].

The study is divided into sections that address the description of the re-identification dataset, tests and results, conclusion, and a thorough methodology of the suggested tracking strategy. Another study uses color segmentation and convolutional neural networks (CNNs) to present DTT (Detection, Tracking, and Tracker) as an efficient vehicle tracking technique [14].

This method is not dependent on any particular object and provides flexibility in tracking other items. An enhanced Shift-variant CNN architecture is also recommended by the study to lessen tracking errors brought on by variations in the target's visual appearance. Furthermore, a method for the tracking and identification of automobiles and pedestrians in video frames is suggested that combines CNNs with the YOLO algorithm. Using YOLO's estimated parameters and object features, this technique employs a tracking region algorithm and a CNN-based classification algorithm to predict the presence of cars and pedestrians. This allows for the accurate identification of moving objects. The study highlights deep convolutional neural networks' potential for object detection and positions them as viable substitutes for conventional FG-BG algorithms in the processing of vehicle videos [15].

Khan et al. recorded a video of an urban commercial roundabout employing a UAV, then generated entry/exit matrices using image processing and tracking techniques. Numerous strategies have been developed for traffic study in several kinds of roads, including roundabouts. Some of these systems make use of UAVs for data gathering and picture processing. In earlier research, camera systems were utilized to record video close to roundabouts; the resulting entry/exit matrices had error levels between 1.5 and 2.0 [16].

Suggested Over the past ten years, DL techniques have demonstrated robust skills in visual object tracking. These techniques are combined with pertinent filtering techniques by algorithms. Siamese networks are well suited for tracking jobs where obtaining large amounts of labeled data is difficult since they perform well in one-time and few-shot learning scenarios. There exist detection-based and non-detection-based multitarget tracking algorithms, along with distinct methodologies for single-target and multitarget tracking. The goal of integrating Convolutional Block Attention Module (CBAM) with Swim Transformer is to improve the YOLO network's knowledge of traffic scenes. Vehicle tracking jobs now have far better real-time capabilities thanks to the shift from R-CNN to YOLO models in visual object recognition and tracking [17]. suggested an assessment of the literature in the cutting-edge field of underwater garbage detection After 100 training cycles, the improved model's training results on the TrashICRA19 dataset are discussed [18]. Compared ensemble-learning methods for IoT environment threat detection and multiclass categorization. The paper discusses how to create effective IDS using bagging and boosting ensemble decision trees. Several ensemble machine learning multiclass classification methods are evaluated using 'TON-IoT' datasets of IoT and IoT devices. The paper assesses network traffic patterns in Internet of Things environments and identifies many classification methods using decision tree ensembles [19].

Focused all of its attention to assessing compressed formats for deep learning detection tasks. The topic of action recognition from videos is covered, exposing the shortcomings of conventional video analytics procedures [20].

Performed in-depth research on methods for predicting traffic congestion and managing time-series datasets. Numerous machine learning techniques, such as hybrid models and LSTM networks, have been investigated for the purpose of forecasting traffic congestion in smart cities.[21].

**Table 1:** Comparative analysis of various machine-learning techniques

| References | Contributions of this survey | Limitations | Algorithms Used | Results | Future scope |
|---|---|---|---|---|---|
| [1] | Evaluates the methodology's accuracy in analyzing each vehicle path separately. | Excludes details regarding the precise outcomes or discoveries of the experiments. | OpenCV. | For the cases 1 and 2, the corresponding root mean square percentage error was determined to be 3.96 and 3.57, respectively, in the speed evaluation. | Taking consideration of further camera settings to improve ground coordinate matching. |
| [2] | Increases the accuracy of vehicle detection by fusing IOU with object attribute data. | Results are comparable to Staple's, but not superior. | Gradient ascent | The usefulness of the suggested model-based object tracking approach is demonstrated by the paper's results on actual traffic scenarios. | Adding a search tree to improve the matching procedure |
| [3] | lowers computing costs by speculating on possible locations for vehicles | Decrease in classification precision with insufficient feature data. | Lightweight feature extraction | Analyzing the suggested algorithm's efficacy Performance and speed evaluation using publicly available datasets. | Analyzing the suggested algorithm's performance and generalizability on more datasets |

| | | | | | |
|---|---|---|---|---|---|
| [4] | Adding the D-GMPHD filter to multiple vehicle tracking based on images, The ability to track early allows for the timely discovery of traffic congestion. | Ineffective track initiation strategies in the multiple vehicle tracking techniques now in use. | The foreground and background are detect using the OTSU technique. | When compared to other techniques, the suggested ECNN-SVM method achieves a higher f-measure. | There is scope to enhance optimization and improvement of the procedure. |
| [5] | Online SA has a 25% performance boost, according to experimental results. | There was no information offered regarding the approach's possible shortcomings or restrictions. | - | The study suggests Real-time tracking of several vehicles bottlenecks. | Larger and more varied datasets can used for additional testing and validation. |
| [6] | Obtains encouraging outcomes with significant margins over the most advanced trackers. | The limits of the suggested approach and any potential negative effects of adding the extension of the low confidence track filtering to the Deep SORT algorithm are not thoroughly analyze in the research. | SAKF(situation assessment Kalman filter | Using online SA within KF improves performance by 25%, according to experimental results. | Including both long- and short-term adjustments separately in the future |
| [7] | Creation of a shift-variant architecture to reduce drift issues.<br>A discriminative model for object tracking that extracts features from dynamic states. | CNN model fails to track targets as their appearance changes because it is not built for long-term occlusions. | DEEP Sort | Performs comparably to batch-mode trackers and better than cutting-edge online trackers. | The suggested approach can enhanced even more by adding more sophisticated detecting methods |

| | | | | | |
|---|---|---|---|---|---|
| [8] | Offers a suggested algorithm for organizing a video detection result into video frames for monitoring cars and pedestrians. | No comparison with other algorithms or approaches currently in use | CNN | Multiple items able to monitor by the approach in congested and complex situations. | The suggested approach's 3-D implementation will be the focus of future research. |
| [9] | Makes use of a deep CNN to detect objects and extract features. | There was no information offered regarding the algorithm's performance or accuracy. | CNN, YOLO | Foreground-background differentiation and object detection predictions, training sets and networks are made available | Extending the capability to encompass bikes and pedestrians to improve comprehension of the allocation of citizen space. Along The path |
| [22] | Implement CNN-based classifier to improve classification accuracy | Limited discussion on the impact of weather conditions on accuracy | Yolo, SSD, CNN | increased accuracy of vehicle classification from 57% to 95.45% | Examine methods to increase the precision of object detection for traffic flow monitoring. |
| [23] | Applying YOLO and other deep learning techniques to enhance vehicle detection | Lack of a comparison study with the current traffic monitoring systems. | YOLO v4, Deep Learning | 99% vehicle recognition rate, less than 20% errors in classification | In the future, this will involve using cutting-edge AI technologies to improve traffic monitoring. |
| [24] | The study improves CAMShift, a technique for multi-vehicle tracking with video cameras. | Conventional CAMShift limitations include window size, insufficient color features, and inefficient prediction. | CAM shift Method | The suggested approach more correctly reflects vehicle movement by adjusting window widths. | examining how environmental influences affect results tracking and creating plans to lessen their impact |
| [25] | Improve network coverage in places where | Lack of data regarding system | GPS,GSM | System recognizes incidents and alerts authorities | One area for development is the inclusion of a |

| | | | | | |
|---|---|---|---|---|---|
| | connection is lacking. | performance in actual accident situations | | via alert messages with the location | wireless camera for accurate help. |
| [26] | formulates the issue as one of one-class classification with anomaly detection & Using deep learning for automatic road accident identification and localization | Not fully explored is the comparison with supervised learning techniques. | LSTM | Achieved significant advances in the localization and automatic detection of traffic accidents. Model's accident detection F1 score was 78.58. | Future objectives include for deploying hardware for quicker video processing and adding more features, such as characteristics of the road, for increased accuracy. |
| [27] | Use motion vector estimation precisely determine the positions of the vehicles. | lacks a thorough comparison with the current car monitoring systems | Faster, R-CNN | 95% of vehicles accurately discovered, 90% are correctly classified, and 92% of tracks generated well. | Improving real-time tracking and vehicle detection accuracy are among the future goals. |
| [28] | Utilize use of modern DL methods for reliable vehicle identification and categorization. | Little attention paid to the difficulties of DL vehicle detection. | Descriptive and inferential analysis | thorough examination of benchmark datasets, performance metrics, and DL approaches | Improving DL methods for vehicle identification and classification is part in the future scope. |
| [29] | Concentrate on next research with more automation. | limited dataset as a result of worldwide bans on automated cars | YOLOv3 ,SSD | When the automated system is in use, there are fewer accidents at high speeds. | Further investigation into higher automation levels is part of the future scope. |
| [30] | Intelligent car tracking device for detecting and identifying theft. Moreover, Alerts for incidents, fires, and GSM module tracking in real time. | Insufficient details regarding the system's implementation costs | GSM,GPS | Detection of gas, fire, smoke, and collisions in moving cars. Notifications of emergencies, such as fires or accidents, to registered numbers. | Future plans aim for automating tasks using AI integration and improving sensor capabilities. |

| | | | | | |
|---|---|---|---|---|---|
| [31] | Introduced the attention mechanism, anti-bottleneck structure, and smooth loss function. | Real-time performance and computing efficiency discussed. & Absence of comparison with alternative cutting-edge car tracking systems | FairMot , Deep learning | Improved roadside view vehicle tracking achieved with the proposed approach. Achieved optimal MOTA outcomes in contrast to popular tracking techniques. | Potential avenues for further research include improving real-time robustness and performance. |
| [32] | Vehicle detection using statistical methods, deep learning, and YOLO-v3. improvements to the multi-scale vehicle detection method and new structural changes | Lack of comprehensive assessment measures for the suggested methods | YoloV3 ,Blob detection | A statistical technique uses reflection patterns to effectively identify and track automobiles. | Improving real-time detection and statistical flow recalculation are among the future scopes objectives. It is projected that multi-scale vehicle detection algorithms will advance. |
| [33] | improved tracking success rate and lower average central-pixel error were attained | Absence of discussion regarding the suggested algorithm's computational complexity | Particle filter, semi supervised learning | The suggested approach produces fewer errors and a higher tracking success rate. | Enhancing deep models and real-time performance are among the future scope's objectives. |
| [34] | Empirical testing to identify ideal parameters, such as the smallest distance between clusters | Trajectories interrupted by missed detections in vehicle detection, and inaccurate tracking data may arise from false alarms. | JPDAF stastical , UKF | In urban settings, lidar sensors provide precise vehicle detection and tracking. The accuracy of the vehicle speed estimate from 0.41 to 0.22 m/s. | Improving sensor fusion methods in the future will increase classification accuracy. |
| [35] | A multi-sensor fusion framework | requires constant internet | Decision tree, | The F1-score for the LR-based stacking | investigating cutting-edge tracking and |

| | | | | | |
|---|---|---|---|---|---|
| | suggested improving the effectiveness of classification. | access; inaccurate outcomes due to equipment malfunction | Logistic Regression | classifier was 0.95. Moreover, The DT-based ML model demonstrated good recall and precision. | vehicle detection methods to increase the precision of speed estimation |
| [36] | Novel structure for tracking single and many cameras. | Recently, there has been little research on vehicle re-identification. | - | Achieved first place in multi-camera tracking and 3D speed estimation. | Potential avenues for future research include improving vehicle re-identification algorithms. |
| [37] | Suggested collision avoidance, speed-based lane changes, and TOA-based localization algorithms | Transmissions of emergency messages delayed due to infrastructure problems. | RFID, GPS, VANET | Simulations and field research showed effectiveness of the suggested system. | Improving the defenses against malevolent assaults on the automotive cloud network |
| [38] | The suggested multi-vehicle tracking algorithm achieves notable gains over earlier techniques by properly predicting the next features of tracked objects based on appearance and motion features. | The study admits that choosing the incorrect car with a similar appearance can result in tracking errors when appearance features are limited without motion. | YOLOv4 | Obtained 86.3% MOTP and 84.5% MOTA on KITTI dataset. enhanced MOTA by 8.6% and MOTP by 9.6% compared to earlier techniques | Future research will involve training on difficult datasets with compact cars. |
| [39] | The proposed ISIEMF framework aims to detect traffic incidents and handle evidence. | Absence of labelled real-world datasets for assessment | LSTM | LSTM conflict resolution results in a 66.24% reduction in network overhead.<br><br>The chance of OBU failure detection peaks at 93% at a speed of 20 ms. | Improved traffic incident detection efficiency is one area of future focus. |

| | | | | | |
|---|---|---|---|---|---|
| [40] | Improved fusion rate, accuracy, and decreased complexity and errors. | Insufficient investigation of scalability problems in large-scale intelligent transportation networks | Deep Learning | Enhanced fusion rate of 0.9741 and accuracy of 92.078%. fusion time, reduced error of 0.0662, and complexity of 0.0717 | Future work will focus on increasing efficiency, lowering complexity, and boosting accuracy. |
| [41] | Suggested a two-phase target association technique to monitor the trajectories of the workers. | Just a brief examination of the comparison using other techniques. | YOLO, ACKF, Kalman filter | Tracking target success rate: 90.1% with occlusion and 97.3% without it | Improving YOLO5 to increase multi-objective tracking precision |
| [42] | Transformers and YOLO models combined in a hybrid Siamese network for tracking | The significance of dynamic vision sensors is briefly discussed | Hybrid Siamese network, CBAM | The model recognizes cars both close by and far away with accuracy. | Future studies could look into dynamic vision sensors and spiking neural networks. |
| [17] | The proposed approach lowers the erroneous track rate and concentrates on moving cars. | Due to thresholds for duration and distance, some trajectories might not merge. | Yolo v4 | Error rate was lower using the suggested strategy than the benchmark. | Examine more sophisticated object tracking methods to enhance the detection of vehicles. |

The primary findings is that authors can actively work on creating algorithms in the future to solve issues with robustness, accuracy, and speed in practical settings. They accomplish this by investigating various strategies such as deep learning architectures, filtering techniques, and feature extraction. Although some research present encouraging findings, there is always room for improvement due to constraints such as the absence of performance comparisons or information on managing long-term occlusions. The overall goal of the project is to develop reliable and effective algorithms for precise vehicle monitoring in a range of applications.

## 3. Object tracking datasets

Algorithms based on deep learning for object tracking require datasets. Pedestrians and automobiles are the primary tracking objects of intelligent cars in real-world road scenarios, and object-tracking databases that include both data have become widely used. Table 2 displays a summary of several example object tracking datasets.

### 3.1 Visual Object Tracking Challenge (VOT) Dataset

The VOT, a prestigious worldwide competition in the area of visual tracking, is the source of the VOT dataset [43]. The dataset has been made available annually since the competition's launch in 2013 and has grown to be the norm for assessing the effectiveness of single item tracking methods. VOT datasets have been employed to assess real-time effectiveness in addition to long-term object tracking, even though their initial purpose was short-term object monitoring. The 60 video clips in the VOT2017 dataset cover a variety of challenge features, including scenario elements like as the camera moves, occlusion, lighting, scale, and motion all vary. This dataset can track tiny objects, and its adjustment data can be precise down to the pixel level with sufficient testing performance. The VOT2018 dataset was assessed using the following metrics: robustness, accuracy, and EAO. The single object tracking approach works best when resilience and precision scores are low and EAO and accuracy levels are high.

### 3.2 LaSOT dataset

The LaSOT dataset [44] was created in collaboration with Temple University, South China University of Technology, and Meitu HiScene Laboratory. It's a huge, superior, long-term object tracking dataset. the biggest single item tracking dataset with deep labelling currently accessible. This dataset comprises 1400 video sequences containing 3.52 million frames, or 2512 frames on average every series. The collection contains 20 sequences for each of the 70 item categories, 16 of which are Sequences are used as training sets and test sets, accordingly, for four sequences. The LaSOT dataset offers detailed natural language requirements and visual bounding box annotations, in contrast to other datasets. Furthermore, this dataset has 14 scene factors, which are more difficult features. Such as changing the perspective (VC), background clutter (BC), aspect ratio (ARC), low resolution (LR), and camera motion (CM).

**Table 2:** Object Tracking Datasets

| Datasets | Time | Tracking objects | Characteristics |
|---|---|---|---|
| UAV [19] | 2016 | Vehicle and ,Pedestrians | Tracking model for tracking a single object in both normal and challenging circumstances |
| OTB [25] | 2013 | Vehicles,Pedestiants | extensive dataset |
| LSVH [35] | 2018 | Vehicle, | Typical problems with appearance-based item classification and detection |
| Tracking Net Dataset [45] | 2018 | humans, animals | vast single object tracking dataset, evaluated using an internet server without test labels |
| nuScenes dataset [47] | 2019 | Vehicles | The initial dataset with a complete sensor kit |
| VOT [45] | 2013 to present | Pedestrians, vehicles, etc. | Mainstream a dataset to track individual items using consistent indications. |
| LaSOT [43] | 2019 | Vehicles | Extensive, superior, long-term single-object tracking dataset |
| UA-DETRAC dataset [44] | 2015 | Vehicles | a challenging multi-object tracking and identification dataset from the real world |
| Mot Challenges Dataset [47] | 2015 | pedestrian | include a number of smaller datasets with various situations and camera perspectives. |
| Waymo Open dataset [44] | 2019 | Human,vehicles | dataset for the segmentation of panoramic videos. |
| Omni dataset [44] | 2020 | Vehicles | extensive synthesis of open dataset |
| BDD100K dataset [44] | 2018 | Vehicles | a vast and varied collection of driving videos |
| ILSVRC [45] | 2012 | Human,animal,vehicle | a huge and varied driving dataset |
| Got10K dataset [46] | 2019 | Vehicles | A sizable multi-scene generic single object monitoring dataset was created using the WordNet semantic hierarchy technique. |
| KITTI Dataset [46] | 2012 | Pedestrians. | Include minimal scene objects with annotations for pedestrians and vehicles. |

### 3.3 Tracking Net dataset

One of the biggest single item tracking datasets available, TrackingNet [46], contains carefully chosen inside scenes from the Youtube-BoundingBoxs dataset, which consists of 30,643 videos with an average length of 16.6 seconds and an overall time period of 140 hours. Of these, 511 are employed for testing, and the remaining 30-132 are utilized for instructional videos. The TrackingNet collection primarily focuses on outside things, such as people, automobiles, and animals. There are 15 challenging features in all, such as similar object (SOB), motion blur (MB), deformation (DEF), background clutter (BC), and camera motion (CM). The object tracking algorithms are assessed online using a server, and no test labels are given.

### 3.4 Got 10 K dataset

The Chinese Academy of Sciences released the GOT-10k dataset [47], a thorough multi-scene generic single object tracking method assessment dataset, and public. It also employs the WordNet semantic hierarchy approach for the first time in a video trajectory dataset. This dataset, which includes 10,000 video sequences, aims to provide more than 1.5 million images of video frames with accurate bounding box point annotations, in order to capture several dynamic things as is feasible in the real world. These images feature 87 distinct motion modes and 563 distinct object types. The GOT-10k dataset consists of three sets: testing, validation, and training. Additionally, by implementing the one-shot approach for monitoring evaluation for the first time, it achieves 0% overlap between the training and testing sets.

Furthermore, this dataset includes a number of difficult features, including quick motion, scale changes, lighting variation, and occlusion (OCC). (FM), and in recent years it has become popularity as a benchmark.

### 3.5 MOT challenge dataset

The most frequently utilized dataset to test different object tracking algorithms, particularly those pertaining to pedestrian tracking, is MOTChallenge [48]. The MOTChallenge dataset, as of late, comprisesSub-datasets MOT15, MOT16/MOT17, and MOT20. This series' first-generation sub-dataset, MOT15, is separated into 2D and 3D datasets. Twenty-two video sequences total—eleven for training and eleven for testing—make up the 2D dataset. The MOT16/MOT17 sub-dataset presents various issues due to its very high pedestrian density. The eight video sequences in the MOT20 sub-dataset feature highly intricate tracking situations with a high pedestrian density, making them ideal for assessing the tracker's overall performance.

### 3.6 Kitti Dataset

The KITTI dataset is a thorough collection of scenarios related to autonomous driving that was produced in partnership with the Karlsruhe Institute of Technology (KIT) and the Toyota Technological Institute at Chicago (TTIC) in Germany. Ninety-nine testing and twenty-one training sequences total—mostly including the tracking of three different kinds of objects—are included in the KITTI Tracking dataset [46]. Up to 15 autos and 30 people may be seen in each frame of the image in this dataset, which was mostly collected from highway, rural, and urban locations. The collection comprises annotation files that provide information on object type tracking, object crossing and blocking, object rotation angle around the Y-axis in the Camera Coordinate, 2D and 3D identification box placement and size, and the level of confidence in identifying an object.

### 3.7 nuScenes dataset

The nuScenes dataset [48] was released by Motional, formerly known as nuTonomy. It is a substantial dataset that is mostly utilized for 3D object tracking and recognition tasks in the autonomous driving domain. The nuScenes collection is more extensive and more difficult to utilize than the KITTI dataset. It includes 23 item categories, 1000 scenarios, 1.4 million camera images, 390,000 frames of LiDAR point cloud data, and 1.4 million 3D annotation boxes. A 5 long-range millimeter wave radars, 32-line rotating LiDAR, 6 panoramic cameras, IMUs, and GPS were employed to gather the dataset in Boston and Singapore during congested, hazardous driving conditions. 150 testing sets, 850 training sets, and 850 validation sets make up the total of 850 scenarios in the dataset.

### 3.8 Waymo Open dataset

Google's self-driving car company Waymo released a sizable public dataset for outside 3D object monitoring and recognition called the Waymo Open dataset [45]. The Waymo Open dataset has 200 frames per sequence for each of the 202 verification and 798 training sequences. The dataset's data came from five LiDARs and five cameras that were positioned at various points on the car. The data was gathered in Phoenix, Mountain View, and San Francisco, California, and covered a variety of conditions, including day and night, dawn and twilight, rainy and sunny days. 11.8 million 2D and 12.6 million 3D bounding boxes are marked in the collection. The Waymo Open dataset provides better tracking algorithm assessment capabilities and is broader, more diverse, and of higher quality than other datasets of a comparable nature.

### 3.9 UA-DETRAC dataset

The majority of the data for the UA-DETRAC [45] large-scale multi-object tracking dataset is gathered in Chinese cities like Beijing and Tianjin along highways and pedestrian overpasses. A Cannon EOS 550D camera was used to film a 10-hour video series at 24 distinct locales and at various lighting perspectives. The complete dataset consists of 140,000 genuine photos and 8,250 manually categorized automobiles, yielding 1.21 million bounding boxes for detected objects. There are many different types of vehicles, including cars, buses, vans, and more. There are further categories for different types of weather, such as cloudy, nighttime, sunny, and precipitation.

### 3.10 BDD100K dataset

The BDD100K dataset, a substantial collection of multi-scene driving movies, is available for public use thanks to the efforts of the AI Laboratory at UC Berkeley (BAIR) [45]. There are 100,000 full HD 720p videos with a duration of around 40 seconds at 30 frames per second. 100,000 images are produced at the 10th second of the video by sampling key frames. Numerous US locations, involving residential areas in the San Francisco Bay Area and New York, freeways, and metropolitan streets, were employed to capture driving footage. It can be separated into days that are bright, rainy, or snowy based on the conditions. Videos can be categorized as dawn/dusk, midnight, or daylight according to the time of day.

### 3.11 Omni dataset

A huge synthetic public dataset named as the Omni-MOT dataset [45] was created by Deakin University, Chang'an University, the University of Western Australia, and other academic institutions and is available to the public. It includes all MOT scenarios and sequences in the Omni-MOT dataset. The CARLA simulator's dataset encompasses clear, cloudy, and rainy-day sceneries in each of the 5 simulated cities. The Omni-MOT dataset, which consists of 250 K tracks, more than 14 million frame images, and 110 million bounding boxes, offers nearly 1200 times number of frames, 210 times number of tracks, and 30 times number of bounding boxes as the MOT17 dataset.

## 4. Result and Discussion

This study emphasizes the value of cutting-edge methods for processing video, especially when it comes to improving the precision of vehicle tracking for intelligent transportation systems. Tracking performance is significantly improved, outperforming conventional techniques, by utilizing novel algorithms such as Deep SORT and YOLO, and integrating online scenario assessment within Kalman filters. These findings promise improved safety and dependability in applications related to public transportation and traffic monitoring, paving the door for more effective and flexible vehicle surveillance systems. In table 3. Shows the evaluation of vehicle tracking algorithm to show accuracy, efficiency, robustness and real time performance. In addition, figure six. shows the performance analysis metrics.

**Table 3:** Evaluation of vehicle tracking algorithm on key performance indicators

| References | Efficiency | Accuracy | Robustness | Real time performance |
|---|---|---|---|---|
| [3] | Low | Low | - | Yes |
| [5] | Medium | Low | Low | No |
| [9] | High | Medium | Low | Yes |
| [24] | Low | Medium | High | Yes |
| [10] | Low | Decreased | Low | - |
| [11] | High | Medium | High | - |
| [28] | Low | Medium | - | Yes |
| [15] | High | Low | Low | Yes |
| [26] | Medium | High | Medium | Yes |
| [22] | High | High | Medium | YES |
| [23] | Low | High | High | Yes |
| [27] | High | High | Low | Yes |

| | | | | |
|---|---|---|---|---|
| [29] | High | Medium | Low | Yes |
| [37] | High | High | Low | Yes |
| [30] | High | High | Low | Yes |
| [31] | High | High | Medium | Yes |
| [36] | High | Medium | Medium | Yes |
| [32] | High | Medium | Low | Yes |
| [33] | Low | High | Low | - |
| [34] | Low | Medium | Medium | Yes |
| [35] | High | High | Medium | Yes |

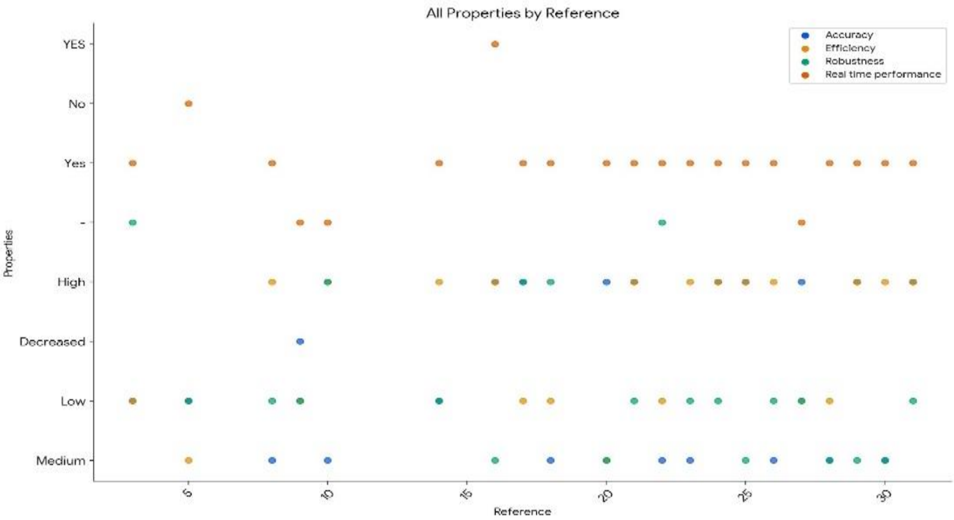


**Figure 6.** Performance evaluation metrics

## 5. Conclusion and Future scope

This paper offers a thorough categorization of object tracking methods, grouping them into four main groups: learning segmentation-, estimation-, and -based, and feature-based. This research specifically emphasizes the object tracking. Nonetheless, investigating various tracking algorithm kinds continues to be a viable field for next studies. Diverse tracking technologies, comprising segmentation-based, feature-based, and learning-based approaches, have the potential to improve tracking performance, robustness, and accuracy in future research. This all-encompassing method of object monitoring has the potential to improve surveillance systems' capabilities and make a positive impact on a number of industries, such as augmented reality applications, video analytics, and autonomous vehicles.

**Funding:** “This research received no external funding”

**Conflicts of Interest:** “The authors declare no conflict of interest.”

## References

**[1]** “Object Tracking Methods:A Review | IEEE Conference Publication | IEEE Xplore.” Accessed: Apr. 17, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8964761

**[2]** “Vehicle tracking with Kalman filter using online situation assessment - ScienceDirect.” Accessed: Apr. 17, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S092188902030436X?casa_token=fi1Z__C2XaAAAAAA:4bcvluZRhjmtoDPw0-ntO96kYAV71QJxuAOpMoTcHx401tE0HmStz3bxJzVuBw8RtCp2M8zE1A

**[3]** D. Koller, K. Daniilidis, T. Thórhallson, and H.-H. Nagel, “Model-based object tracking in traffic scenes,” in Computer Vision — ECCV’92, vol. 588, G. Sandini, Ed., in Lecture Notes in Computer Science, vol. 588. , Berlin, Heidelberg: Springer Berlin Heidelberg, 1992, pp. 437–452. doi: 10.1007/3-540-55426-2_49.

**[4]** "Object tracking and detection techniques under GANN threats: A systemic review - ScienceDirect." Accessed: Apr. 17, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494623002429

**[5]** "Vehicle Tracking System based on Videotaping Data - ScienceDirect." Accessed: Apr. 17, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877042814001487

**[6]** "Multi object tracking: a survey." Accessed: Apr. 17, 2024. [Online]. Available: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11878/118780I/Multi-object-tracking-a-survey/10.1117/12.2602901.short#_=_

**[7]** Y. Zhao, H. Shi, X. Chen, X. Li, and C. Wang, "An overview of object detection and tracking," in 2015 IEEE International Conference on Information and Automation, Aug. 2015, pp. 280–286. doi: 10.1109/ICInfA.2015.7279299.

**[8]** S. M., N. P., and G. P., "Survey on Vehicle Detection and Tracking Techniques in Video Surveillance," Int. J. Comput. Appl., vol. 160, no. 7, pp. 22–25, Feb. 2017, doi: 10.5120/ijca2017913086.

**[9]** A. P. Shukla and M. Saini, "'Moving Object Tracking of Vehicle Detection': A Concise Review," Int. J. Signal Process. Image Process. Pattern Recognit. vol. 8, no. 3, pp. 169–176, Mar. 2015, doi: 10.14257/ijsip.2015.8.3.15.

**[10]** G. Guido, V. Gallelli, D. Rogano, and A. Vitale, "Evaluating the accuracy of vehicle tracking data obtained from Unmanned Aerial Vehicles," Int. J. Transp. Sci. Technol., vol. 5, no. 3, pp. 136–151, Oct. 2016, doi: 10.1016/j.ijtst.2016.12.001.

**[11]** Kavitha, P. Subha, R. Priya, R. "An Implementation Of Statistical Feature Algorithms For The Detection Of Brain Tumor," Journal of Journal of Cognitive Human-Computer Interaction, vol. 1, no. 2, pp. 57 - 62, 2021. DOI: DOI: https://doi.org/10.54216/JCHCI.010202

**[12]** C. Ranjeeth Kumar and R. Anuradha, "RETRACTED ARTICLE: Feature selection and classification methods for vehicle tracking and detection," J. Ambient Intell. Humaniz. Comput., vol. 12, no. 3, pp. 4269–4279, Mar. 2021, doi: 10.1007/s12652-020-01824-3.

**[13]** X. Zhang, S. Hu, H. Zhang, and X. Hu, "A real-time multiple vehicle tracking method for traffic congestion identification," KSII Trans. Internet Inf. Syst. TIIS, vol. 10, no. 6, pp. 2483–2503, 2016, doi: 10.3837/tiis.2016.06.003.

**[14]** "Vehicle Tracking Using Deep SORT with Low Confidence Track Filtering | IEEE Conference Publication | IEEE Xplore." Accessed: Apr. 17, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8909903

**[15]** S. Shruthi, "Vehicle Tracking using Convolutional Neural Network," 2011.

**[16]** V. P, P. D. K, S. S, and A.Tamizharasi, "Video Analysis of Vehicle and Pedestrian Using Neural Network," Ann. Romanian Soc. Cell Biol., pp. 4727–4733, May 2021.

**[17]** E. Avşar, "Moving vehicle detection and tracking at roundabouts using deep learning with trajectory union," Multimed. Tools Appl., vol. 81, no. 5, pp. 6653–6680, Feb. 2022, doi: 10.1007/s11042-021-11804-0.

**[18]** X. Liu, W. Q. Yan, and N. Kasabov, "Moving vehicle tracking and scene understanding: A hybrid approach," Multimed. Tools Appl., Nov. 2023, doi: 10.1007/s11042-023-17618-6.

**[19]** C. Gupta, N. Singh Gill, and P. Gulia, "SSDT: Distance Tracking Model Based on Deep Learning," Int. J. Electr. Comput. Eng. Syst., vol. 13, no. 5, pp. 339–348, Jul. 2022, doi: 10.32985/ijeces.13.5.2.

**[20]** C. Gupta, N. S. Gill, P. Gulia, and J. M. Chatterjee, "A novel finetuned YOLOv6 transfer learning model for real-time object detection," J. Real-Time Image Process., vol. 20, no. 3, p. 42, Apr. 2023, doi: 10.1007/s11554-023-01299-3.

**[21]** S. Sangeeta et al., "Video Object Detection from Compressed Formats for Modern Lightweight Consumer Electronics," IEEE Trans. Consum. Electron. Vol. PP, pp. 1–1, Jan. 2023, doi: 10.1109/TCE.2023.3325480.

**[22]** "Information | Free Full-Text | A Hybrid Univariate Traffic Congestion Prediction Model for IoT-Enabled Smart City." Accessed: Apr. 22, 2024. [Online]. Available: https://www.mdpi.com/2078-2489/14/5/268

**[23]** J. Azimjonov and A. Özmen, "A real-time vehicle detection and a novel vehicle tracking systems for estimating and monitoring traffic flow on highways," Adv. Eng. Inform., vol. 50, p. 101393, Oct. 2021, doi: 10.1016/j.aei.2021.101393.

**[24]** "Development of AI-Based Vehicle Detection and Tracking System for C-ITS Application." Accessed: Apr. 21, 2024. [Online]. Available: https://www.hindawi.com/journals/jat/2021/4438861/

**[25]** "Automatic multi-vehicle tracking using video cameras: An improved CAMShift approach | KSCE Journal of Civil Engineering." Accessed: Jan. 15, 2024. [Online]. Available: https://link.springer.com/article/10.1007/s12205-013-0263-7

**[26]** Badjajian, N. Barry, W. "Deep Learning Driven Automated Red Palm Weevil Detection Using Sparrow Search Optimization," Journal of International Journal of Advances in Applied Computational Intelligence, vol. 6, no. 2, pp. 16-27, 2024. DOI: https://doi.org/10.54216/IJAACI.060202

**[27]** "CAVIDS: Real time intrusion detection system for connected autonomous vehicles using logical analysis of data - ScienceDirect." Accessed: Jan. 15, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214209623000827?casa_token=_dDLagae_c4AAAAA:dpnVVfNNgw3pfYJNHuWIKzapQ8pFLOq6r4x-YDuiXLSaTP46GQkO5fY9W4o-iNA-UmtRdAlR2w

**[28]** K. Pawar and V. Attar, "Deep learning based detection and localization of road accidents from traffic surveillance videos," ICT Express, vol. 8, no. 3, pp. 379–387, Sep. 2022, doi: 10.1016/j.icte.2021.11.004.

**[29]** M. Shehata, R. Abo-Alez, F. Zaghlool, and M. Abou-Kreisha, "Deep Learning Based Vehicle Tracking in Traffic Management," Int. J. Comput. Trends Technol., vol. 67, pp. 5–8, Mar. 2019, doi: 10.14445/22312803/IJCTT-V67I3P102.

**[30]** M. A. Berwo, "Deep Learning Techniques for Vehicle Detection and Classification from Images/Videos: A Survey," Sensors, vol. 23, no. 10, Art. no. 10, Jan. 2023, doi: 10.3390/s23104832.

**[31]** V. K. Y. Chan, "Descriptive and inferential statistics of serious accidents involving considerably automated vehicles—a necessity of smart cities," Procedia Comput. Sci., vol. 219, pp. 856–863, Jan. 2023, doi: 10.1016/j.procs.2023.01.360.

**[32]** P. S. Teja, N. S. S. Reddy, P. K. Pandugu, and P. Vangari, "Smart Vehicle Monitoring And Tracking System," E3S Web Conf., vol. 391, p. 01099, 2023, doi: 10.1051/e3sconf/202339101099.

**[33]** "Sustainability | Free Full-Text | Vehicle Tracking Algorithm Based on Deep Learning in Roadside Perspective." Accessed: Jan. 15, 2024. [Online]. Available: https://www.mdpi.com/2071-1050/15/3/1950

**[34]** "Vehicle tracking and detection techniques using IoT - ScienceDirect." Accessed: Jan. 15, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214785321046733?casa_token=Y4i_57nRiGoAAAAA:aUJzfLuSYHDHtdJSvecFYpi3bHxXqR3o2wGqWd1Y-HVFQlZsQ7e-lb6x6Zse_8B4fW12aNbtOA

**[35]** "Visual Vehicle Tracking Based on Deep Representation and Semisupervised Learning." Accessed: Jan. 15, 2024. [Online]. Available: https://www.hindawi.com/journals/js/2017/6471250/

**[36]** "Vehicle Tracking and Speed Estimation from Roadside Lidar | IEEE Journals & Magazine | IEEE Xplore." Accessed: Apr. 21, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9200682

**[37]** N. Kumar, D. Lohani, and D. Acharya, "Vehicle accident sub-classification modeling using stacked generalization: A multisensor fusion approach," Future Gener. Comput. Syst., vol. 133, pp. 39–52, Aug. 2022, doi: 10.1016/j.future.2022.03.005.

**[38]** Z. Tang, G. Wang, H. Xiao, A. Zheng, and J.-N. Hwang, "Single-Camera and Inter-Camera Vehicle Tracking and 3D Speed Estimation Based on Fusion of Visual and Semantic Features," presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018, pp. 108–115. Accessed: Apr. 21, 2024. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2018_workshops/w3/html/Tang_Single-Camera_and_Inter-Camera_CVPR_2018_paper.html?ref=https://coder.social

**[39]** "Smart vehicle monitoring and assistance using cloud computing in vehicular Ad Hoc networks - ScienceDirect." Accessed: Jan. 15, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S204604301730031X

**[40]** M. S. Abdallah, D. S. Han, and H. Kim, "Multi-Vehicle Tracking Using Heterogeneous Neural Networks for Appearance and Motion Features," Int. J. Intell. Transp. Syst. Res., vol. 20, no. 3, pp. 720–733, Dec. 2022, doi: 10.1007/s13177-022-00320-6.

**[41]** A. O. Philip and RA. K. Saravanaguru, "Multisource traffic incident reporting and evidence management in Internet of Vehicles using machine learning and blockchain," Eng. Appl. Artif. Intell., vol. 117, p. 105630, Jan. 2023, doi: 10.1016/j.engappai.2022.105630.

**[42]** "Multi-sensor information fusion for efficient smart transport vehicle tracking and positioning based on deep learning technique | The Journal of Supercomputing." Accessed: Jan. 15, 2024. [Online]. Available: https://link.springer.com/article/10.1007/s11227-021-04115-6

**[43]** D. Cai et al., "Multi-objective tracking for smart substation onsite surveillance based on YOLO Approach and AKCF," Energy Rep., vol. 9, pp. 1429–1438, Oct. 2023, doi: 10.1016/j.egyr.2023.05.103.

**[44]** Baderiya, P. Gupta, C. Dubey, S. "A Review on Software Fault Detection Mechanisms and Fault Prevention Mechanisms in Networks," Journal of International Journal of Wireless and Ad Hoc Communication, vol. 6, no. 2, pp. 34-42, 2023. DOI: https://doi.org/10.54216/IJWAC.060203

**[45]** H. Fan et al., “LaSOT: A High-Quality Benchmark for Large-Scale Single Object Tracking,” presented at the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 5374–5383. Accessed: May 16, 2024. [Online]. Available: https://openaccess.thecvf.com/content_CVPR_2019/html/Fan_LaSOT_A_High-Quality_Benchmark_for_Large-Scale_Single_Object_Tracking_CVPR_2019_paper.html

**[46]** J. Cao, H. Zhang, L. Jin, J. Lv, G. Hou, and C. Zhang, “A review of object tracking methods: From general field to autonomous vehicles,” Neurocomputing, vol. 585, p. 127635, Jun. 2024, doi: 10.1016/j.neucom.2024.127635.

**[47]** M. Muller, A. Bibi, S. Giancola, S. Alsubaihi, and B. Ghanem, “TrackingNet: A Large-Scale Dataset and Benchmark for Object Tracking in the Wild,” presented at the Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 300–317. Accessed: May 16, 2024. [Online]. Available: https://openaccess.thecvf.com/content_ECCV_2018/html/Matthias_Muller_TrackingNet_A_Large-Scale_ECCV_2018_paper.html

**[48]** L. Huang, X. Zhao, and K. Huang, “GOT-10k: A Large High-Diversity Benchmark for Generic Object Tracking in the Wild,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 43, no. 5, pp. 1562–1577, May 2021, doi: 10.1109/TPAMI.2019.2957464.

**[49]** L. Leal-Taixé, A. Milan, I. Reid, S. Roth, and K. Schindler, “MOTChallenge 2015: Towards a Benchmark for Multi-Target Tracking.” arXiv, Apr. 08, 2015. Accessed: May 16, 2024. [Online]. Available: http://arxiv.org/abs/1504.01942