Accident Analysis and Prevention 154 (2021) 106019

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

A real-time video surveillance system for traffic pre-events detection

Anima Pramanik [a], Sobhan Sarkar [b,*], J. Maiti [a]

[a] Department of Industrial & Systems Engineering, Indian Institute of Technology Kharagpur, Kharagpur, 721302, India
[b] Management Science, Business School, University of Edinburgh, 29 Buccleuch Place, Edinburgh, EH8 9JS, UK

ARTICLE INFO

Keywords:
Pre-event analysis
Anomaly detection
Video surveillance system
Road safety

ABSTRACT

In this study, a conceptual framework is proposed for the development of a video surveillance-based system for improving road safety. Based on the framework, a set of algorithms are developed which are capable of detecting various traffic pre-events from traffic videos, such as speed violation, one-way traffic, overtaking, illegal parking, and wrong drop-off location of passengers. After detecting the pre-events, an alarm will be automatically generated in the control room which helps to take precautionary measures to avoid any potential mishap on road, thereby, improving the road safety. In previous studies, a single system can handle either one or two pre-events. Whereas, in our present study, five anomalies can be detected in a single system using five different algorithms. Our study further contributes to the detection of "wrong drop-off location of passengers". The effectiveness of the developed algorithms is demonstrated over 132 traffic videos acquired from an integrated plant in India. Some additional comparative studies for overtaking and illegal parking are done using two benchmark datasets, namely 'CamSeq01' and 'ISLab-PVD'. Through an extensive study, it can be concluded that our developed algorithms are superior to some state-of-the-art algorithms in the detection of pre-events on road.

1. Introduction

Road plays an important role in the transportation system which helps people and goods move from one place to another. However, it poses a serious concern to the society due to the occurrence of road events. According to the World Health Organization (WHO), approximately 1.3 million people succumb to death every year worldwide due to road events or crashes (WHO, 2020). As per the report of Annual Global Road Crash Statistics (Foundation, 2020), nearly 4.7 lakhs injuries happen due to road crash in India in the year of 2018–2019. Therefore, road crash becomes one of the leading causes of death globally. This problem becomes much worsen under the pressure of an increasing number of vehicles commuted daily on roads. Road crash, however, cannot take place at random. There must be a chain of pre-events that leads to the occurrence of a road crash. Pre-events, such as speed violation, one-way traffic, overtaking, illegal parking, wrong drop off location of passengers, and so on actuate an emergency situation like accidents or crashes. To minimize the number of such crashes in roads, two basic tasks are usually required to be performed: (i) detection of traffic pre-events (i.e., cause of road crash), and (ii) undertaking of corrective actions before the road crash happened. Timely precautionary measures can offer an adequate safeguard to human lives and enhance the level of road safety. These measures can be taken after the correct detection of traffic pre-events. Therefore, an automatic surveillance system becomes extremely necessary under such circumstances, which can deploy machine learning (ML) algorithms more efficiently to analyze the pre-event scenarios to minimize the occurrence of road events. It is noteworthy to mention that the conventional ML models, including support vector machine (Goh and Ubeynarayana, 2017), artificial neural network (de Naurois et al., 2018), clustering (Pramanik et al., 2021a; Sarkar et al., 2018a) and decision tree (de Oña et al., 2013), as compared to the traditional models (Paul et al., 2005), have been successfully used in diverse application areas, such as occupational safety (Sarkar et al., 2017, 2018b, 2019; Sarkar and Maiti, 2020), healthcare (Souri et al., 2020), and so forth.

Several sensory modalities, including Radio Detection And Ranging (RADAR), Light Detection And Ranging (LIDAR), and Computer Vision (CV) are being used to capture/analyze the road event scenarios. Of late, cameras are available at a cheaper rate. Additionally, videos contain rich information as compared to RADAR and LIDAR applications (Datondji et al., 2016). Moreover, CV algorithms can be used effectively to analyze the trajectory of an object of interest (OoI) in the traffic environment (Datondji et al., 2016). With all the aforesaid reasons, current research interests align more with the use of CV in detecting the anomalous traffic

* Corresponding author.
E-mail addresses: apramanik17@gmail.com (A. Pramanik), sobhan.sarkar@ed.ac.uk (S. Sarkar), jhareswar.maiti@hotmail.com (J. Maiti).

https://doi.org/10.1016/j.aap.2021.106019
Received 2 October 2020; Received in revised form 11 December 2020; Accepted 29 January 2021
Available online 31 March 2021
0001-4575/

pre-events. An anomaly can be considered as an event which is deviated from what is normal or standard or expected. If any anomaly is observed on road, video surveillance can be used to detect such anomaly by making an intelligent transportation system (ITS). ITS automatically monitors and categorizes the perceived behavior of traffic flow and consequently triggers alerts. Such a system can be useful in effectively for sending a warning signal automatically to the control room to take preventive actions to avoid road crash or to mitigate its effects. The video surveillance system can be categorized into two systems: (i) a basic system, and (ii) a smart system. The basic system (Diehl, 2000) is commonly used for recording or monitoring purposes. Whereas, the smart system (Thomas et al., 2017) is used for automatically detecting the abnormal events by analyzing the behavior of targets using CV-based algorithms. The potential strengths related to the implementation of video surveillance system on roads are huge to enhance road safety in both *reactive* and *proactive* ways (Cocca et al., 2016). In the *reactive* approach (Kongsvik et al., 2012), video technology may be useful to collect road crash data. The availability of huge amount image data for a road crash may help an analyst to understand the nature of road crash and identify the most effective corrective actions or apply traffic rules to avoid its re-occurrence (Conche and Tight, 2006). On the other hand, the video surveillance system may be useful to augment road safety in a *proactive* manner (Mackenzie et al., 2002), i.e., not waiting for a road crash to happen but rather trying to detect the anomalous traffic pre-events prior to their occurrences. Hence, the video surveillance-based proactive approach for road safety management is more useful than the reactive one.

The proactive safety management-based approach depends on behavior-based safety (BBS) method (Matthews, 1997), automatic detection of anomaly. In the BBS method, subjects (i.e., people) are enforced to train about safe practices by analyzing the behavior of traffic anomaly and assessing the risk, thereby reducing the probability of traffic pre-events and hence road crash. As video clips are the rich source to examine safety performance and provide many advantages to live observation, CV-based algorithms can be used to model the pre-events that can cause the road crashes. Videos can be used for training process to explore the behavior of un-natural traffic scenarios. As a consequence, traffic rules can be modified to reduce the number of road accidents/crashes. However, there exists a possible drawback of applying video surveillance in road safety management system. The higher the number of monitored areas are present, the larger the number of resources are needed to implement this system which eventually increases the complexity in managing road safety.

As already discussed that road safety management system mainly depends on the correct detection of traffic pre-events (i.e., anomaly), and adaptation of corrective actions after detecting these traffic pre-events. Anomaly can be categorized into behavior-based anomaly (Raptis and Sigal, 2013), and scene-based anomaly (Thomas et al., 2017; Pramanik et al., 2019). Anomalies can be related to vehicles, pedestrians, their interactions, or environment. Anomaly related to either vehicle or pedestrian is termed as a *behavior-based anomaly*. Whereas, interactions between vehicle-vehicle, vehicle-pedestrian, pedestrian-environment, and vehicle-environment are termed as a *scene-based anomaly*. As already discussed, road accidents happen due to neglecting the occurrence of pre-event(s). Detection of such accidents is very domain-specific. Based on the characteristics of road (i.e., either straight road or intersection), the causes of such accidents may vary. In case of road intersection, crashes mainly happen due to the presence of several pre-events, for examples, traffic signal violation, overtaking, wrong turn, illegal parking, and so on. Whereas, in case of straight road, crashes mainly occur due to the existence of pre-events, such as speed violation, overtaking, one-way traffic, illegal parking, wrong drop off location of passengers, etc. Although a large number of studies have been carried out either on road intersection or on straight road for pre-events detection, there still remains a dearth of studies on the improvement of road safety management system by analyzing the pre-events. However, it is very difficult to define a generic system to capture different types of traffic pre-events either on straight road or on road- intersection. One example of such a straight road is a highway. With the rapid development of our country's highway construction, highway traffic management scale expands rapidly; thereby requiring a higher level of highway traffic safety. Traffic pre-events, such as 'speed violation', 'overtaking', 'one-way traffic', 'illegal parking', and 'wrong drop off location of passengers' are majorly attributed to the occurrence of crashes in the straight road or highway environment. Although a lot of researches have done for many of the aforesaid traffic pre-events, but these are restricted to either simple traffic flow or simulated data. To address this issue and to enhance the road safety, we have considered the detection of aforesaid traffic pre-events as a research topic. As per the authors knowledge, detection of 'wrong drop off location of passengers' is a new research topic in the domain of road safety.

Pre-events occurred in the straight road belong to the behavior-based anomaly. Such behaviors are categorized as irregular behavior (Zhang and Liu, 2007), uncommon behavior (Wiliem et al., 2008), unusual behavior (Jiang et al., 2009), and abnormal behavior (Kratz and Nishino, 2009). Inference about "what is going on" in a scene leading to the behavior classification is drawn by analyzing several object-level features, including pose, movement, and gesture. The similar type of behavior can be considered as either normal or abnormal based on the context of scenario. For instance, the speed limit of moving vehicle in highway is different than that of the urban road. Therefore, in a real-time scenario, the nature of abnormality for any kind of event may vary from one frame to another frame of any given video. This results in a lack of labeled data for training of the CV-based model. Under such circumstances, anomalous traffic event detection is very challenging. Therefore, in most of the cases, one needs to adopt unsupervised learning for modeling the normal scenario. Several factors, such as anomaly class information, nature of anomaly, and availability of data are useful to develop a video surveillance system for the detection of various traffic pre-events.

In the present study, we have defined a conceptual framework for improving road safety. This framework is intended to model different traffic pre-events (probable cause of road crash), and then take corrective action to prevent the occurrence of a road crash. Different learning algorithms are developed for the detection of traffic pre-events, namely speed violation, one-way traffic, overtaking, illegal parking, and wrong drop-off location of passengers. We have named these algorithms as speed violation detection (SVD), one-way traffic detection (OTD), overtaking detection (OD), illegal parking detection (IPD), and wrong drop off location of passenger detection (WDLPD). These algorithms are integrated within one surveillance system. Therefore, these five types traffic pre-events can be detected using only one system instead of using five different systems. These algorithms are trained using the characteristic features of traffic pre-events, events, and post-events. Pre-event is an insight into the scenario that define the possible causes of the event; whereas event is a type of incident that happens in the road, and post-event is a consequence of the incident. All these events can be detected by analyzing the behavior of trajectory for the OoI. Here, OoIs represent the foreground objects that are present in the video frames. Trajectory conveys the position of an object over a sequence of frames. Both spatial and temporal information are used to form the trajectory of a detected object in video frames. By automatically detecting the pre-events, road safety management system should activate the alarm, and send this alarm to the control room for taking immediate action; thereby augmenting the level of road safety. Therefore, it significantly reduces the human effort.

As said earlier, our conceptual framework is based on correct detection/modeling of traffic anomaly (pre-events) and adaptation of corrective actions. Until the traffic anomaly can be detected, suitable action cannot be taken place. Therefore, the effectiveness of this framework is primarily based on the correct detection of traffic pre-events. Accurate detection of traffic pre-events is based on the aforesaid developed algorithms that are integrated within the surveillance system. To prove the effectiveness of these developed algorithms, several comparisons with the state-of-the-art methodologies have been demonstrated over the real-time videos acquired from a plant. We have also used another two datasets, namely CamSeq01 (Cambridge, 2007) and ISLab-PVD (Jo et al., 2017) to demonstrate the effectiveness of OD and IPD algorithms for the detection of overtaking and illegal parking, respectively. As the aforesaid two datasets are only accessible freely, they are used in our study. For two traffic anomalies, namely speed violation and one-way traffic, no such dataset is freely available. Since there is no literature present on the detection of "wrong drop off location of passengers", no comparative study is possible to carry out for this anomaly detection.

The rest of the article is organized as follows: Section 2 presents related works on this domain. A conceptual framework for road safety management is described in Section 3. Proposed methodology for the detection of traffic pre-events is illustrated in Section 4. A case study is presented in Section 5. In Section 6, results are discussed. Finally, Section 7 concludes this study with limitations and scope for the future works.

2. Related works

Anomalous traffic pre-events, such as speed violation, one-way traffic, overtaking, illegal parking, and wrong drop-off location of passengers are based on the abnormality in vehicular behavior. To analyze the behavior of on-road vehicles, vehicle detection and tracking are prerequisites. A vehicle is considered as a category of object. Therefore, first, we review the recent approaches used for object detection and tracking, which is described in Section 2.1, and then, we review the relevant existing approaches used for vehicular behavior-based anomaly detection, which is presented in Section 2.2.

2.1. Object detection and tracking

Relevant literature on object detection and tracking are reported in the following Sections 2.1.1 and 2.1.2, respectively.

2.1.1. Object detection

Existing literature on object detection are categorized into deep learning (DL) (Girshick, 2015; Pramanik et al., 2021b) and image processing (IP)-based approaches (Blanc et al., 2007). Of them, DL-based approaches become the most popular in recent years. Most popular DL-based approaches for object detection are RCNN (Girshick et al., 2014), YOLO (Redmon et al., 2016), and their variants. RCNN is a two-stage detector, whereas, YOLO is a one-stage detector. RCNN and its variants first predict probable object region in the image and then this region is used for object classification. On the other hand, YOLO and its variants are used to classify objects without region proposal task. Two-stage detectors achieve higher detection accuracy. Whereas, one-stage detectors achieve high inference speed. The function of these detectors (i.e., DL approaches) depends on the computational power of graphical processing units (GPUs), which may not always be affordable. Moreover, in DL-based approaches, labeled data is required for training of an algorithm. However, there is a lack of labeled data in the real-time traffic monitoring system. In contrast, IP-based approaches for object detection and tracking can be easily run using the Central Processing Unit (CPU) only. Moreover, IP-based approaches can enhance detection and tracking results by adopting various pre-processing tasks before the feature extraction and classification tasks. Therefore, IP-based approaches are relatively preferable in object detection in real-time scenario. In principle, object detection has two steps: (i) defining foreground region proposal, and (ii) recognizing objects present in the proposed regions. Background subtraction is used in various studies for obtaining foreground regions over the image/video frame (Cheon et al., 2012; Stauffer and Grimson, 1999). Most of them are eventually based on measuring of a local symmetry of an image patch or difference in multiple images (Ha and Lee, 2010). In Pramanik et al. (2020), object regions are extracted using Histogram of gradients (HoGs) of convolution features and statistical application on the gradient data of different feature maps. Local symmetry can be defined as a spatial/color/-edge/temporal similarity of pixels within an image/video frame.

In recent years, there has been a transition observed from a simpler image feature (i.e., edge) generation to a robust feature generation for background modeling. These feature sets are common in computer vision literature and allow for direct classification and detection of objects in images. HOGs (Cheon et al., 2012) and Haar-like (Sun et al., 2006) features are extremely well represented in the object (e.g., vehicle) detection literature. HOG features (Cheon et al., 2012) are extracted by first evaluating edges over the image and thereafter, discretizing and binning the orientations of the edge intensities into a histogram. This histogram is used as a feature vector. Although the HOG features (being the descriptive image features) exhibit a good amount of detection performance in a variety of computer vision tasks, including vehicle detection; they are, however, found slower in computation. The most popular background modeling method is the Gaussian Mixture Model (GMM) (Stauffer and Grimson, 1999), which helps in modeling of pixel values over a span of time by a weighted mixture of Gaussian. However, it is indeed a difficult task to achieve good results by applying the GMM over real-time scenarios. On the other hand, foreground regions can be obtained in the image/video frame using granulation methods (Chakraborty and Pal, 2017; Chakraborty et al., 2013). Granulation is the method of forming clusters/granules within the image. Granules could be of either equal or unequal sizes. For object detection, unequal sized granules are useful to obtain reliable contents (i.e., object regions). The effectiveness of granulation for the foreground region proposal is enumerated more in details in Chakraborty and Pal (2017). Region growing based on the local symmetry of pixels is done here for finding clusters or granules within the image/video frame. These granules represent the foreground regions. This technique is robust and unsupervised. Therefore, in this study, the concept of granulation is used for object detection in video frames. The concept of the study from Chakraborty et al. (2013) is adopted here for object detection. After detection, detected objects are used for tracking. Recent literature on object tracking are discussed in the next section.

2.1.2. Object tracking

Object tracking follows object detection. There are several approaches for object tracking mentioned in earlier studies. Depending on the object representation, which ranges from pixel level to feature level to object level, tracking approaches are categorized into five classes, namely: (i) deep learning-based (Wojke et al., 2017), (ii) region-based (Yilmaz et al., 2006), (iii) contour-based (Saunier and Sayed, 2006), (iv) model-based (Sobral and Vacavant, 2014), and (v) feature-based (Barth and Franke, 2009). In recent years, deep learning-based approaches (Wojke et al., 2017; Shen et al., 2018) for multi-object tracking are becoming popular with great success. However, these are restricted to the computing power of GPUs and they are semi-supervised in nature. Whereas, real-time tracking is unsupervised. One of unsupervised tracking is region-based approach (Yilmaz et al., 2006). Here, regions can be defined as connected image parts with distinguishing common properties, such as intensity, color or texture statistics. Region-based tracking aims at tracking objects (e.g., vehicles) according to the variations of regions in images (Yilmaz et al., 2006). However, this approach is restricted to the congested traffic conditions, complex deformation or cluttered background, where objects (i.e., vehicles) are partially occluded by one another instead of being spatially isolated. Another unsupervised tracking is contour-based approach that represents objects by their contours, which are nothing but their boundaries, and updates these contours dynamically at each time increment (Saunier and Sayed, 2006). Contour-based tracking offers a more efficient description of the object (e.g., vehicle) than region-based tracking by reducing both the computational time, and the complexity. However, this tracking approach is unable to solve the occlusion problem.

Other unsupervised tracking is model-based approach that matches the projected models (i.e., detected objects) in the current frame with the suitable existing trajectories. This characteristic allows to recover trajectories and models and in particular, the pose of the object (e.g., vehicle) with higher accuracy (Sobral and Vacavant, 2014). The major weakness of this approach is the need for an accurate geometric object model. It is, however, unrealistic to obtain such models for all moving objects in traffic condition. Therefore, additional cues/features are usually added to the projected models. Feature-based tracking utilizes the principle that object can be represented by a set of features, instead of an entire region/contour/model. This refers to the group of methods that perform tracking first by extracting features from the independent image frame and then, matching those features over the frames. Here, features can be selected as representative parts of the object (e.g., vehicle), such as corners, lines, or any typical shapes. This technique is effective as long as the selected features are robust and can be distinguishable, even if the object remains partially occluded at some points in the video sequence. In this technique, Bayesian algorithms are used for proper data association over the extracted features (Moqqaddem et al., 2011; Barth and Franke, 2009). Data association is the allocation of the detected object (in the current frame) to a trajectory based on similar features. Some popular algorithms for data association include traditional Kalman filter (Moqqaddem et al., 2011) and extended Kalman filter (Barth and Franke, 2009). By reviewing all the aforesaid approaches, feature-based approaches for object tracking is considered as the state-of-the-art in this study.

After detection and tracking of objects, object-level features are analyzed for anomaly classification. The literature of vehicular behavior-based anomaly detection is stated in the following section.

2.2. Vehicular behavior-based anomaly detection

Analysis of vehicular behavior is, in fact, a pressing challenging of late. It can be categorized into context, manoeuvers, and trajectories. Context-based approaches are related to types of road, i.e., whether the anomalies are detected at either urban road or highways. In Sivaraman et al. (2011), a context-specific spatio-temporal model is developed to capture the anomaly in highway driving. In this approach, clustering is done over the observed trajectories of vehicles in highways. In Cherng et al. (2009), a dynamic visual model is developed of the driving environment with saliency alerting the system to unusual and critical on-road conditions. Histograms of scene flow vectors (i.e., Spatio-temporal information) are used in Geiger and Kitt (2010) to classify the driving environment into intersection and non-intersection driving. The motion model is used in Jazayeri et al. (2011), which models the distribution of vehicle in image plane using the prior information of vehicle detection.

In manoeuvre-based approach, the term, 'manoeuvre' refers to the behavior of an on-road vehicle (e.g., overtaking, illegal parking, etc.). On-road vehicle behavior is modeled in Gindele et al. (2010) as a Markov process and inferred using a dynamic Bayesian network using tracking observations. However, the experimental evaluation is performed using simulation data. In Alonso et al. (2008), overtaking behavior is detected by identifying vehicles in the blind spot of the ego vehicle (i.e., automated driving vehicle). Overtaking behavior is detected in Zhu et al. (2006) for vehicles in front of the ego vehicle. Another anomaly is illegal parking that can result in a traffic jam. Several studies have been carried out on illegal parking (Stauffer and Grimson, 1999; Zivkovic, 2004; Sun et al., 2006), most of which are based on the separation of either foreground or background. The study by Stauffer and Grimson (1999) explores the extraction of foreground vehicles by using background modeling strategy, however, this method is restricted to simple traffic only. In Zivkovic (2004), foregrounds are extracted using a Gaussian mixture model, though it is unable to address the occlusion problem. In Sun et al. (2006), a scalable histogram of gradient features extraction followed by support vector machine (SVM) classification is used for foreground modeling; however, this is severely affected by weather condition. All the aforesaid methods used for the detection of either overtaking or illegal parking are restricted to the simple traffic flow only.

Trajectory-based approaches are mostly used for detecting abnormal events as compared to aforesaid two approaches. In Wiest et al. (2012), variational Gaussian mixture modeling is used to classify and predict the long-term trajectories of vehicles by using simulated data. In Rad et al. (2010), long term trajectories are analyzed to detect both speed violation and one-way traffic. However, this technique is restricted to the simple traffic flow only. Although various algorithms are developed for different traffic pre-event detection. These are restricted to temporal information only. But spatio-temporal information is more effective in real-time scene description as compared to temporal information only (Chakraborty and Pal, 2017). Therefore, in our study, we have defined various rules using spatio-temporal information for the detection of different traffic pre-events. Based on the above-mentioned studies discussed, some research issues and contributions from our studies are identified and mentioned below.

2.3. Research issues

(i) Research on the application of video surveillance system for the improvement of road safety is limited.
(ii) Research on traffic pre-event detection using spatio-temporal information is still not much explored.
(iii) Most of the approaches are context-specific. For example, algorithm designed for the detection of overtaking cannot be used for the speed violation.
(iv) Most of the approaches are restricted to either simple traffic flow or simulated data.
(v) Two or more traffic pre-events cannot be detected using a single system. The reason is that each system consists of a task-specific single algorithm. Therefore, it is essential to develop such a system that effectively combines two or more algorithms so that the system cannot be restricted only to a particular context.
(vi) Most of the researches on implementation of video surveillance system for road safety are based on reactive approach which is less effective than proactive approach.

2.4. Contributions

(i) We have developed a conceptual framework for a video surveillance system to enhance the level of road safety in an automated way. This framework includes five steps: (a) analyze the behavior of traffic pre-events, (b) identification of places for installing the Closed-circuit television (CCTV), (c) development of algorithms for modeling the traffic pre-events, (d) training and validation of the developed algorithms, and (e) test operation. The test operation consists of two basic stages: *first*, testing is done over live stream video and *second*, take corrective actions after detecting the traffic anomaly in this video. This framework is generic.
(ii) A set of five algorithms, namely SVD, OTD, OD, IPD, and WDLPD are integrated within one surveillance system to detect traffic pre-events, such as speed violation, one-way traffic, overtaking, illegal parking, and wrong drop off location of passengers.
(iii) A set of features and rules are generated for the detection of aforesaid traffic pre-events. Features used in each pre-event detection may not be the same always. Therefore, based on the nature of pre-events, we have defined a set of features. Using these features, we have further defined a set of rules to detect the traffic pre-events/anomalies in more efficient ways.

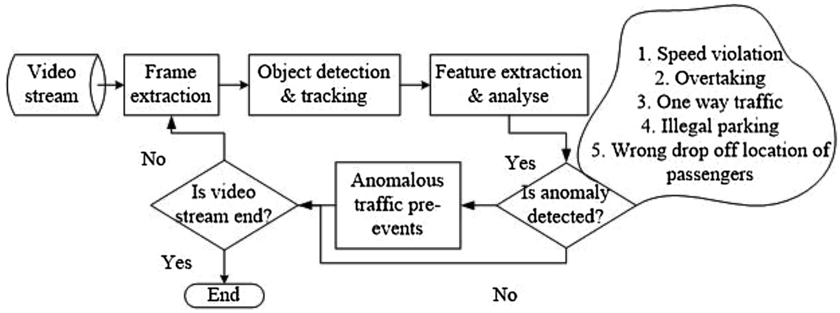


Fig. 1. Methodology of traffic pre-event detection.

(iv) These algorithms are based on the combination of both spatial and temporal features which is much informative than considering either spatial or temporal feature.
(v) All these algorithms are unsupervised in nature. As per author's knowledge, the research on wrong drop off location of passengers is new in road safety domain.
(vi) Implementation of video surveillance system to enhance the road safety is based on proactive approach which is much effective than reactive approach.

3. A conceptual framework for traffic surveillance system

A conceptual framework for traffic surveillance system is developed to enhance the road safety. This conceptual framework consists of five stages: (i) analyze the behavior of traffic pre-events, (ii) identification of places for installing the CCTV, (iii) development of algorithms for traffic pre-events modeling, (iv) training and validation of the developed algorithms, and (v) test operation. These phases are explained below.

3.1. Analyze the behavior of traffic pre-events

First of all, it is necessary to understand the types of anomalies occurred on roads. For the analysis of traffic pre-events, a multidisciplinary team is initially formulated. The team consists of safety professionals and technical experts. Safety professionals have thorough knowledge about the traffic pre-events, and technical experts have expertise in video processing. Safety professionals should know the way by which videos are collected, and the nature of each traffic pre-events. Based on these two sources of knowledge, technical experts can develop algorithms for modeling traffic pre-events/anomalies. Then, by analyzing the traffic pre-events detected, safety professionals can identify the unsafe conditions and behaviors, and correct them to prevent incidents and injuries. Since unsafe condition or behavior can be considered as the predecessor/precursors of future incidents, their timely detection and correction are critical for traffic safety improvement. Here, unsafe behavior is defined as an action, such as a procedure, a task or an activity that does not conform to general operation process. It may result in property damages or personal injuries. Whereas, the unsafe condition is a state characterizing an environment with a potential to cause damage. Two examples of unsafe behavior include high speeding, and one-way traffic; whereas, low visibility and rainy weather belong to unsafe condition.

3.2. Identification of places for installing the CCTV

In the next phase, places are identified, where CCTVs should be installed to monitor the traffic scenarios. CCTV is nothing but the digital camera that can capture the safety-relevant targets. Therefore, based on the requirement, safety professionals should decide the number of cameras to be installed. All the cameras are linked with one storage device which depends on the number of video streams to be recorded, the duration for video data collection, and the video quality.

3.3. Development of algorithms for modeling of traffic pre-events

The third phase is the modeling phase. In this phase, the behavior of anomalous traffic pre-events is modeled. Any event that differs from the normal behavior is termed as an anomaly. Such traffic pre-events that are modeled in this study are speed violation, one-way traffic, overtaking, illegal parking, and wrong drop off location of passengers. We have developed a methodology for detecting/ modeling these anomalies. This is explained in Section 4.

3.4. Training and validation of the developed algorithms

Next phase is the training and validation of the algorithms that are developed under the proposed methodology. Training and validation are done before applying the methodology in the final operation. A total of 300 videos are extracted from (24×7) video-stream and used for the training process. Further, these algorithms are validated using different 132 videos.

3.5. Test operation

The operation has two steps: (i) test over live stream video and (ii) take corrective actions after detecting the traffic anomaly. After training and validation, initially, the methodology is used for testing operation over 132 videos. After obtaining the promising testing accuracy, this methodology is tested over (24×7) live stream-videos using camera IP address for automatic identification of traffic pre-events. Then, a list to support the testing results is generated automatically. The list contains a set of information about each observation, namely the location, timestamp, and type of the pre-event. At each location, one camera is set. Therefore, using the location information of the camera, the video surveillance system can predict the location-stamp of the event. Similarly, using the time-stamp information of camera, the system can predict the time-stamp of the event. Our methodology involves several algorithms that are used to analyze the behavior of traffic pre-event. Based on the analyzing results, the methodology can predict the type of pre-events.

It is noteworthy to mention that the video-based analysis takes into account the pre-event scenarios, thereby enhancing the capability to recognize the upcoming event. Once the most safety-critical targets (pre-events) are identified automatically, an alarm is generated in the control room to take preventive measures. Then, safety professionals should take corrective action to prevent the probable road crash at the location of detected traffic pre-events. In this way, road safety can be enhanced. From the aforesaid discussion, it is evident that the effectiveness of this conceptual framework depends on the correct modeling of traffic pre-events. Methodologies for detecting five types traffic pre-events are stated in the next section.

4. Methodology for detecting/modeling of traffic pre-events

As mentioned earlier, we have developed a conceptual framework for the traffic surveillance system to improve road safety. The main objective of this system is the correct detection of traffic pre-events/anomalies and then, corrective actions taken by safety professionals. After the detection of a traffic anomaly (i.e., pre-event), a video summary corresponds to this anomaly is generated from the live-streaming, and stored in the database. This reduces the human effort for finding the video summary of corresponding traffic pre-events from live streaming. This video summary can be used in future for either training or inspection purpose. This task is done in real-time and repeated until the video stream ends. However, after detecting the anomaly, an alarm is automatically generated in the control room to take preventive measures for avoiding the road crash. The preventive measure should be taken in off-line mode and on-time to prevent the road incident; thus, enhancing road safety. From the aforesaid discussion it is evident that until any traffic pre-event is detected, no action can be taken by safety professionals. Therefore, the effectiveness of the traffic surveillance system primarily depends on the correct detection of traffic pre-events. Flowchart for the detection of traffic pre-events is shown in Fig. 1.

From Fig. 1, it is seen that the frames are extracted from the video stream. Then, object detection and tracking are done over these frames. This is the prerequisite of this study. Thereafter, object-level features, namely velocity, area, and position are extracted for each detected object in the current frame. These features are analyzed to identify anomalous traffic pre-events from the video stream. Methodology of object detection & tracking and case-specific traffic pre-event detection are stated in the following Section 4.1.

4.1. Object detection and tracking

Object detection refers to the scanning and searching for an object in a frame or video; whereas, in object tracking, objects are tracked solely based on their trajectories. Object detection is followed by object tracking. Methods for object detection and tracking are discussed in Sections 4.1.1 and 4.1.2, respectively.

4.1.1. Object detection

In this study, spatio-temporal neighborhood granulation is done to detect moving objects in video frames. Granulation is nothing but a clustering technique. Combination of spatio-neighborhood and temporal granules (clusters) represents object regions more effectively than either considering spatial or temporal granule. As spatio-neighborhood granules represent different sized static objects pixels and temporal granules represent moving objects pixels. Due to the presence of noise, some undesired pixels (background) may belong to temporal granules that can create ambiguity in defining moving objects. Therefore, the commonality between these two granules truly represents moving objects regions in video frames. The concept of Chakraborty et al. (2013) is adopted in our study for generating the spatio-temporal neighborhood granules for object detection. This method is unsupervised and provides effective results in video object detection. It is not restricted to only simple traffic flow, and can be easily run using the Central Processing Unit (CPU) which is much affordable.

Object detection can be done on both gray image and color image. Gray image represents one intensity for each pixel location, whereas, color image presents three intensities for each pixel location. Although gray image is good for many image processing task, but sometimes, it could not represent the picture character (i.e., overlapped objects). The core part of moving object detection is image segmentation task. Better the segmentation mean higher the detection accuracy. Meaningful granulation is very important to generate the granules which truly represent the natural scenario. Crisp granules are formed using gray image. Crisp granules make the computation much faster. But in real-life applications, information system is not always crisply separable rather overlapping in nature. This issue can be deal with spatio-color neighborhood granules. Both spatial and color nearness are used to form spatio-color neighborhood granules which represent static object regions in still images. Temporal granules represent moving object regions in video frames. The similarities between both spatio-color neighborhood and temporal granules are considered during the formation of spatio-temporal granules which have arbitrary shape and size. It is proved in Chakraborty and Pal (2017) that spatio-temporal granules lead to meaningful segments with different shapes and sizes as compared to gray level image segmentation. Therefore, in this study, we have used color image-based spatio-temporal granules for moving object detection.

Here, spatio-color granules are formed using the spatial and color similarity of pixels in a frame. The main idea is: the difference between the maximum and minimum pixel intensities in a granule (cluster) should be greater than a certain threshold T, which can be defined as $T = (Q_3 - Q_1)/2$. Here, Q_3 and Q_1 denote the third and first quartiles of pixel-level distribution, respectively. In this process, a (3×3) window slides over the video frame (F_t), and it is also checked whether the aforesaid criterion is matched or not. If the criterion is matched, then, this spatial window containing pixels are considered as foreground pixels. Thereafter, the region of this spatial window grows by grouping its neighborhood windows which have nearly equal average intensities. This results in spatio-color neighborhood granules (G_s) in the video frame F_t. Any two windows, say, x and y in spatio-color neighborhood granule are related to the Eq. (1), as defined:

$$G_s = (x, y) \in F_t | I(x) - I(y)| < T_1 \quad \&(x, y) \text{ are8} - \text{connected} \tag{1}$$

where $I(.)$ denotes the average intensity of the window $(.)$, and T_1 indicates the threshold used for grouping the neighborhood windows. Temporal granules are formed using the three-point estimation (Chakraborty et al., 2013). This method is adopted as it is occlusion-free and is not affected by low illumination. In this process, a set of N frames are processed at a time. Based on the information within the frames, three points, such as A, B, and C are estimated, which are defined in the following Eqs. (2), (3), and (4), respectively.

$$A = \max(F_1, F_2, ..., F_t, ..., F_N) \tag{2}$$

$$B = \text{median}(F_1, F_2, ..., F_t, ..., F_N) \tag{3}$$

$$C = \min(F_1, F_2, ..., F_t, ..., F_N) \tag{4}$$

In addition, another two information, mean/background (μ) and standard deviation (σ) are also estimated for background modeling. Both μ and σ can be defined by the following Eqs. (5) and (6) (Chakraborty et al., 2013):

$$\mu = (A + 4B + C)/6 \tag{5}$$

$$\sigma = (A - C)/6 \tag{6}$$

Any pixel (p) in the video frame is considered as foreground, if it follows the criteria, as expressed in Eq. (7):

$$p = \begin{cases} \text{foreground}, & \text{if} \quad (p - \mu) > 3 \times \sigma \\ \text{background}, & \quad \text{otherwise}. \end{cases} \tag{7}$$

The commonality between the spatio-color neighborhood and temporal granules is considered as foreground object in a video frame. After detecting the objects, bounding boxes are fitted over it and are used for tracking. The tracking process is described in the next section.

4.1.2. Object tracking

After detection, object tracking is done. Though object tracking can be done using spatio-temporal granulation, a fine-tuning is required for obtaining robust tracking, in case of complex scenarios. Here, we have used the concept of feature-based tracking. Our tracking process has two steps: the first step is the state-space estimation and the second step is the data association. In the first step, Kalman filter prediction of states of detected objects (track-lets) in 8-dimensional state space is done. As said earlier, detected objects are fitted with bounding boxes. This state-space

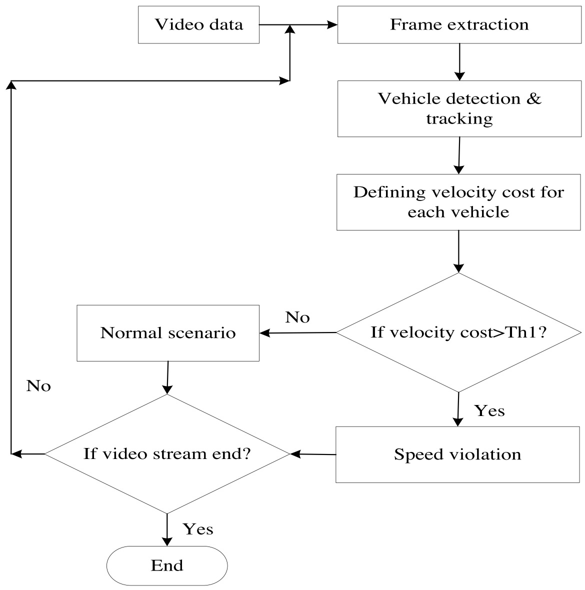


Fig. 2. Flowchart for the detection of speed violation.

contains the information of detected bounding boxes. The 8-dimensional state information includes the centre position of the bounding box in image coordinates, its aspect ratio and height, and their respective velocities. Thereafter, proper data association is done using the 8-dimensional state information of detected objects to form the trajectories. Proper data association means each object (detected objects in the current frame) gets its proper association to a trajectory. A trajectory contains the spatial information of an object over frames. Each spatial information in the trajectory is called track-let. Let O_i, T_j, and M be the i^{th} detected object in the current frame, j^{th} track-let, and the total number of track-lets in the previous frame, respectively. Then, O_i is associated with its nearest track-let present in the previous frame using Eq. (8).

$$O_i \rightarrow T_j, \text{if} \quad \text{dis}(O_i, T_j) = \min\{\text{dis}(O_i, T_1), \text{dis}(O_i, T_2), ..., \text{dis}(O_i, T_M)\} \tag{8}$$

where $\text{dis}(O_i, T_j)$ defines the distance between O_i and T_j. After object detection and tracking, object-level features (OLFs), such as speed/velocity, area, position, and orientation are extracted for the detected object in the current frame. Thereafter, these features are used for modeling the abnormal traffic pre-events. Characteristics of traffic pre-events are defined in the next section.

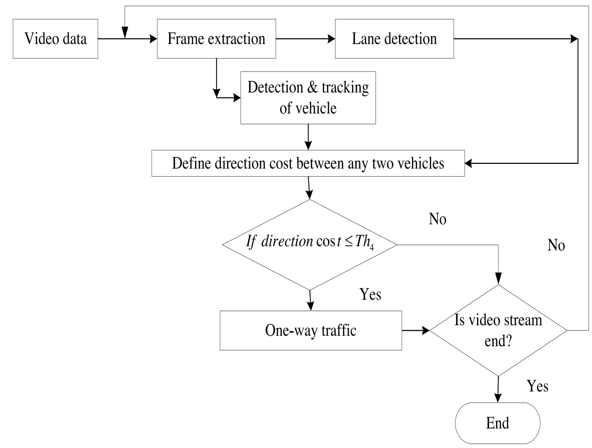


Fig. 3. Flowchart for the detection of one-way traffic.

4.2. Traffic pre-events detection

Traffic pre-events that are considered in this study are (i) speed violation, (ii) overtaking, (iii) one-way traffic, (iv) illegal parking, and (v) wrong drop-off location of passengers. It is already stated that object detection and tracking is the prerequisite of traffic pre-event detection. After the detection of an object, bounding-box is fitted over the detected object. The spatial and temporal information of this bounding-box, such as position, velocity, direction, and area are used as features for further operation. Methodologies for all the aforesaid traffic anomalies (i.e., pre-events) are briefly described in the following sections:

4.2.1. Speed violation

One example of traffic rules violation is speed violation, which is referred to as high speeding. Dynamic feature, such as velocity is extracted using the trajectory information of the detected object and it is used for the detection of speed violation. The flowchart of speed violation is depicted in Fig. 2. From this figure, it is seen that the frames are extracted from the video stream. Then, during the detection and tracking task, bounding-box is fitted over the detected object. Coordinate of the centroid of such box is used to obtain the velocity of the detected object. Velocity of each detected object (vehicle) is obtained in terms of pixels per second. This is done during object tracking. The velocity of a tracked object can be defined by the following Eqs. ((9, 10), (11)):

Algorithm 1 Pseudo-code of speed violation.

Input: $\{X^{t-1}, Y^{t-1}\} = \{(x_1^{t-1}, y_1^{t-1}), (x_2^{t-1}, y_2^{t-1}), ..., (x_n^{t-1}, y_n^{t-1})\}$: A set of centroid co-ordinates for n detetcted objects at $(t-1)^{th}$ frame; $\{X^t, Y^t\} = \{(x_1^t, y_1^t), (x_2^t, y_2^t), ..., (x_m^t, y_m^t)\}$: A set of centroid co-ordinates for m detetcted objects at $(t)^{th}$ frame; and threshold $= Th_1$.

Output: $\{SV\}^t$: A set to store the position of vehicles at t^{th} frame that violate speed limit.

```
Initialize the set {SV}^t ← φ
for i=0 to m do
    compute v_i^t for t^th vehicle using Eq. (11)
    if v_i^t > Th_1 then
        {SV}^t ← (x_i^t, y_i^t)
    end
end
return {SV}^t
End
```

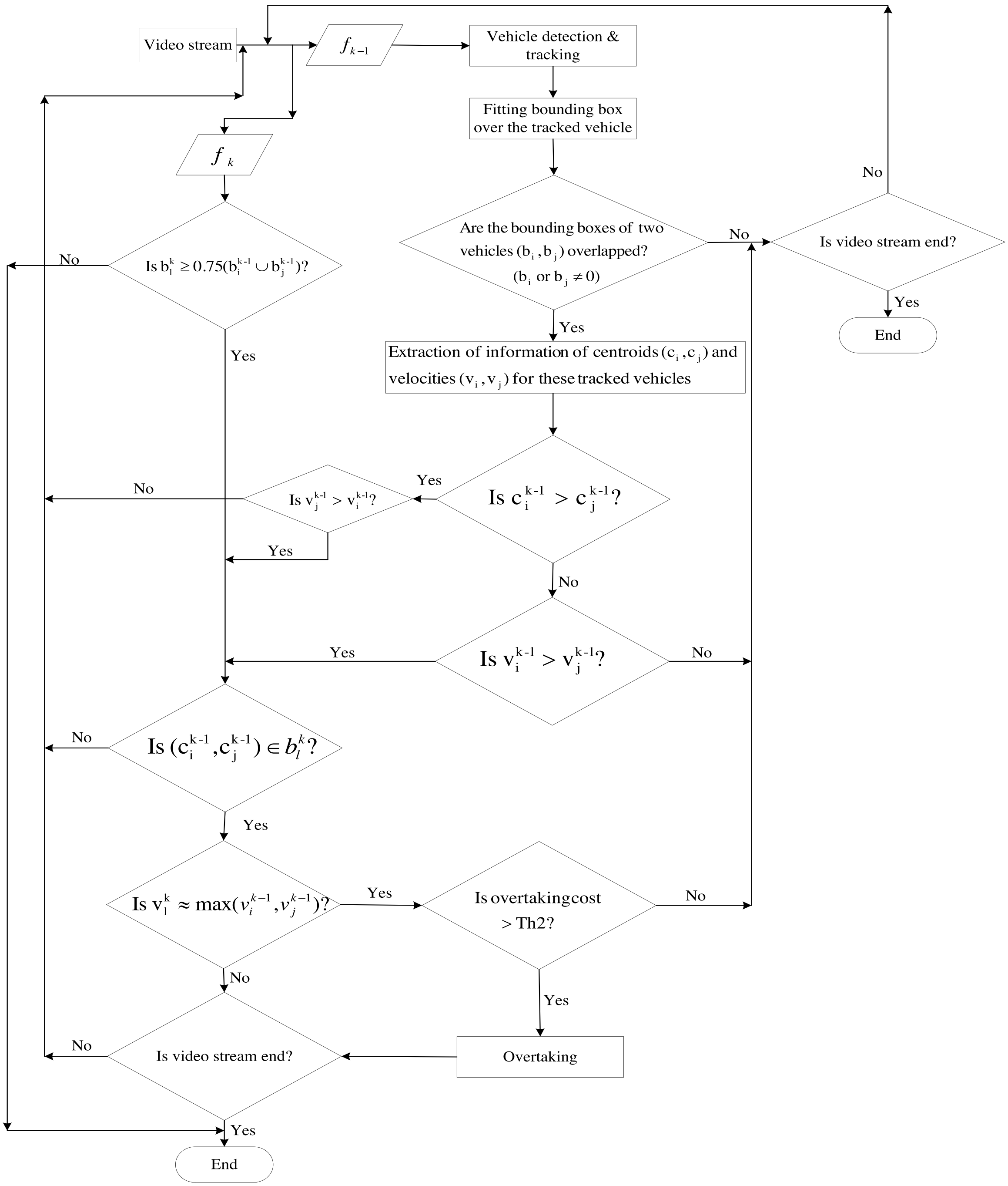


Fig. 4. Flowchart for the detection of overtaking.

$$vx_i^t = |x_i^t - x_i^{t-1}| \times \text{fps} \tag{9}$$

$$vy_i^t = |y_i^t - y_i^{t-1}| \times \text{fps} \tag{10}$$

$$v_i^t = \sqrt{(vx_i^t)^2 + (vy_i^t)^2} \tag{11}$$

where vx_i^t and vy_i^t represent the velocity components of the i^{th} vehicle in horizontal (x)- and vertical (y)-axis, respectively at t^{th} time, x_i^t and y_i^t indicate the centre position components of the i^{th} vehicle in x- and y-axis, respectively at t^{th} time, and v_i^t denotes the velocity of i^{th} vehicle at t^{th} time. The algorithm detects speed violation if i^{th} vehicle is over speeding, i.e., speed (velocity) of the vehicle exceeds a pre-defined threshold, say Th_1. This algorithm is named as speed violation detection (SVD). The pseudo-code for the detection of speed violation is shown in Algorithm 1.

4.2.2. *One-way traffic*

One-way or oncoming traffic is another kind of traffic pre-event. In a one-way road, if vehicles come from two opposite directions, then, it is called one-way traffic. The flowchart for the detection of one-way traffic is illustrated in Fig. 3, and the algorithm used for this purpose is named as oncoming traffic detection (OTD). From Fig. 3, it is evident that object detection and tracking is done over the frames extracted from the video stream. Thereafter, the OTD algorithm defines the direction of movement of the vehicle. We have assumed that if a vehicle moves from the right side to the left side of a frame, then, the direction of movement is considered positive; otherwise, it is considered negative. Let $(\mathrm{vx}_i^t, \mathrm{vy}_i^t)$ and $(\mathrm{vx}_j^t, \mathrm{vy}_j^t)$ be the velocity components of i^{th} and j^{th} vehicles along horizontal (x-axis) and vertical (y-axis) direction for t^{th} frame. Then, the direction of the velocity for both i^{th} and j^{th} vehicles are defined as $\theta_i = \tan^{-1}(\mathrm{vy}_i^t/\mathrm{vx}_i^t)$ and $\theta_j = \tan^{-1}(\mathrm{vy}_j^t/\mathrm{vx}_j^t)$ respectively. Here, first, we manually define the region of interest (RoI) as road detection over the video frames. Then, this RoI is divided into four equal regions. The centroid position of the RoI is considered as $(0, 0)$ and co-ordinates of the aforesaid four regions are revised accordingly. For one-way traffic detection, we have defined the direction cost $(\mathrm{dir}_{i,j}^t)$ between i^{th} and j^{th} vehicles. Direction cost is the difference between direction movements of any two vehicles. $\mathrm{dir}_{i,j}^t$ is expressed in the following Eq. (12).

$$\mathrm{dir}_{i,j}^t = \exp((\theta_i - \theta j)\Big/360) + \exp - (\sqrt{(x_i^t - x_j^t)^2 + (y_i^t - y_j^t)^2}) \tag{12}$$

If $\mathrm{dir}_{\mathrm{ij}}^t \geq \mathrm{Th}_4$, then it is considered as one-way traffic violation. The pseudo-code for the detection of one-way traffic violation is shown in Algorithm 1.

4.2.3. Overtaking

To detect overtaking, our proposed algorithm checks whether one moving vehicle is occluded by another moving vehicle or not. This algorithm is named as overtaking detection (OD). The flowchart for overtaking detection is shown in Fig. 4. From this figure, it is seen that bounding-boxes are fitted over the detected objects after detection and tracking task. Thereafter, 3-dimensional information (i.e., area (a), velocity (v), and centroid (x, y) of bounding-box) of each detected object are extracted. Let a_i^{t-1} and a_j^{t-1} be the area of bounding-boxes for i^{th} and j^{th} vehicles at $(t-1)^{\mathrm{th}}$ frame, respectively. Then, the common area (say, intersection over union (IoU)) shared by these two vehicles is considered as a feature. If there are some non-zero values for IoU, then, OD checks the velocity and centroid of these two vehicles for obtaining the overtaking cost function. If the area of IoU gradually increases over consecutive frames, then, there is a high possibility of the occurrence of overtaking. In any frame (say, t^{th} frame), if the two vehicles (say, i^{th} and j^{th} vehicles) are fully occluded by each other, then, one bounding-box is obtained at this location instead of two boxes. To ensure whether an overtaking occurs or not, we define three conditions and overtaking cost function, as stated below.

$$\begin{aligned}
&\text{Condition1a :}, \quad b_i^{t-1} \cap b_j^{t-1} \neq 0, _\text{d is}(\sqrt{(x_i^{t-1} - x_j^{t-1})^2 + (y_i^{t-1} - y_j^{t-1})^2}), \text{and} \quad v_j^{t-1} < v_i^{t-1} \\
&\text{Condition1a :}, \quad b_i^{t-1} \cap b_j^{t-1} \neq 0, _\text{d is}(\sqrt{(x_i^{t-1} - x_j^{t-1})^2 + (y_i^{t-1} - y_j^{t-1})^2}), \text{and} \quad v_j^{t-1} > v_i^{t-1} \\
&\text{Condition2 : If} b_l^t \geq (b_i^{t-1} \cup b_j^{t-1}), (x_i^{t-1}, y_i^{t-1}), (x_j^{t-1}, y_j^{t-1}) \in b_l^t, \text{and} \quad v_l^t \approx \max(v_i^{t-1}, v_j^{t-1})
\end{aligned} \tag{13}$$

Algorithm 2 Pseudo-code of one-way traffic.

Input: $\{X^{t-1}, Y^{t-1}\} = \{RoI\}$: A set of co-ordinates for lane; $\{X^t, Y^t\} = \{(x_1^t, y_1^t), (x_2^t, y_2^t), ..., (x_m^t, y_m^t)\}$: A set of centroid co-ordinates for m detetcted objects at $(t)^{th}$ frame; and threshold $= Th_4$.

Output: $\{OT\}^t$: A set to store the position of vehicles at t^{th} frame that violate the rule of one-way traffic.

Initialize the set $\{OT\}^t \leftarrow \phi$

for i=0 *to* m **do**
  if $Object_i^t \in \{RoI\}$ **then**
    for j=0 *to* m **do**
      if $i \neq j$ **then**
        if $Object_j^t \in \{RoI\}$ **then**
          compute dir_{ij}^t using Eq. (12)
          if $dir_{ij}^t > Th_4$ **then**
            $\{OT\}^t \leftarrow (x_i^t, y_i^t)$
          end
        end
      end
    end
  end
end

return $\{OT\}^t$

End

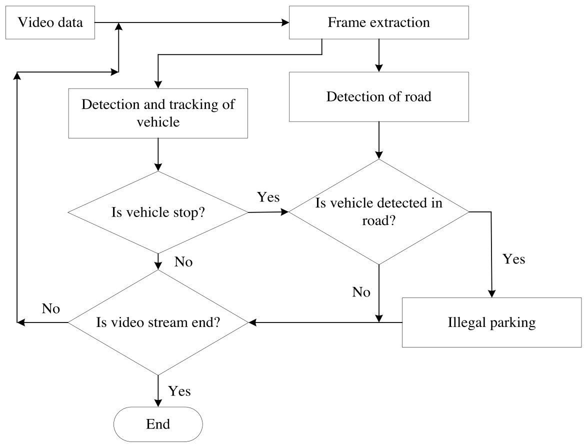


Fig. 5. Flowchart for the detection of illegal parking.

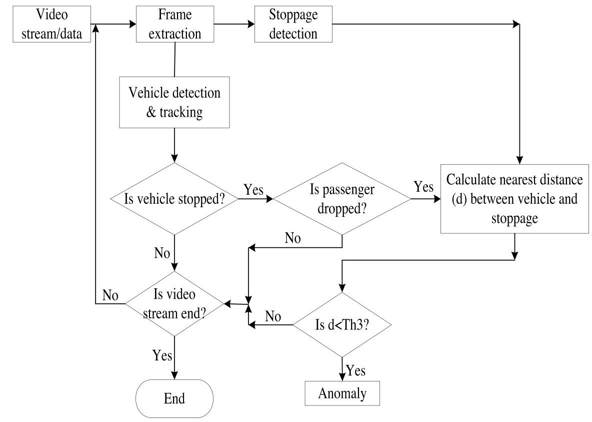


Fig. 6. Flowchart for the detection of the wrong drop-off the location of passengers.

Algorithm 3 Pseudo-code of overtaking.

Input: $\{X^{t-1}, Y^{t-1}\} = \{(x_1^{t-1}, y_1^{t-1}), (x_2^{t-1}, y_2^{t-1}), ..., (x_n^{t-1}, y_n^{t-1})\}$: A set of centroid co-ordinates for n detetcted objects at $(t-1)^{th}$ frame; $\{X^t, Y^t\} = \{(x_1^t, y_1^t), (x_2^t, y_2^t), ..., (x_m^t, y_m^t)\}$: A set of centroid co-ordinates for m detetcted objects at $(t)^{th}$ frame; $\{b^{t-1}\} = \{b_1^{t-1}, b_2^{t-1}, ..., b_n^{t-1}\}$: A set of bounding boxes over n detetcted objects at $(t-1)^{th}$ frame; $\{b^t\} = \{b_1^t, b_2^t, ..., b_m^t\}$: A set of bounding boxes over m detetcted objects at $(t)^{th}$ frame, $(i, j) \in n$.

Output: $\{OV\}_1^t, \{OV\}_2^t$: Two sets to store the position of vehicles at t^{th} frame that involve overtaking operation.

Initialize the set $\{OV\}_1^t \leftarrow \phi$
Initialize the set $\{OV\}_2^t \leftarrow \phi$
for *l=0 to m* **do**
  if $(b_i^{t-1} \cap b_j^{t-1}) \neq \phi$ **then**
    compute v_i^{t-1} and v_j^{t-1} using Eq. (11)
    if $((x_i^{t-1} > x_j^{t-1})$ *and* $(y_i^{t-1} > y_j^{t-1})$ *and* $(v_j^{t-1} > v_i^{t-1}))$ *or* $((x_j^{t-1} > x_i^{t-1})$ *and* $(y_j^{t-1} > y_i^{t-1})$ *and* $(v_i^{t-1} > v_j^{t-1}))$ **then**
      if $b_l^t \geq 0.75(b_i^{t-1} \cup b_j^{t-1})$ *and* $(x_i^{t-1}, y_i^{t-1}), (x_j^{t-1}, y_j^{t-1}) \in b_l^t$ **then**
        compute v_l^t using Eq. (6)
        if $v_l^t \approx max(v_i^{t-1}, v_j^{t-1})$ **then**
          compute O_{ij}^{t-1} using Eq. (14)
          if $O_{ij}^{t-1} > O_{ij}^t$ **then**
            $\{OV\}_1^t \leftarrow (x_i^{t-1}, y_i^{t-1})$
            $\{OV\}_2^t \leftarrow (x_j^{t-1}, y_j^{t-1})$
          end
        end
      end
    end
  end
end
return $\{OV\}_1^t$ and $\{OV\}_2^t$
End

$$o_{ij}^{t-1} = \exp - \left(\left(b_i^{t-1} \cap b_j^{t-1} \right) \Big/ \left(b_i^{t-1} \cup b_j^{t-1} \right) \right) + \exp - \left(\sqrt{\left(x_i^{t-1} - x_j^{t-1}\right)^2 + \left(y_i^{t-1} - y_j^{t-1}\right)^2} \right) \tag{14}$$

where $\mathrm{dis}\left(\sqrt{\left(x_i^{t-1} - x_j^{t-1}\right)^2 + \left(y_i^{t-1} - y_j^{t-1}\right)^2}\right)$ represents the distance between centroids of i^{th} and j^{th} vehicles at $(t-1)^{\text{th}}$ frame and v_i^{t-1} denotes the velocity of i^{th} vehicle at $(t-1)^{\text{th}}$ frame. v_i^{t-1} as defined in Eq. (11). To detect overtaking, we have defined two rules:

Rule 1: If conditions $1a$ and 2 are satisfied with $o_{ij}^{t-1} > o_{ij}^{t}$, then, i^{th} vehicle overtakes j^{th} vehicle.

Rule 2: If conditions $1b$ and 2 are satisfied with $o_{ij}^{t-1} > o_{ij}^{t}$, then, j^{th} vehicle overtakes i^{th} vehicle.

The pseudo code of overtaking, OD is shown in Algorithm 3. As already mentioned that if i^{th}- vehicle overtakes j^{th} vehicle, then, during overtaking, one larger bounding box is formed over these two vehicles. Let it occurs at the t^{th} frame and after a certain time (say, t_1), these two vehicles will be separated, and fitted with two different bounding boxes. Therefore, overtaking cost between these two vehicles will fall suddenly. To confirm the overtaking, following the Rule 1 and 2, we also check the conditions if $b_i^{t-1} \approx b_i^{t+t_1}$ and $b_j^{t-1} \approx b_j^{t+t_1}$, where b_i^{t-1} indicates the area of bounding box fitted over the i^{th} vehicle at $(t-1)^{\text{th}}$ time.

4.2.4. Illegal parking

Parking vehicles in a restricted area (e.g., the zone where signs are posted in crosswalks or sidewalks or some area dictated by traffic laws) is called illegal parking. The flowchart for the detection of an illegally parked vehicle is shown in Fig. 5, and the algorithm, which is used for the detection of illegal parking, is named as illegal parking detection (IPD). From this figure, it is seen that the frames are extracted from the video stream. Then, object (vehicle) detection and tracking are done over these frames. At the same time, the parking zone is also detected over the video frames. Both training and testing are done for the detection of the parking zone. There are two classes images, such as an image with a parking zone, and an image with no parking zone are used for the training process. The training process has two steps: (i) extraction of unique features from the input images using HoGs (Karami et al., 2017), and ii) training of bi-class support vector machine (SVM) using these extracted features. For the testing process, one window of size (3 × 3) slides over the video frame for obtaining multiple regions. Each sliding window represents a region, where the HoG algorithm is applied over each region to extract unique features. The resultant feature vector corresponding to each region is fed to the trained SVM for parking zone classification. The regions classified as parking zones are merged to form the actual parking zone in the test video frame. At the same time, road is detected using hough line transformation and corrected component analysis. Co-ordinates that represent the outer area of road are stored. Thereafter, distances between center of parking region and all the said co-orniates are calculated. Out of these distances, minimum distance is considered as threshold Th_5. For illegal parking detection, first, we have checked whether moving vehicle is stopped or not. If vehicle is stopped, then the distance between vehicle and the center of parking area is measured. if this distance is greater than threshold Th_5, then it is termed as illegal parking. The pseudo-code of IPD is shown in Algorithm 4.

4.2.5. Wrong drop off location of passengers

In public transport, vehicles pick passengers from one stoppage and drop them at another stoppage. Occasionally, the passengers are dropped at wrong locations, which seriously violate the road safety rules. The flowchart for the detection of wrong drop off location of passengers is displayed in Fig. 6. The algorithm used for the detection of the wrong drop off location of passengers is named as WDLPD. From this figure, it is seen that the object detection and tracking are done over each frame extracted from a video stream. On the other hand, a bus stoppage is detected using SVM trained with HoG (Karami et al., 2017) feature. These features are extracted from the images containing the stoppage and no stoppage. Thereafter, the WDLPD algorithm classifies the object. If the object is classified as a vehicle, then, WDLPD algorithm checks whether the vehicle is stopped at the stoppage location or not. If the vehicle is stopped at the wrong place, then, this algorithm further checks if moving objects present at the same location or not. If moving objects exist, then, the algorithm tries to classify these objects. If these moving objects are classified as persons, then, it is considered as an anomaly (i. e., passengers drop at the wrong place). The pseudo-code for WDLPD is shown in Algorithm 5.

Algorithm 4 Pseudo-code of illegal parking.

Input: $\{X^t, Y^t\} = \{(x_1^t, y_1^t), (x_2^t, y_2^t), ..., (x_m^t, y_m^t)\}$: A set of centroid co-ordinates for m detetcted objects at $(t)^{th}$ frame; (x_r, y_r): center co-ordinate of parking region, and threshold $= Th_5$.

Output: $\{IPV\}^t$: A set to store the position of illegally parked vehicles at t^{th} frame.

```
Initialize the set {IPV}^t ← φ
for i=0 to m do
    compute v_i^t using Eq. (11)
    if v_i^t==0 then
        compute dis_i^t = sqrt((x_i^t − x_r)^2 + (y_i^t − y_r)^2)
        if dis_i^t > Th_5 then
            {IPV}^t ← (x_i^t, y_i^t)
        end
    end
end
return {IPV}^t
End
```

(a) spatial granule **(b)** temporal granule **(c)** spatio-temporal granule

Fig. 7. Results of spatio-temporal granulation.

Algorithm 5 Pseudo-code of wrong drop off location of passengers.

Input: $\{X^t, Y^t\} = \{(x_1^t, y_1^t), (x_2^t, y_2^t), ..., (x_m^t, y_m^t)\}$: A set of centroid co-ordinates for m detetcted objects at $(t)^{th}$ frame; (x_r, y_r): center co-ordinate of parking region/bus stoppage, $\{b^t\} = \{b_1^t, b_2^t, ..., b_m^t\}$: A set of bounding boxes over m detetcted objects at $(t)^{th}$ frame, and threshold $= Th_5$.

Output: $\{WDL\}^t$: A set to store the position of illegally parked vehicles at t^{th} frame.

Initialize the set $\{WDL\}^t \leftarrow \phi$

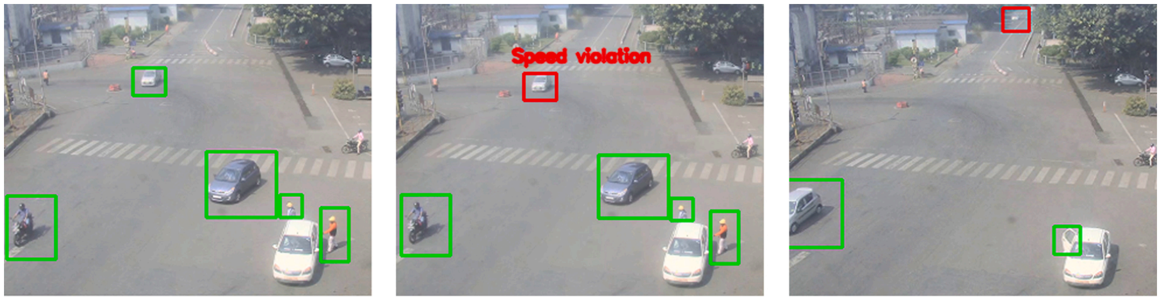
for *i=0 to m* **do**
  compute v_i^t using Eq. (11)
  if $v_i^t == 0$ **then**
    compute $dis_i^t = \sqrt{(x_i^t - x_r)^2 + (y_i^t - y_r)^2}$
    if $dis_i^t > Th_5$ **then**
      for *j=0 to m* **do**
        if $i \neq j$ **then**
          compute v_j^t using Eq. (11)
          if $v_j^t \neq 0$ **then**
            compute $dis_{i,j}^t = \sqrt{(x_i^t - x_j^t)^2 + (y_i^t - y_j^t)^2}$
            if $dis_{i,j}^t \in b_i^t$ **then**
              $\{WDL\}^t \leftarrow (x_i^t, y_i^t)$
            end
          end
        end
      end
    end
  end
end

return $\{WDL\}^t$

End

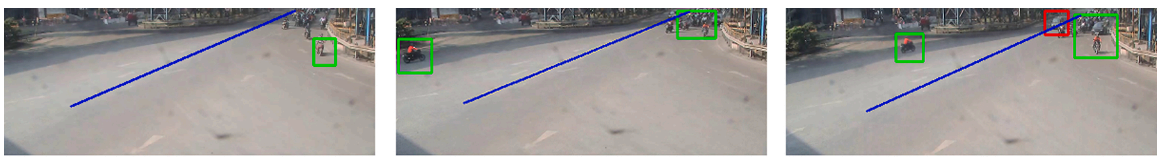
(a) Frame no. 91 **(b)** Frame no. 97

Fig. 8. Results of object detection and tracking.



(a) Frame no. 38 (b) Frame no. 73 (c) Frame no. 91

Fig. 9. Speed violation.



(a) Frame no. 18 (b) Frame no. 35 (c) Frame no. 36

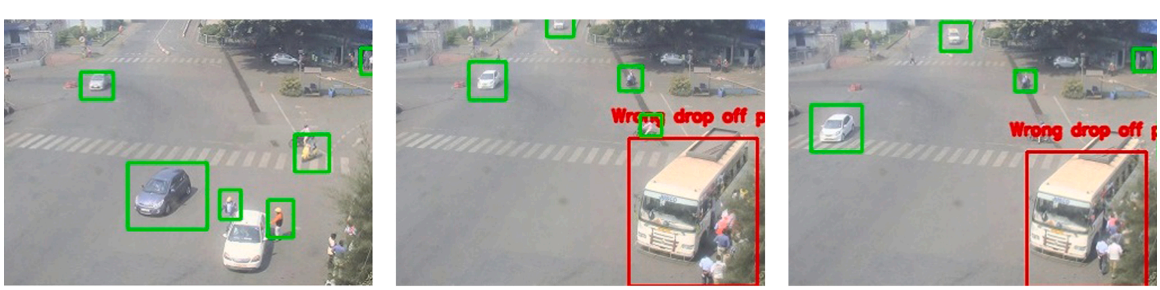
Fig. 10. One way traffic.

(a) Frame no. 84 (b) Frame no. 114 (c) Frame no. 259

Fig. 11. Overtaking.

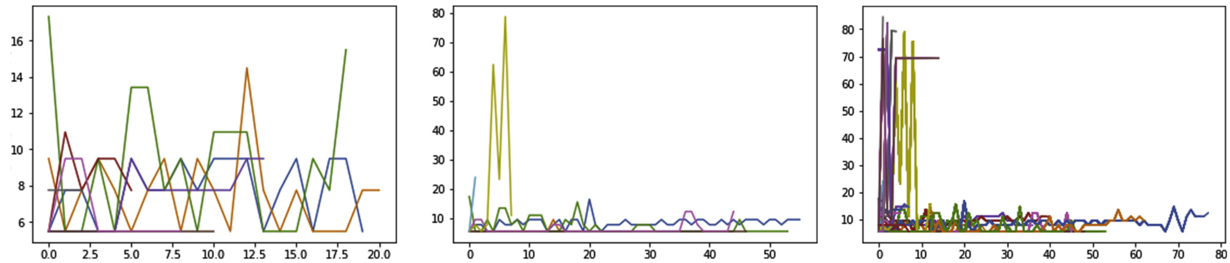
(a) Frame no. 29 (b) Frame no. 81 (c) Frame no. 225

Fig. 12. Illegal parking.



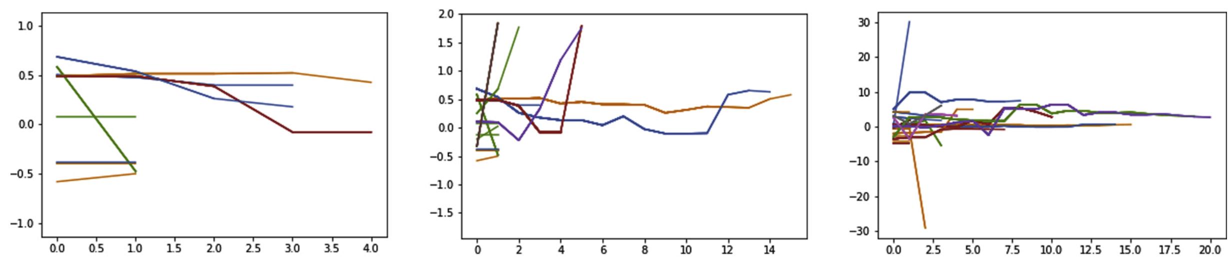
(a) Frame no. 53 (b) Frame no. 186 (c) Frame no. 192

Fig. 13. Wrong drop off location of passengers.



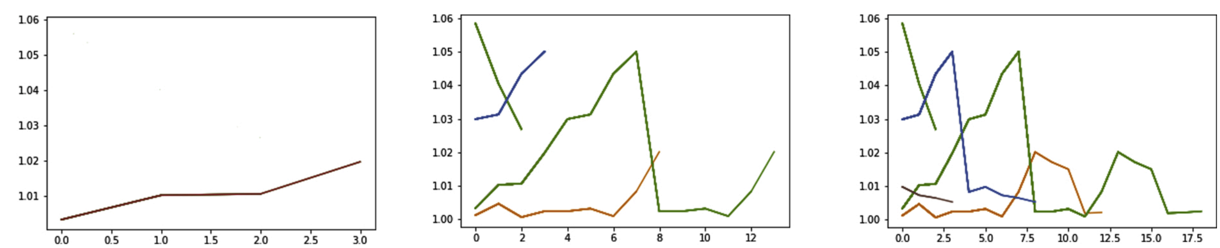
(a) Frame no. 38 (b) Frame no. 73 (c) Frame no. 91

Fig. 14. Velocity costs of different objects in different frames.



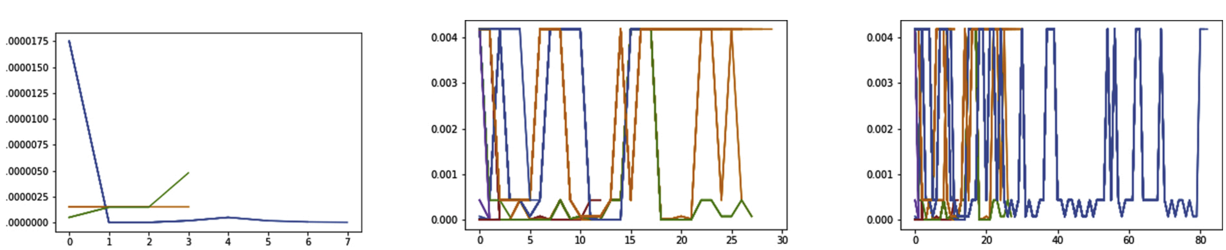
(a) Frame no. 18 (b) Frame no. 35 (c) Frame no. 36

Fig. 15. One-way costs of different objects in different frames.



(a) Frame no. 84 (b) Frame no. 114 (c) Frame no. 259

Fig. 16. Overtaking costs for different objects in different frames.



(a) Frame no. 29 (b) Frame no. 81 (c) Frame no. 225

Fig. 17. Illegal parking costs for different objects in different frames.

5. A case study

As a case study, we have used traffic video data acquired from an integrated plant in India. The effectiveness of our proposed methodology for traffic pre-event modeling is demonstrated with an extensive study over this video data. Based on the methods described in Section 4, first, we analyze the characteristics of different types of traffic pre-events occurred in the plant road. These characteristics are used to detect the abnormal behavior of traffic flow and are represented by some conditions and/or rules. A total of 30 s surveillance cameras of resolution 8 MP each are installed on the poles mounted at the side of each road in the plant to capture/monitor the traffic flow. Continuous video stream data is captured by these cameras. The video data related to traffic anomalies are considered as a rich source of information than traditional traffic data (in the form of categorical or continuous or text). Therefore, the video surveillance system acts as an integral part of advanced technology-driven safety management system (SMS) for the plant. The video data is continuously monitored by road safety professionals in the surveillance room to check whether any anomaly takes place in any roadside or not. In addition, some safety professionals are present at the roadside area to monitor the daily traffic flow. Incidents sometimes remain unnoticed by the professionals who are present in the surveillance room. Under such circumstances, they look for assistance look for from the persons present in the roadside to get information about the actual incidents. Based on the type of incidents, penalties are charged to the person responsible for the incident. Moreover, the video summary of the incident is used for future inspection and training of the driver. This enhances road safety. As the aforesaid process is extremely human dependent, a traffic pre-event may be skipped by both safety professionals of the control room and roadside area, which may adversely impact the level of road safety. Therefore, it is essentially required to develop a video surveillance system embedded with different algorithms that can automatically detect traffic pre-events from the streaming videos. Whenever any incident happens on roads, it will be automatically logged into the surveillance database, thereby significantly reducing human efforts.

In this study, a total of five types of traffic pre-events (i.e., anomalies), namely speed violation, one-way traffic, overtaking, illegal parking, and wrong drop off location of passengers are considered. Definitions of these anomalies are stated as follows: (i) *Speed violation:* It refers to over speeding (ii) *Overtaking:* One vehicle overtakes another one (iii)*One-way traffic:* Car coming from the opposite direction is termed as an anomaly. (iv)*Illegal parking:* Vehicles are parked in an illegal or restricted area (v)*Wrong drop off location of passengers:* passengers are dropped at wrong locations. The effectiveness of the developed algorithms for the detection of aforesaid traffic pre-events is stated in the next section.

6. Results and discussions

In this section, the experimental setup, the dataset used, and experimental results are discussed in details.

6.1. Experimental setup

The algorithms (refer to Section 4) used in this study are coded in Python 3.6 (Anaconda) in Windows 10 using an Intel *i*5 processor clocked at 2.50 GHz with 8 GB of memory. The libraries used are ‘cv2’, ‘NumPy’, and ‘imutils’. For object detection, a threshold $T_1 = 1$ and $N = 3$ are considered (Chakraborty et al., 2013).

For the detection of speed violation, threshold Th_1 is used. Here, only one parameter needs to be selected. Grid search is the most effective in such cases. Of various approaches, Brute force (Schaeffer et al., 1993) is the most effective and well-known grid search approach. Therefore, it has been used in this study to select the optimal value of Th_1. In this strategy, first, one traffic video (consists of 231 frames) is collected. This video contains both normal and abnormal (i.e., high speeding vehicle) scenarios. Second, the velocity cost of each vehicle present in this video is calculated. Graphs of velocity costs are shown in Fig. 14. Graphs in Fig. 14 are represented by different colored line charts. These colored lines depend on the number of objects (i.e., vehicles) detected in the video. Each color in these graphs represents the velocity cost of each detected vehicle. Figs. 14(a) and 14 (b) represent the velocity costs of all detected vehicles from Frame no. 1 to Frame no. 38 and Frame no. 1 to Frame no. 73. From Fig. 14(a), it is seen that the range of velocity costs for all detected vehicles is 8 to 25. But from Fig. 14(b), it is evident that the one vehicle poses high velocity cost (greater or equal to 60) and the velocity costs of other vehicles are in between 8 to 25. Therefore, the vehicle haing velocity cost of 60 is act as an outlier here, and this value is considered as threshold Th_1.

For the detection of one-way traffic, threshold Th_4 is used. Here, also Brute force (Schaeffer et al., 1993) strategy is used to select the optimal value of Th_4. In this strategy, first, one traffic video (consists of 87 frames) is collected. This video contains both normal and abnormal (i.e., one-way traffic) scenarios. Second, the direction costs between vehicles present in this video are calculated. Graphs of direction costs are shown in Fig. 15. Graphs in Fig. 15 are represented by different colored line charts. The colored lines depend on the number of vehicles detected in the video. Each color in these graphs represents the direction cost between two detected vehicles. Figs. 15(b) and 15 (c) represent the direction costs between all detected vehicles from Frame no. 1 to Frame no. 35 and Frame no. 1 to Frame no. 36. From, Fig. 15(b), it is seen that the range of direction costs between all detected vehicles is -0.5 to 2. But from Fig. 15(c), it is evident that the that the range of direction costs between all detected vehicles is -30 to 30. Here, the direction cost between two vehicles is abruptly changed. This proves that abnormality must happen in this frame. For defining threshold Th_4, first we have calculated average of these two ranges. Then, the difference between these two averages is considered as threshold Th_4. $\mathrm{Th}_4 = (60/2) - (2.5/2) = 28.75$. For two vehicles, if the modulus value of their direction cost is greater than Th4, then this is considered as abnormal event (i.e., one-way traffic).

Datasets used in this study are briefly described in the next section.

6.2. Datasets used

The effectiveness of the proposed methodology is demonstrated over two benchmark datasets, namely CamSeq01 dataset (Cambridge, 2007), and Illegally Parked Vehicle Dataset (ISLab-PVD) (Jo et al., 2017), and traffic videos acquired from a plant in India. Among these datasets, CamSeq01 contains a sequence of 50 images for overtaking detection. Whereas, dataset ISLab-PVD (Jo et al., 2017) is considered as a benchmark for illegally parked vehicle detection. This dataset contains 16 video sequences with various challenging scenarios, including crowded scenes, different lighting conditions, vehicles of varying sizes, and night-time conditions. There is a scarcity in the availability of benchmark datasets for rash driving detection. Rash driving is categorized into speed violation, one-way traffic, and overtaking. Traffic videos from a plant contain all types of aforesaid rash driving scenarios, illegal parking, and wrong drop off location of passengers. We have acquired and used a total of 132 plant traffic videos, each having a duration of 2 to 8 minutes. Three challenges, such as occlusion, low resolution, and complex traffic flows are there in these videos. We consider these traffic pre-events because these activities may cause a road crash. There is no benchmark dataset available for the “wrong drop off locations of passengers”. Due to the lack of benchmark datasets, developed algorithms are tested over the traffic surveillance videos acquired from the plant. The experimental results are elaborated in the following section.

6.3. Experimental results

Experiments are conducted to evaluate the effectiveness of the

Table 1
Comparative results of speed detection and violation.

Algorithm	Accuracy (%)	Speed (fps)
CVS	71.3	14
VSD	77.6	15
DSE	83.9	18
SVD	88.1	21

Table 2
Comparative results of one-way traffic.

Algorithm	Accuracy (%)	Speed (fps)
TOTV	76.2	9
WADD	81.6	15
OTD	88.1	19

Table 3
Comparative results of overtaking.

Algorithm	Accuracy (%)	Speed (fps)
OVD	82.6	12
VDDA	87.1	15
OD	92.1	19

developed algorithms using plant traffic videos and two benchmark image datasets, CamSeq01 and ISLab-PVD. After the collection of videos, they are pre-processed to remove noises. After noise removal, processed video frames are used for object detection and tracking, and traffic pre-events detection. Object detection and tracking is the pre-requisite of traffic pre-events detection. Result for object detection and tracking over video frames is presented in Section 6.3.1. The results of traffic pre-events detection over plant traffic videos are discussed in Section 6.3.2, and finally, a detailed comparative study between the developed algorithms and state-of-the-art algorithms for speed violation, one-way traffic, overtaking, and illegal parking detection is stated in Section 6.3.3. However, the effectiveness of our developed algorithms for the detection of aforesaid anomalies are proved by adopting the statistical significance test and robustness checking, as stated in Sections 6.3.4 and 6.3.5, respectively. Moreover, a comparative study using two benchmark datasets for detection of overtaking and illegal parking is presented in Section 6.3.6.

6.3.1. Results for object detection and tracking

Video frames acquired from the surveillance camera are originally color images. To preserve the intensity of each pixel of these images, the grey conversion is done. Grey image is not affected by luminance. To reduce noise from grey images, the binary conversion is done using smoothing operation. First, a (3×3) spatial window is moved throughout each pixel location in the original image. Then, if the intensity of any pixel location is greater than threshold T_b, this location is filled with pixel intensity of 255, otherwise, the location is filled with pixel intensity of 0. T_b is the average of pixel intensities present in (3×3) spatial window. The binary images produced are blurred. Due to noise, blurred images may contain holes. To reduce the size of holes in the images, binary images are converted to dilated images (Pramanik et al., 2018), which are noise-free/ less noisy processed images.

As discussed in Section 4.1.1, spatio-temporal granulation is performed over the dilated image for obtaining the regions of the object. Results of this granulation over a video frame are depicted in Fig. 7. Here, Figs. 7(a), 7 (b), and 7 (c) represent the spatio-color neighborhood, temporal, and spatio-temporal granules over a video frame, respectively. After the formation of spatio-temporal granules, bounding boxes are fitted over these detected granules for object detection. Object detection and tracking results over two different frames of the same video are shown in Figs. 8(a) and 8 (b). After detection and tracking,

Table 4
Comparative results of illegal parking.

Algorithm	Accuracy (%)	Speed (fps)
IPSF	82.9	14
PDT	87.4	16
IPD	89.3	24

Table 5
Results of the statistical tests.

Algorithm	P-value	Conclusion
SVD vs. DSE	2.0914e-23 (P-value $\leq$ 0.05)	SVD is significant over DSE
SVD vs. VSD	2.0909e-23 (P-value $\leq$ 0.05)	SVD is significant over VSD
SVD vs. CVS	2.0909e-23 (P-value $\leq$ 0.05)	SVD is significant over CVS
OTD vs. TOTV	8.2719e-14 (P-value $\leq$ 0.05)	OTD is significant over TOTV
OTD vs. WADD	7.3184e-14 (P-value $\leq$ 0.05)	OTD is significant over WADD
OD vs. OVD	3.6102e-08 (P-value $\leq$ 0.05)	OD is significant over OVD
OD vs. VDDA	3.2481e-08 (P-value $\leq$ 0.05)	OD is significant over VDDA
IPD vs. IPSF	5.0161e-11 (P-value $\leq$ 0.05)	IPD is significant over DSE
IPD vs. PDT	5.2518e-09 (P-value $\leq$ 0.05)	IPD is significant over VSD

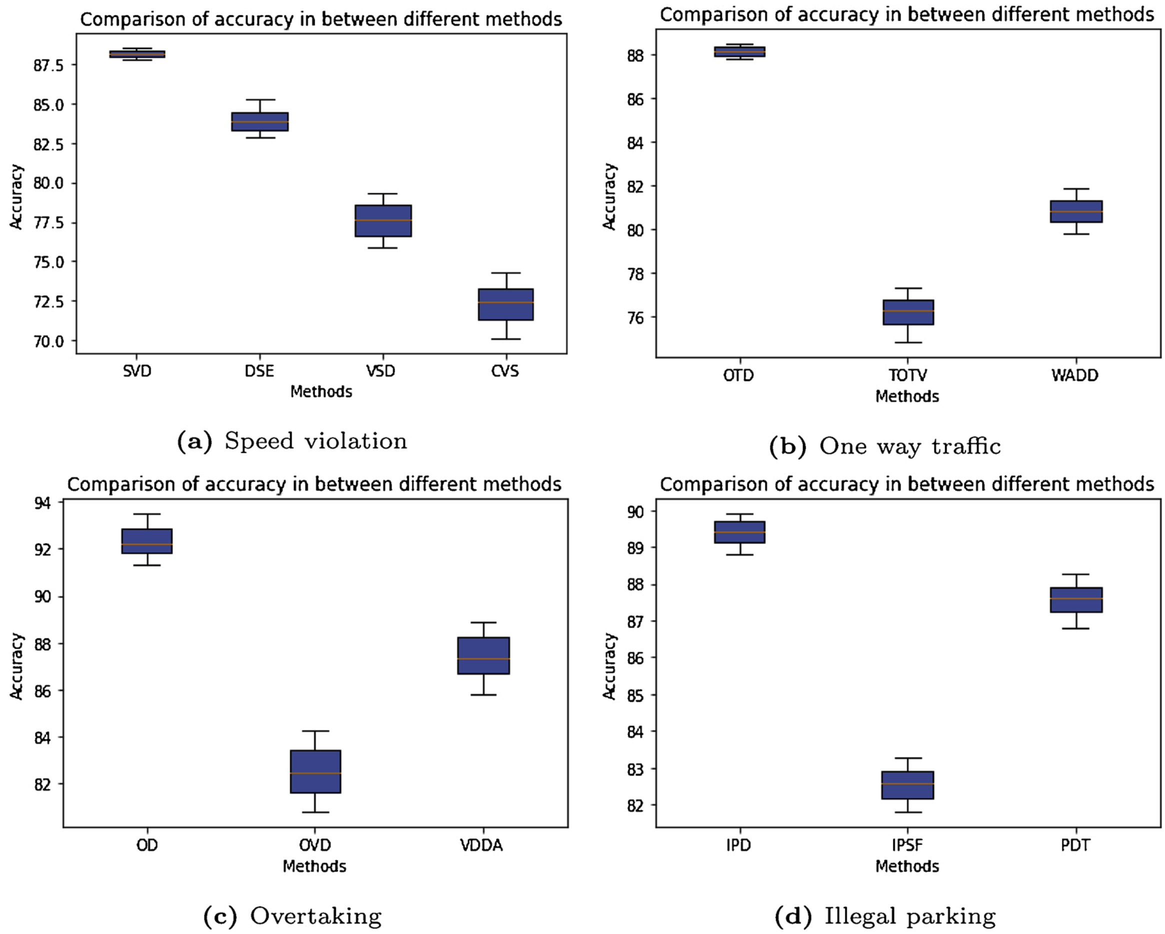
object-level features are extracted and analyzed for traffic pre-events detection. The results of different traffic pre-events detection over plant videos are explained in the following section.

6.3.2. Results of traffic pre-events detection

A detailed experiment is carried out over 132 traffic videos acquired from a plant. Of them, a total of 30 contain normal scenarios, and 102 have abnormal or anomalous traffic events, namely speed violation, one-way traffic, overtaking, illegal parking, and wrong drop off location of passengers. For instance, the detection results of these type of anomalies over different video frames are depicted in Figs. 9–13. Each figure contains three image frames, where the first frame shows the pre-anomaly scenario and second and third frames (except Fig. 10) represent the anomalous scenario. As discussed in Section 4, various cost functions are developed for each detected object (i.e., vehicle) in the video frame to model/detect these five types of traffic pre-events (anomalies). Graphs of the cost functions for different detected vehicles in different frames corresponding to the same videos are shown in Figs. 14–17.

(i) **Speed violation:** Fig. 9 shows the detection results of speed violation over different frames of one video. Velocity cost is defined to model the speed violation. Graphs of velocity costs for different detected objects (i.e., vehicles) over different frames of the same video are shown in Fig. 14. The graphs in Figs. 14(a), 14 (b), and 14 (c) correspond to the detection results of Figs. 9(a), 9 (b), and 9 (c), respectively. In Figs. 14(a), 14 (b), and 14 (c), the x-axis represents the frame number over which velocity cost of at least one detected object is non-zero, and the y-axis represents the velocity costs of all detected objects. Figs. 9(a) and 14 (a) are corresponding to the frame no. 38 of a video. Similarly, Figs. 9(b) and 14 (b), and Figs. 9(c) and 14 (c) correspond to frame nos. 73 and 91, respectively from the same video. Graphs in Fig. 14 are represented by different colored line charts. The colored lines depend on the number of moving objects detected in the video. Each color in these graphs represents the velocity cost of each detected moving object. From Fig. 14, it is observed that the trend of velocity costs for more than seven moving objects. It implies that a total of seven moving objects, at least, are detected either in this frame or in its previous frames.

Likewise, velocity costs are also obtained for all these detected objects. The graph in Frame no. 38 (refer to Fig. 14) represents all velocity costs for all detected moving objects from Frame no. 1 to



(a) Speed violation

(b) One way traffic

(c) Overtaking

(d) Illegal parking

Fig. 18. Robustness checking using box-plots.

Frame no. 38. Similarly, Figs. 14(b) and 14 (c) represent the velocity costs for all detected moving objects from Frame no. 1 to 73, and Frame no. 1 to 91, respectively. All the graphs of Fig. 14 illustrate how the velocity costs of detected moving objects vary from the frame by frame. From Fig. 14(a), it is evident that the velocity costs of all detected moving objects are less than the threshold Th_1, thereby resulting as a normal scenario. If speed violation occurs in any frame, the area is marked by red color bounding-box; otherwise, moving objects are marked by green color bounding-box. Based on this graph of Fig. 14(a), the detection result is shown in Fig. 9(a), where we can see that all detected moving objects are marked by green color bounding-boxes, which represent the normal scenario. From Fig. 14(b), threshold Th_1 is measured automatically. Based on this threshold, one speed violation is occurred in the Frame no. 73. The result of a speed violation is shown in Fig. 9(b), where it is observed that one detected moving object is marked by red colored bounding-box. This indicates a speed violation. In addition, from Fig. 14 (c), it is seen that velocity costs (represented by different colors) of three detected moving objects exceed the threshold Th_1. These means from Frame no. 1 to 91, at least, a total of three speed violations occur. The speed violation detection result is shown in Fig. 9(c), where it is seen that one detected moving object is marked by red colored bounding-box, which implies a one-speed violation occurs in the Frame no. 91, and other two-speed violations occur in between Frame no. 73 to 91.

(ii) **One way traffic:** Both Figs. 10 and 15 exhibit the detection results and graphs of one way traffic. Figs. 10(a), 10 (b), and 10 (c) show the detection results in Frame nos. 18, 35, and 36, respectively. In the case of one-way traffic, the lane is marked manually by blue color over the road, as shown in Fig. 10. On the other hand, Figs. 15(a), 15 (b), and 15 (c) represent the trend of direction costs between all detected moving objects from Frame nos. 1 to 18, 1 to 35, and 1 to 36, respectively. In Figs. 15(a), 15 (b), and 15 (c), the x-axis represents the frame numbers over which one-way cost between two detected moving objects is non-zero, and the y-axis represents the direction costs for all detected moving objects. Each color in these graphs represents the direction cost between two detected moving objects. From Figs. 15(a) and 15 (b), it is evident that direction costs between all detected moving objects are less than threshold Th_4. This means no anomaly (one-way traffic) occurs from Frame nos. 1 to 18 and 1 to 35. One-way traffic detection results are shown in Figs. 10(a) and 10 (b), where all detected moving objects are marked by green colored bounding-boxes. This results in normal scenarios over the two frames, Frame nos. 18 and 35. From Fig. 15(c), it is evident that the one-way cost between two detected moving objects exceeds the threshold Th_4, which results in one-way traffic

Table 6
Comparative study over benchmark datasets.

Algorithm	Anomaly	Dataset	Accuracy (%)	Speed (fps)
OVD	Overtaking	CamSeq01	76.2	9
VDDA	Overtaking	CamSeq01	81.6	15
OD	Overtaking	CamSeq01	88.1	19
IPSF	Illegal parking	ISLab-PVD	82.9	14
PDT	Illegal parking	ISLab-PVD	87.4	16
IPD	Illegal parking	ISLab-PVD	89.3	24

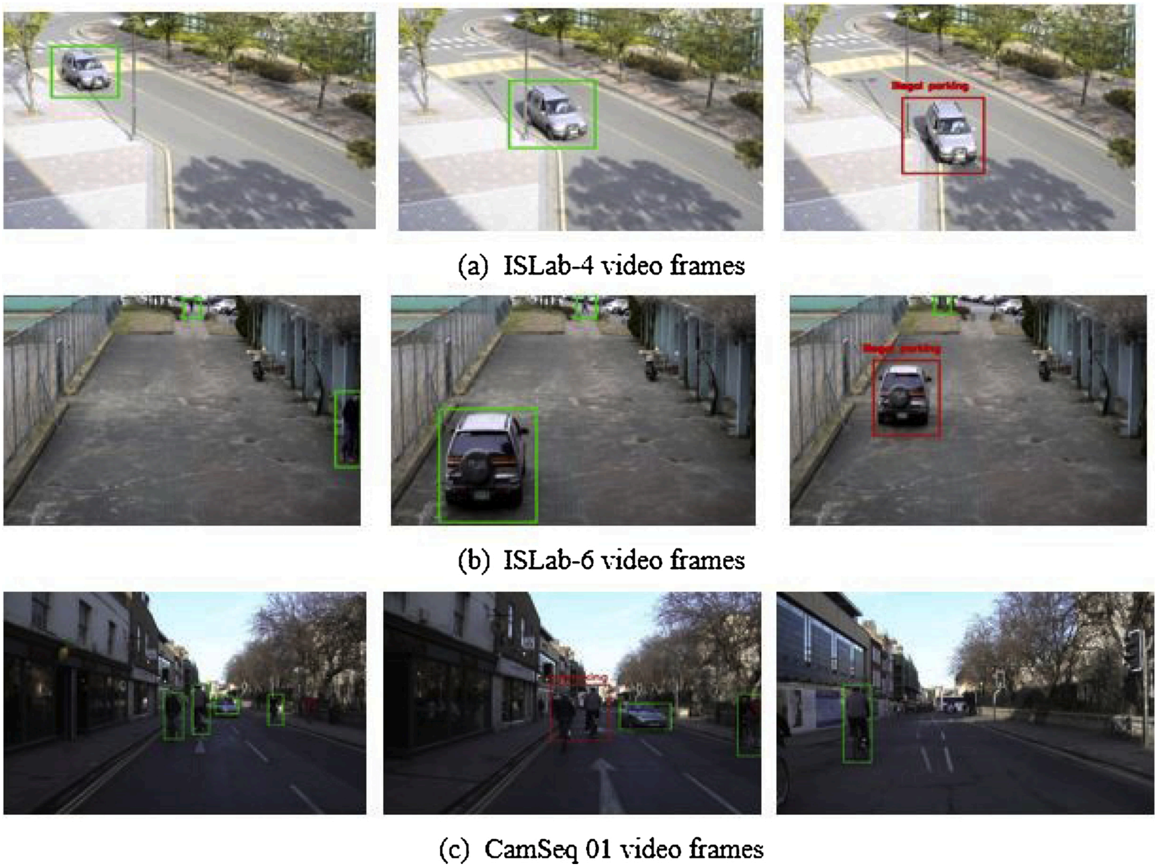


Fig. 19. Results of illegal parking and overtaking detection over benchmark datasets ISLab-PVD and CamSeq01.

detected in Frame no. 36. The detection result of one-way traffic is shown in Fig. 10(c), where the detected moving object violates one-way traffic rule is marked by red colored bounding-box.

(iii) **Overtaking:** Figs. 11 and 16 show the detection results and graphs of overtaking, respectively. Figs. 11(a), 11 (b), and 11 (c) exhibit the overtaking detection results in Frame nos. 84, 114, and 259, respectively. On the other hand, Figs. 16(a), 16 (b), and 16 (c) represent the trend of overtaking costs for all detected moving objects from Frame nos. 1 to 84, 1 to 114, and 1 to 259, respectively. Graphs in Fig. 16 are represented by different colored line charts. The number of colored lines depends on the number of moving objects detected in the video. Each color in these graphs represents the overtaking cost of each detected moving object. From Fig. 16(a), it is evident that the trend of all overtaking costs between all detected moving objects in the Frame nos. 1 to 84 is increasing. That means no overtaing is happened. Overtaking detection result on Frame no. 84 is shown in Fig. 11(a), which represents the normal scenario. From Figs. 16 (b) and 16 (c), it is evident that the trend of at least one overtaking cost between two detected moving objects in the Frame nos. 1 to 114 and 1 to 259 is in zigzag nature. Therefore, overtaking is happened. Overtaking detection results are shown in Figs. 11(b) and 11 (c). From these two figures, it is seen that overtaking occurs in Frame nos. 114 and 259. During overtaking, one detected moving object overtakes other detected moving object and hence, both are marked by red-colored bounding-boxes, as seen in Figs. 11(b) and 11 (c).

(iv) **Illegal parking:** Both Figs. 12 and 17 exhibit the detection results and graphs of illegal parking. Figs. 12(a), 12 (b), and 12 (c) reveal the illegal parking detection results in Frame nos. 29, 81, and 225, respectively. On the other hand, Figs. 17(a), 17 (b), and 17 (c) represent the trend of parking costs for all detected objects from Frame nos.1 to 29, 1 to 81, and 1 to 225, respectively. Graphs in Fig. 17 are represented by different colored line charts. The number of colored lines depends on the number of objects detected in the video. Each color in these graphs represents the parking cost of each detected object. From Fig. 17(a), it is evident that no parking cost exceeds the threshold Th_5 from Frame no. 1 to 29. Illegal parking detection result at Frame no. 29 is shown in Fig. 12(a), which represents the normal scenario. From Figs. 17 (b) and 17 (c), it is evident that parking costs exceed threshold Th_5. Illegal parking detection results are shown in Figs. 12(b), 12 (c). From these two figures, it is seen that illegal parking occurs in Frame nos. 81 and 225. Illegally parked vehicles are marked by red-colored bounding-boxes, as shown in Figs. 12(b), 12 (c).

(v) **Wrong drop off location of passengers:** The characteristics of illegal parking detection and wrong drop off location of passengers are the same, therefore, we have considered the same threshold (Th_5) to detect these two pre-events. After detecting the illegally parked vehicles, our WDLPD algorithm checks whether trajectories of detected moving objects (i.e., persons) are present at the parking place or not. If present, then it is termed as the wrong drop off location of passengers. Detection results of such anomaly are shown in Figs. 13(a), 13 (b), and 13 (c) for Frame nos. 53, 186, and 192, respectively. Here, the anomaly is marked by the red-colored bounding-box, whereas detected moving objects in the normal scenario are marked by the green-colored bounding-boxes.

6.3.3. *Comparative study*

To prove the effectiveness of our algorithms, we have conducted a comparative study between developed algorithms and some state-of-the-art algorithms for the detection of four types of traffic pre-events (i.e., speed violation, one-way traffic, overtaking, and illegal parking) in 132 traffic videos acquired from the plant. This is the first time, when anomaly, named the "wrong drop off location of passengers" is analyzed. Therefore, no comparative study is done for the detection of this anomaly. Accuracy and speed are considered as performance metrics for this study. In case of speed violation detection from traffic surveillance videos, there are few studies available, and all are useful for detecting or estimating the speed. Before going for speed violation detection, we have also detected the speed of vehicles. Hence, we have conducted a comparison between our proposed algorithm and other existing algorithms for speed detection based on two performance metrics, speed (frame per second (fps)) and accuracy (%). For speed detection and violation, three recent state-of-the-art methods, a combination of saturation and value (CVS) (Rad et al., 2010), vehicle speed detection (VSD) (Wu et al., 2009), and detection and speed estimation (DSE) (Kumar and Kushwaha, 2016) are used in this analysis. The results are reported in Table 1. From this table (last row), it is seen that SVD is superior to others in terms of both speed and accuracy. Similarly, for doing a comparative study on one-way traffic detection, we have considered two state-of-the-art algorithms, namely tracking oncoming and tracking vehicles (TOTV) (Barth and Franke, 2010) and wrong-way driver detection (WWDD) (Monteiro et al., 2007). The results of the comparisons are shown in Table 2. From this table, it is seen that our algorithm (OTD) is superior to TOTV and WWDD algorithms in terms of both speed and accuracy.

In the case of overtaking detection, we have considered two algorithms, namely overtaking vehicle detection (OVD) (Chanawangsa and Chen, 2013) and vehicle detection for driving assistance (VDDA) (Satzoda and Trivedi, 2014) in comparison. Since these algorithms are the most useful in overtaking detection, they are considered in our analyses. The comparative results from analyses are shown in Table 3. From this table, it is observed that our algorithm (OD) is superior to the state-of-the-art methods in terms of both speed (fps) and accuracy (%). From this table, it is also evident that the conditions and rules that we have defined for overtaking detection are effective as our method gives 92.1% accuracy for overtaking detection. Finally, for doing the comparative study on illegal parking, we have considered two most useful algorithms in this experiment, namely illegally parked vehicle by seed fill algorithm (IPSF) (Sarker et al., 2015) and parking detection transformation (PDT) (Lee et al., 2009). Comparative results are reported in Table 4. From this table, it is seen that our algorithm is superior to both IPSF and PDT with respect to both speed (fps) and accuracy (%).

As already said that the traffic pre-event/anomaly 'wrong drop off location of passengers' is new in the domain of road safety. Therefore, no comparion can be carried out for this anomaly. Our proposed WDLPD algorithm achieves 84.1% accuracy and speed of 24 fps in detecting the anomaly 'wrong drop off location of passengers' in road. From the aforesaid study, it is evident that all proposed algorithms are good in detecting the traffic pre-events. As it is said, object detection and tracking is the pre-requisite of the traffic pre-event detection. From this study, it can also be concluded that the approaches that are adopted for object detection and tracking are effective. Although aforesaid algorithms are used for the detection of traffic pre-events in plant road, they can be useful in any kind of road. Statistical significance test for the developed algorithms is defined in the next section.

6.3.4. Statistical significance test

Based on the above-mentioned results reported in Tables 1 to 4, developed algorithms, SVD, OTD, OD, and IPD are found to be the best for the detection of speed violation, one-way traffic, overtaking, and illegally parking. Besides, we have conducted a Wilcoxon signed-rank test (Sarkar et al., 2019, 2020) to check whether these developed algorithms for the detection of traffic pre-events are statistically significant than other state-of-the-art algorithms. Wilcoxon signed-rank test is conducted using accuracy. The test results are reported in Table 5, which reveals that there exist significant differences among the performances of developed algorithms and some state-of-the-art algorithms since the corresponding p-values are less than 0.05. From the first three rows of Table 5, it is evident that SVD is statistically significant than DSE, VSD, and CVS for the detection of speed violation. Likewise, OTD is significant than TOTV and WADD for the detection of one way traffic (refer to row nos. 4 and 5 of Table 5). Another algorithm OD is statistically significant than OVD and VDDA for the detection of overtaking (refer to rows 6 and 7 of Table 5). Finally, from the last two rows of Table 5, it is evident that IPD is statistically significant than IPSF and PDT. Besides with statistical significance test, the robustness of the developed algorithms is also investigated, which is discussed in the next section.

6.3.5. Robustness checking

A total of 132 plant traffic videos are used to check the robustness of each developed algorithm for each traffic pre-event detection. As a consequence, a total of 132 accuracy values are obtained for each algorithm. Using these accuracies, a box-plot is generated and exhibited in Figs. 18(a), 18 (b), 18 (c), and 18 (d). Fig. 18(a) shows the box plots of accuracies obtained from the experiments on speed violation. Similarly, Figs. 18(b), 18 (c), and 18 (d) show the box plots of accuracies obtained in the algorithms used for the detection of one-way traffic, overtaking, and illegal parking, respectively. It is revealed from Fig. 18(a) that the CVS algorithm for speed violation detection shows the minimum accuracy with a moderate range of dispersion. On the other hand, the maximum accuracy with the lowest degree of dispersion is yielded by the SVD algorithm. Similarly, from Figs. 18(b), 18 (c), and 18 (d), it is evident that TOTV, OVD, and IPSF algorithms for one-way traffic, overtaking, and illegal parking show the minimum accuracy with the highest degree of dispersion. Whereas, OTD, OD, and IPD algorithms for the same traffic pre-events show the maximum accuracy with the lowest degree of dispersion. Hence, SVD, OTD, OD, and IPD are considered to be the robust algorithms and also, the best ones among the other respective state-of-the-art algorithms used for the detection of speed violation, one-way traffic, overtaking, and illegal parking, respectively. A comparative study using two benchmark datasets for overtaking and illegal parking detection is presented in the next section.

6.3.6. A comparative study using benchmark datasets

As there are few studies available on rash driving (speed violation, one-way traffic, and overtaking), the number of available benchmark datasets is also limited. For overtaking detection, there is a benchmark dataset, namely CamSeq01 (Cambridge, 2007). Besides, for detecting illegal parking, there is an open-source benchmark dataset (ISLab-PVD) (Jo et al., 2017) available. In most cases, authors use their dataset(s) for their analyses and do not provide free access to download their datasets. Therefore, our analyses are restricted to only two above-mentioned datasets. Hence, the comparative study is carried out for the detection of these two traffic pre-events (i.e., overtaking and illegal parking) to make a fair comparison.

A comparative study between OD, OVD, and VDDA is carried out for overtaking using the dataset CamSeq01. Similarly, another comparative study is done between IPD, IPSF, and PDT for illegal parking detection using the dataset ISLab-PVD. The results are presented in Table 6. From this table (refer to first three rows), it is observed that OD is superior to some state-of-the-art algorithms for overtaking detection over the dataset CamSeq01 in terms of both speed (fps) and accuracy (%). From the same table (refer to last three rows), it is also reported that IPD is the best as compared to some existing algorithms for illegal parking detection over the dataset ISLab-PVD in terms of both speed and accuracy. The experimental results for the detection of overtaking and illegal parking over these two datasets using our methods are shown in Fig. 19.

As stated earlier (refer to Section 1), road crash is happened due to the occurrence of various traffic pre-events, such as speed violation, one-way traffic, overtaking, illegal parking, and wrong drop off location of passengers. After detecting these traffic pre-events, an alarm will be automatically generated in the control room for taking corrective actions, which may prevent the road crash to occur. If corrective actions cannot be taken on-time, road crash or accident may occur. Even, under the state of such situation, if we can automatically detect the road crash and subsequently generate an alarm to the control room to send immediate help to the accident place, we can then mitigate the consequences of the event. This will eventually improve the level of road safety.

7. Conclusion

Events on roads are growing globally each year as roads are used as vital mode of transportation. Therefore, maintaining road safety is of paramount importance nowadays. Any road event usually happens due to the existence of various traffic pre-events, for examples, speed violation, one-way traffic, overtaking, illegal parking, and wrong drop off locations of passengers. Therefore, correct detection of these traffic pre-events and adaptation of timely corrective actions are essential for enhancing road safety. The most important pre-requisite of road safety is the correct detection of traffic pre-events/anomalies. As per the authors' knowledge, this is the first time when the detection of traffic anomaly called "the wrong drop off location of passengers" is considered as a research topic. Moreover, in previous studies, only one or two traffic anomalies are detected using a single algorithm and this is restricted to context-specific and simple traffic flow. To address these issues, we have developed a video surveillance-based system to enhance the level of road safety in an automated way. We have also developed a conceptual framework for functioning this video surveillance system in real-time. This framework involves five steps: (i) analysis of the traffic pre-events, (ii) identification of places for installing the CCTV, (iii) development of algorithms for modeling the traffic pre-events, (iv) training and validation of the developed algorithms, and (v) operation. The operation consists of two basic stages: (i) test over live stream video, and (ii) take corrective actions after detecting the traffic anomaly. Overall, the framework is generic in nature. Various algorithms are developed for detecting the traffic pre-events, namely speed violation, one-way traffic, overtaking illegal parking, and wrong drop off location of passengers. A set of features and rules are generated for the detection of aforesaid traffic pre-events. Our developed algorithms are named as SVD for "speed violation detection", OTD for "one-way traffic detection", OD for "overtaking detection", IPD for "illegal parking detection", and WDLPD for the "wrong drop off location of passengers detection". Object detection and tracking is the pre-requisite of the traffic pre-event detection. We have adopted the approach presented in Chakraborty et al. (2013) for object detection.

To demonstrate the effectiveness of our proposed algorithms, a comparative study with some different state-of-the-art methodologies is conducted using two benchmark datasets and 132 traffic videos acquired from a plant in India. These two benchmark datasets are 'CamSeq01' and 'ISLab-PVD'. Two performance metrics, speed (fps) and accuracy (%) are used for the evaluation of algorithms in our comparisons. From the experiments conducted in this study, it is evident that SVD is superior to the combination of saturation and value (CVS) (Rad et al., 2010), vehicle speed detection (VSD) (Wu et al., 2009), and detection and speed estimation (DSE) (Kumar and Kushwaha, 2016) algorithms for the detection of speed violation. Similarly, OTD outperforms than tracking oncoming and tracking vehicles (TOTV) (Barth and Franke, 2010) and wrong-way driver detection (WWDD) (Monteiro et al., 2007) algorithms for one-way traffic detection. On the other hand, OD is better than overtaking vehicle detection (OVD) (Chanawangsa and Chen, 2013) and vehicle detection for driving assistance (VDDA) (Satzoda and Trivedi, 2014) algorithms for overtaking detection. Finally, illegal parking detection (IPD) algorithm is found to be better than the illegally parked vehicle by seed fill (IPSF) (Sarker et al., 2015) and parking detection transformation (PDT) (Lee et al., 2009) algorithms for the detection of illegal parking. In addition, detection of "wrong drop off locations of passengers" is a new application, as per the authors' knowledge; therefore, no such comparative study is available in this research. We have also demonstrated that our developed algorithms are statistically significant and robust than the aforesaid state-of-the-art methods. The developed framework is generic in nature and can be used in any road context to detect traffic pre-events, namely speed violation, one-way traffic, overtaking, illegal parking, and wrong drop off locations of passengers.

Similar to other studies, the present one also suffers from a few limitations. Although a lot of traffic pre-events are responsible for road events, our attention of this research focuses only on the above-mentioned five types of traffic pre-events. Additionally, the present study is restricted to monocular vision only. Moreover, clutter effect is not handled in this research. As a future scope, a framework can be developed which can address both monocular and stereo vision in traffic pre-event detection.

CRediT author statement

Anima Pramanik: Conceptualization, Methodology, Model building, Implementation/Experimentations, Model validation, Writing – Reviewing and Editing.

Sobhan Sarkar: Conceptualization and analysis, Writing – Organization of paper and reviewing & editing.

J Maiti: Conceptualization, supervision, reviewing, and editing.

Acknowledgments

The work is done under UAY project, Government of India (GOI), funded by Ministry of Education & Ministry of Steel (GOI), and Tata Steel Limited. We acknowledge the Centre of Excellence in Safety Engineering and Analytics (CoE-SEA) (www.iitkgp.ac.in/department/SE), IIT Kharagpur and Safety Analytics & Virtual Reality (SAVR) Laboratory (www.savr.iitkgp.ac.in) of Department of Industrial & Systems Engineering, IIT Kharagpur for experimental/computational and research facilities for this work. We would like to thank the management of the plant for providing relevant data and their support and cooperation during the study.

Conflict of interest: The authors declare no conflict of interest.

References

Alonso, J.D., Vidal, E.R., Rotter, A., Muhlenberg, M., 2008. Lane-change decision aid system based on motion-driven vehicle tracking. IEEE Trans. Veh. Technol. 57 (5), 2736–2746.

Barth, A., Franke, U., 2009. Estimating the driving state of oncoming vehicles from a moving platform using stereo vision. IEEE Trans. Intell. Transp. Syst. 10 (4), 560–571.

Barth, A., Franke, U., 2010. Tracking oncoming and turning vehicles at intersections. 13th International IEEE Conference on Intelligent Transportation Systems 861–868.

Blanc, N., Steux, B., Hinz, T., 2007. Larasidecam: a fast and robust vision-based blindspot detection system. 2007 IEEE intelligent vehicles symposium 480–485.

Cambridge, U., 2007. Camseq01 Dataset: Cambridge Labeled Objects in Video. University of Cambridge, UK. http://mi.eng.cam.ac.uk/research/projects/VideoRec/CamSeq01/.

Chakraborty, D.B., Pal, S.K., 2017. Neighborhood rough filter and intuitionistic entropy in unsupervised tracking. IEEE Trans. Fuzzy Syst. 26 (4), 2188–2200.

Chakraborty, D., Shankar, B.U., Pal, S.K., 2013. Granulation, rough entropy and spatiotemporal moving object detection. Appl. Soft Comput. 13 (9), 4001–4009.

Chanawangsa, P., Chen, C.W., 2013. A novel video analysis approach for overtaking vehicle detection. 2013 International Conference on Connected Vehicles and Expo (ICCVE) 802–807.

Cheon, M., Lee, W., Yoon, C., Park, M., 2012. Vision-based vehicle detection system with consideration of the detecting location. IEEE Trans. Intell. Transp. Syst 13 (3), 1243–1252.

Cherng, S., Fang, C.-Y., Chen, C.-P., Chen, S.-W., 2009. Critical motion detection of nearby moving vehicles in a vision-based driver-assistance system. IEEE Trans. Intell. Transp. Syst. 10 (1), 70–82.

Cocca, P., Marciano, F., Alberti, M., 2016. Video surveillance systems to enhance occupational safety: a case study. Saf. Sci. 84, 140–148.

Conche, F., Tight, M., 2006. Use of cctv to determine road accident factors in urban areas. Accid. Anal. Prev. 38 (6), 1197–1207.

Datondji, S.R.E., Dupuis, Y., Subirats, P., Vasseur, P., 2016. A survey of vision-based traffic monitoring of road intersections. IEEE Trans. Intell. Transp. Syst. 17 (10), 2681–2698.

de Naurois, C.J., Bourdin, C., Bougard, C., Vercher, J.L., 2018. Adapting artificial neural networks to a specific driver enhances detection and prediction of drowsiness. Accid. Anal. Prev. 121, 118–128. https://doi.org/10.1016/j.aap.2018.08.017.

de Oña, J., López, G., Abellán, J., 2013. Extracting decision rules from police accident reports through decision trees. Accid. Anal. Prev. 50, 1151–1160.

Diehl, C.P., 2000. Toward efficient collaborative classification for distributed video surveillance. Citeseer. Ph.D. thesis.

Foundation, S.L., 2020. Road Crash Statistics 2017. https://savelifefoundation.org/wp-content/uploads/2018/10/Road-Crash-Statistics-2017.pdf.

Geiger, A., Kitt, B., 2010. Object flow: a descriptor for classifying traffic motion. 2010 IEEE Intelligent Vehicles Symposium 287–293.

Gindele, T., Brechtel, S., Dillmann, R., 2010. A probabilistic model for estimating driver behaviors and vehicle trajectories in traffic environments. 13th International IEEE Conference on Intelligent Transportation Systems 1625–1631.

Girshick, R., Donahue, J., Darrell, T., Malik, J., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition 580–587.

Girshick, R., 2015. Fast r-cnn. Proceedings of the IEEE international conference on computer vision 1440–1448.

Goh, Y.M., Ubeynarayana, C.U., 2017. Construction accident narrative classification: An evaluation of text mining techniques. Accid. Anal. Prev. 108, 122–130. https://doi.org/10.1016/j.aap.2017.08.026.

Ha, J.-E., Lee, W., 2010. Foreground objects detection using multiple difference images. Opt. Eng. 49 (4), 047201.

Jazayeri, A., Cai, H., Zheng, J.Y., Tuceryan, M., 2011. Vehicle detection and tracking in car video based on motion model. IEEE Trans. Intell. Transp. Syst. 12 (2), 583–595.

Jiang, F., Wu, Y., Katsaggelos, A.K., 2009. A dynamic hierarchical clustering method for trajectory-based unusual video event detection. IEEE Trans. Image Process. 18 (4), 907–913.

Jo, K.-H., et al., 2017. Islab-pvd: Illegally Parked Vehicle Dataset. http://mi.eng.cam.ac.uk/research/projects/VideoRec/CamSeq01/.

Karami, E., Prasad, S., Shehata, M., 2017. Image Matching Using Sift, Surf, Brief and Orb: Performance Comparison for Distorted Images arXiv preprint arXiv:1710.02726.

Kongsvik, T., Fenstad, J., Wendelborg, C., 2012. Between a rock and a hard place: accident and near-miss reporting on offshore service vessels. Saf. Sci. 50 (9), 1839–1846.

Kratz, L., Nishino, K., 2009. Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models. 2009 IEEE Conference on Computer Vision and Pattern Recognition 1446–1453.

Kumar, T., Kushwaha, D.S., 2016. An efficient approach for detection and speed estimation of moving vehicles. Procedia Comput. Sci. 89 (2016), 726–731.

Lee, J.T., Ryoo, M.S., Riley, M., Aggarwal, J., 2009. Real-time illegal parking detection in outdoor environments using 1-d transformation. IEEE Trans. Circuits Syst. Video Technol. 19 (7), 1014–1024.

Mackenzie, C., Xiao, Y., Hu, P., Seagull, F.J., Hammond, C., Bochicchio, G., Chiu, W., O'Connor, J., Gerber-Smith, L., Dutton, R., 2002. Video clips as a data source for safety performance. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol. 46. SAGE Publications Sage CA, Los Angeles, CA, pp. 1414–1417.

Matthews, G.A., 1997. The values-based safety process: improving your safety culture with a behavioral approach. Prof. Saf. 42 (8), 40.

Monteiro, G., Ribeiro, M., Marcos, J., Batista, J., 2007. Wrongway drivers detection based on optical flow. 2007 IEEE International Conference on Image Processing, Vol. 5 pp. V-141.

Moqqaddem, S., Ruichek, Y., Touahni, R., Sbihi, A., 2011. A spectral clustering and kalman filtering based objects detection and tracking using stereo vision with linear cameras. 2011 IEEE Intelligent Vehicles Symposium (IV) 902–907.

Pramanik, A., Gorai, A., Sarkar, S., Gupta, P., 2018. A novel feature extraction-based human identification approach using 2d ear biometric. 2018 IEEE Applied Signal Processing Conference (ASPCON) 168–172.

Pramanik, A., Sarkar, S., Maiti, J., 2019. Oil spill detection using image processing technique: an occupational safety perspective of a steel plant. Emerging Technologies in Data Mining and Information Security 247–257.

Paul, P.S., Maiti, J., Dasgupta, S., Forjuoh, S.N., 2005. An epidemiological study of injury in mines: implications for safety promotion. Int. J. Inj. Control Saf. Promot. 12 (3), 157–165. https://doi.org/10.1080/15660970500088763.

Pramanik, A., Djeddi, C., Sarkar, S., Maiti, J., et al., 2020. Region proposal and object detection using hog-based cnn feature map. 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI) 1–5.

Pramanik, A., Sarkar, S., Maiti, J., Mitra, P., 2021a. RT-GSOM: Rough Tolerance Growing Self-Organizing Map. Inf. Sci. https://doi.org/10.1016/j.ins.2021.01.039.

Pramanik, A., Pal, S.K., Maiti, J., Mitra, P., 2021b. Granulated rcnn and multi-class deep sort for multi-object detection and tracking. IEEE Trans. Emerg. Topics Comput. Intell. https://doi.org/10.1109/TETCI.2020.3041019.

Rad, A.G., Dehghani, A., Karim, M.R., 2010. Vehicle speed detection in video image sequences using cvs method. Int. J. Phys. Sci. 5 (17), 2555–2563.

Raptis, M., Sigal, L., 2013. Poselet key-framing: a model for human activity recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2650–2657.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 779–788.

Sarker, M., Mostafa, K., Weihua, C., Song, M.K., 2015. Detection and recognition of illegally parked vehicles based on an adaptive gaussian mixture model and a seed fill algorithm. J. Inf. Commun. Converg. Eng. 13 (3), 197–204.

Sarkar, S., Pateshwari, V., Maiti, J., 2017. Predictive model for incident occurrences in steel plant in India. 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT) 1–5.

Sarkar, S., Ejaz, N., Maiti, J., 2018a. Application of hybrid clustering technique for pattern extraction of accident at work: A case study of a steel industry. 4th International Conference on Recent Advances in Information Technology (RAIT) 1–6. https://doi.org/10.1109/RAIT.2018.8389052.

Sarkar, S., Verma, A., Maiti, J., 2018b. Prediction of occupational incidents using proactive and reactive data: a data mining approach. Industrial Safety Management. Springer, Singapore, pp. 65–79.

Sarkar, S., Raj, R., Vinay, S., Maiti, J., Pratihar, D.K., 2019. An optimization-based decision tree approach for predicting slip-trip-fall accidents at work. Saf. Sci. 118, 57–69. https://doi.org/10.1016/j.ssci.2019.05.009.

Sarkar, S., Pramanik, A., Maiti, J., Reniers, G., 2020. Predicting and analyzing injury severity: a machine learning-based approach using class-imbalanced proactive and reactive data. Saf. Sci. 125, 104616.

Sarkar, S., Maiti, J., 2020. Machine learning in occupational accident analysis: A review using science mapping approach with citation network analysis. Saf. Sci. 131, 104900. https://doi.org/10.1016/j.ssci.2020.104900.

Satzoda, R.K., Trivedi, M.M., 2014. Overtaking & receding vehicle detection for driver assistance and naturalistic driving studies. 17th International IEEE Conference on Intelligent Transportation Systems (ITSC) 697–702.

Saunier, N., Sayed, T., 2006. A feature-based tracking algorithm for vehicles in intersections. The 3rd Canadian Conference on Computer and Robot Vision (CRV'06), pp. 59–59.

Schaeffer, J., Lu, P., Szafron, D., Lake, R., 1993. A re-examination of brute-force search. Proceedings of the AAAI Fall Symposium on Games: Planning and Learning 51–58.

Shen, J., Liang, Z., Liu, J., Sun, H., Shao, L., Tao, D., 2018. Multiobject tracking by submodular optimization. IEEE Trans. Cybern. 49 (6), 1990–2001.

Sivaraman, S., Morris, B., Trivedi, M., 2011. Learning multi-lane trajectories using vehicle-based vision. 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops) 2070–2076.

Sobral, A., Vacavant, A., 2014. A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos. Comput. Vis. Image Underst. 122, 4–21.

Souri, A., Ghafour, M.Y., Ahmed, A.M., Safara, F., Yamini, A., Hoseyninezhad, M., 2020. A new machine learning-based healthcare monitoring model for student's condition diagnosis in Internet of Things environment. Soft Comput. 24, 17111–17121. https://doi.org/10.1007/s00500-020-05003-6.

Stauffer, C., Grimson, W.E.L., 1999. Adaptive background mixture models for real-time tracking. Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149), Vol. 2 246–252.

Sun, Z., Bebis, G., Miller, R., 2006. Monocular precrash vehicle detection: features and classifiers. IEEE Trans. Image Process. 15 (7), 2019–2034.

Thomas, S.S., Gupta, S., Subramanian, V.K., 2017. Event detection on roads using perceptual video summarization. IEEE Trans. Intell. Transp. Syst. 19 (9), 2944–2954.

WHO, 2020. Road Traffic Injuries. https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries.

Wiest, J., Höffken, M., Kreßel, U., Dietmayer, K., 2012. Probabilistic trajectory prediction with gaussian mixture models. 2012 IEEE Intelligent Vehicles Symposium 141–146.

Wiliem, A., Madasu, V., Boles, W., Yarlagadda, P., 2008. Detecting uncommon trajectories. 2008 Digital Image Computing: Techniques and Applications 398–404.

Wojke, N., Bewley, A., Paulus, D., 2017. Simple online and realtime tracking with a deep association metric. 2017 IEEE International Conference on Image Processing (ICIP) 3645–3649.

Wu, J., Liu, Z., Li, J., Gu, C., Si, M., Tan, F., 2009. An algorithm for automatic vehicle speed detection using video camera. 2009 4th International Conference on Computer Science & Education 193–196.

Yilmaz, A., Javed, O., Shah, M., 2006. Object tracking: a survey. Acm Comput. Surv. (CSUR) 38 (4), 13–20.

Zhang, Y., Liu, Z.-J., 2007. Irregular behavior recognition based on treading track. 2007 International Conference on Wavelet Analysis and Pattern Recognition, Vol. 3 1322–1326.

Zhu, Y., Comaniciu, D., Pellkofer, M., Koehler, T., 2006. Reliable detection of overtaking vehicles using robust information fusion. IEEE Trans. Intell. Transp. Syst. 7 (4), 401–414.

Zivkovic, Z., 2004. Improved adaptive gaussian mixture model for background subtraction. Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004., Vol. 2 28–31.