IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received January 26, 2022, accepted February 10, 2022, date of publication February 15, 2022, date of current version February 25, 2022.
Digital Object Identifier 10.1109/ACCESS.2022.3151973

PredictV: A Vehicle Prediction Scheme to Circumvent Occluded Frames

SHIKHA TUTEJA, SHASHI PODDAR, DIVYA AGRAWAL, AND VINOD KARAR
Academy of Scientific and Innovative Research (AcSIR), Ghaziabad 201002, India
Central Scientific Instruments Organisation (CSIO), Chandigarh 160030, India
Corresponding author: Shikha Tuteja (shikhatuteja85@gmail.com)

ABSTRACT Although several methods to overcome traffic congestion on the roads exist, occlusion is still a major bottleneck in processing the traffic images and need researchers attention. In the rural and urban areas, heavy congestion on the roads has become the leading cause of occlusion. The PredictV method works on the prediction principle based on existing values of blob width and height. This scheme uses blob detection for the first frame and predict other vehicles based on percentage increment in the different blob parameters to locate the vehicle in the next frame. With the help of this approach, occlusion is handled easily and it prefers to use predicted points to detect vehicles on the frame. The suggested approach is implemented via a MATLAB simulator. It is tested on a large dataset containing 7152 frames of 6 different videos from the Urban Tracker and KoPer datasets. In total, there are 46876 vehicles present on the frame at first, and the blob detection and Kalman filtering approaches have detected 58% and 73% of vehicles respectively, whereas the detection rate is 82% with this proposed approach. The occlusion rate is only 17% on average, which has been reduced via this proposed approach and was 24% and 43% using previous approaches. The performance of this method has been quantitatively measured using MOTA and MOTP parameters and is found to have very good accuracy in vehicle position prediction.

INDEX TERMS Traffic monitoring, vehicle detection, occlusion, prediction.

I. INTRODUCTION

Nowadays, traffic on the roads has increased due to various reasons where the increase in population is the foremost reason. Also, the preference given to personal vehicles becomes the cause of this issue. To avoid congestion or handle the issues generated by congestion, intelligent traffic monitoring systems are now trending. These systems help to analyze the traffic on the roads and for other processes. Over the past several years, these traffic monitoring systems capture the video using different equipment like cameras, sensors, and many others to inspect the number of vehicles on the road.

The methods developed for traffic monitoring systems based on Image processing systems included vehicle detection and tracking. Many researchers proposed techniques for the same purpose. Like to track vehicles, [1] use the Mean-Shift algorithm, but this method does not detect occlusion so, the GMM model has been used to predict the vehicle, which works based on training. With this method, they might track vehicles, but the computations gradually increased and became the major drawback for the proposed method. [2] utilized this mean shift algorithm earlier but used it on the targeted blocks' subblocks. They also calculated the confidence values to find out the exact position of the targets. Similarly, the other technique is developed by [3], where they also use the concept of division of the blocks into subblocks, but they calculated weight values and make use of these values. The results of this proposed method were not good when tested in real-time systems.

Occlusion becomes one of the primary causes of not detecting vehicles accurately from the frames. So, to detect vehicles, different methods have been proposed, even in the presence of occlusion. Like, SIFT features-based approach and template matching proposed by [4]. The other feature-based approach was proposed by [1], [5], and [6]. In this, they solve only the partial occlusion problem using the RDHOGPF feature. The different geometric methods were proposed

The associate editor coordinating the review of this manuscript and approving it for publication was Gulistan Raja.

by [7]–[9] to deal with the occlusion. These are based on the concave points based on which they divided the blocks into sub-blocks and then removed the occlusion. These methods are also very computationally expensive.

It means that occlusion is still an issue for vehicle tracking systems, and researchers should pay attention. So, this work focuses on the same and proposed a novel approach to predicting the vehicles to avoid occlusion. The next section of the paper also discusses existing prediction-based techniques developed by the researchers, followed by the proposed approach and the pseudo-code in the next section. In the end, discussion of the results is done based on different parameters.

II. RELATED WORK

To detect the vehicles on the roads precisely for different purposes is the primary intention of the research. Many researchers have tried different methods to identify the vehicles on the roads, but the result was not the same. In this part, the work is going to be presented, which is done by researchers. Different elements are taken into consideration while proposing these prediction-based methods.

A. CLASSIFICATION-BASED APPROACHES

Reference [10] have recently recommended a two-stage data-driven method. The dataset's generation was done using a safety pilot model development program from which features and labels were extracted using summary statistics and Fast Fourier transform methods. For the selection of optimized features, a random forest algorithm was used. In a two-stage methodology, the raw prediction was first based on the decision tree method followed by the Gaussian mixture model. With the help of clustering, the GMM assigned the weights to predict vehicles on its base. Motion factor is the main element in this suggested approach which signifies the vehicles' movement from frame to frame. The dataset used for training has 51 unique vehicles out of 587812 instances. Confusion metrics were used to figure out results from this method that defines the accuracy of the prediction. The results of this approach were assessed with and without PCA, and it had observed that more balanced prediction results were taken with PCA. Effectiveness had been shown with the implementation of this technique. The computation cost of the GMM model, along with its training, is high because it has effectual output. Some other techniques had been proposed [11], in that a lane change behavior and proposed a maneuver algorithm had been considered to estimate the existence of a longitudinal trajectory. It was used to allow the lateral motion of a lane change maneuver. Some assumptions are used in the lane change maneuver algorithm to find the position of the vehicles, such as a sensor system. In addition, to follow the planned trajectories, prediction, and decision-making systems, the low-level control systems had been used. The main focus of this suggested method was to select an appropriate inter-vehicle traffic gap and time instance for initializing a lane change maneuver's lateral motion, which generates safe and smooth lane change trajectories. The complexity was increased due to multiple constraints in it even though this method has worked well.

Similarly, to detect the lane-change behavior and prediction of vehicles, the SVM-based approach [12] was design based on the feature vector. Firstly, the sensor system was installed in the primary vehicle for implementing this proposed system, consisting of a position sensor (RT3003) and six laser scanners (ibeo LUX). To determine when or when not a target vehicle would change the lane, these input sensors help. Moreover, this technique is made up of driving-intention estimation and vehicle-trajectory prediction. At first, there are three different features: the distance from the centreline, the lateral velocity, and the potential feature combined to form a feature vector. To change the lane, strategies are identified by the vehicle- trajectory prediction. The performance of the suggested technique was evaluated based on precision, recall, and f-score. The outcome of this method has proved that it can detect a lane change 1.74 s before the target vehicle crosses the centreline on average, and its accuracy is 98.1%.

For the motion prediction [7] the other unsupervised-based learning method had been utilized. For training, camera phones are mounted on the cars' dashboards and performed the city-scale structure-from-motion to restructure the trajectories accurately. For localization, this data can be collected in which a novel fully-fledged pipeline was used to track the positions of neighboring cars continuously around a camera-equipped vehicle. It has been used as a situation-awareness module to plan algorithms used to predict and react to the motion of the other traffic participants. Moving further, this has been combined with a CNN-based car detector for proposing a motion detector system. To implement the system's performance, 10M pictures of the dataset have been collected, and as an outcome, for the prediction of the vehicles efficiently, a high-accuracy localization was performed.

In addition, for learning the vehicles' movements on roads better via prediction, the Deep reinforcement learning technique [13] was also used. In this work, two different methods were used to investigate the abilities of the DQNs to learn exploratory behavior, and for this, they used Sumo Simulator. For optimization, the RMSProp algorithm was utilized. The experimentation had been done with and without occlusion, the dataset to calculate the %Success, %Occlusion, Average time, and Average break. The average DQN Time-to-Go before reaching the goal was 28% faster than TTC, while DQN Sequential was 19% faster than TTC. On that account, the DQN approaches are proved more beneficial to reduce congestion on roads because these are navigating intersections effectively. In an occluded environment, their simulation results were not productive. Reference [14] proposed the other probabilistic approach To classify the occlusion region. In this method, three steps are involved, classification and motion prediction followed by risk assessment in the end. In this, the performance had been evaluated even in the occluded real-time environments. For evaluating the risk, the alleyway scenario was applied in the simulation. The chances of collision risk in the potential collision vehicle are estimated by collision speed range and the probabilistic vehicle speed model.

Deep learning techniques are widely used for different systems and for avoiding occlusion on roads; this approach was figured by [15] for vehicle prediction. Bayesian filtering technique is used firstly after the grid map is passed to the Deep convolutional neural network to predict the vehicle's velocity. For improving the system's performance, automatic labels and some other features are added, proving positive for the system as it has been found that the proposed system can model intricate interactions. The other neural network, along with the wavelet-based prediction model, was proposed by [16]. In this particular technique, the decomposition has been done after that, data is divided into high and lower sequences, and in the end, it is reconstructed accordingly using wavelet. Then Gated Recurrent Unit model and ARMA have been used to train low-frequency sequences and high-frequency sequences. Hence, the results of these two methods have been combined to get a single vector that defines the system's effectiveness for predicting prediction accuracy.

B. FEATURE AND FILTER-BASED APPROACHES

The other method has been suggested by [17] for the accurate prediction of vehicles. The proposed algorithm uses a Kalman filter and fuses the information provided by it to segment vehicles under the occlusion state and effectively track vehicles. Targets had been achieved with the help of segmentation. To evaluate the performance of the recommended system, it was tested for video via a scale of 1920 $\times$ 1080 and a frame rate of 25. The vehicle trajectories extracted by the algorithm in this paper were compared with manual calibration to verify the algorithm's effectiveness. The comparison of the proposed method was made with the BDM method, and it was observed from the results that the proposed method outperforms the BDM method. The computational complexity of BDM is lower than that of the proposed algorithm; even its performance is not up to the mark. So, even with the issue of complexity, its performance was considered suitable. To predict vehicles, the other Kalman filter-based approach was proposed by [18], and in that method, the first frame, the center point, and the radius of the reference object are selected.

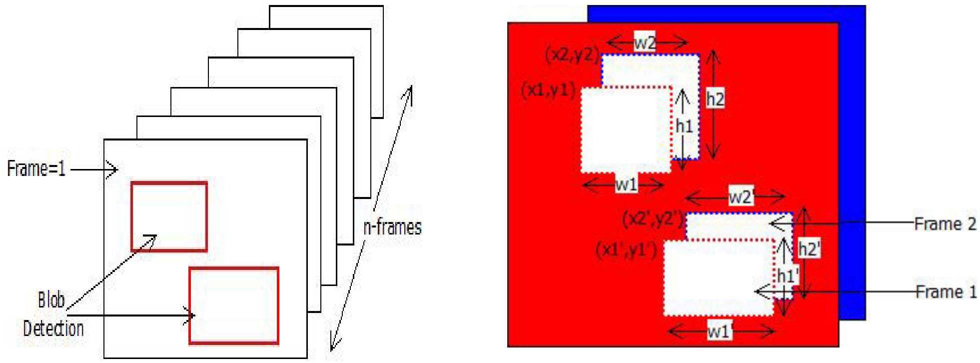
The suggested approach uses a Kalman filter for estimating the object's center in the next frame. The previous information has been used to determine the estimation of the object to predict its trajectory, and they also use the feature of DRIFT in the case of partial occlusion. To identify the object from the given sample, these features are considered the main factor in this technique. With feature extraction, intensity-based histograms had been used; after that, partition was done using Gaussian ringlets. In addition to this, to point out the central rings, weight factor was applied to histograms, and four rings had been used for this evaluation. 32 bin histograms were made using each ring, and to reduce the feature size without losing its accuracy, the ring number and histogram size are selected. In this work, to improve the illumination in the imagery, a nonlinear image enhancement function was also used, allowing better tracking when the object is occluded by very harsh shadows that negatively affect classification. Here, the self tunable transformation function (STTF) enhancement algorithm was selected as the pre-processing algorithm. The benefits of this approach over the other existing methods have been with the results. One of the other approaches based on the Kalman filter also utilizes the invariant descriptor presented by [19]. Invariant feature descriptors were extracted in this technique, which was integrated with an adaptive prediction model afterward. The testing had been performed using the PETS'2001 dataset, containing samples with a resolution of 768 $\times$ 576 pixels. This suggested method had performed very well in partial and complete occlusion conditions compared to the conventional Kalman Filter technique. To solve the problem of occlusion

Reference [20] also proposed a filter-based prediction approach to use particle filters and the KCF algorithm [21] and solve occlusion. Some of the prediction approaches are based on background. For Instance, the taillight adjustment-based approach that adjusts the bounding boxes were proposed by [22]. In this algorithm, color features, Mean Shift Vector, Scale Adjustment, and color histogram had been used to extract the samples' information. Moreover, a back-propagation neural network had been trained by these feature set for vehicle candidate verification. For testing, 104 images had been used, and the simulation was also done using MATLAB. Overall, the performance was good, but it gets affected when vehicles are of the same color.

C. OTHER METHODS

The intersection method has been used to locate vehicles by using some other prediction approaches, and the same technique was proposed by [6]. For calculating the weight factor and moving objects under occlusion, the IoT scheme with a virtual measurement model has been extended. They have also validated the performance of this suggested system with the use of a 2D LODAR and 3D LIDAR-based benchmark system. The accuracy of the prediction has improved along with the improved performance of this recommend technique, even in the presence of occlusion.

For tacking multiple vehicles at one time on the road, an adaptive partial occlusion segmentation method (APPOS) had been proposed by [23]. First of all, the three-frame difference method has been implemented via foreground extraction on video frame sequence. With this, the vehicles can be obtained by threshold-based segmentation to locate their regions of closed contours. The threshold had been determined by value tuning on different videos after that occlusion region was noticed with that guess, their foregrounds were going to be merged when vehicles occluded. In the next step, the occluded region had been segmented then, based on candidate region and followed by the histogram-based method to improve the Overlaying optical flow level. Moving further, the counter's optical flow had been applied for rough



(a) Objects on the First Frame (b) Objects on two different Frames

FIGURE 1. **Position tracking-initial steps.**

evaluations of the occluded vehicles. The end step of this method is line scanning which uses color contrast for locating vehicles. There is a total of four videos that have been used for evaluating the performance of this suggested technique. A fixed camera 720 × 488 had been used to capture the resolution of the practical traffic video. The performance parameters were opted by the researchers with accuracy and the effectiveness of the occlusion segmentation under different traffic scenarios. This method proved beneficial as it was good to handle the detected occlusion, revealed with results.

Reference [24] proposed a very different approach where occlusion prediction was computed based on position, width, length, and angle. This method uses these predictions to track the occlusion state, and it has been tested for the vehicle and the pedestrians, making this approach unique and compelling. This method is computationally high, and this is the only pitfall of it.

There are different methods and factors considered by these above-discussed methods used for the prediction purpose. The existing techniques have given worthy outcomes for tracking vehicles that are based on predictions. There are some drawbacks: high computation cost, pretty expensive, the vehicle's color, and many others still present in the detection/tracking systems. While considering these factors, researchers should focus on an efficient approach to track vehicles on roads.

III. PROPOSED WORK

To handle the occlusion issue, this work has been suggested for developing an efficient vehicle detection technique. Many methods are proposed for detecting vehicles from the road traffic, but their accuracy has been reduced due to occlusion. In this work, an improved blob detection method has been used for the new approach named PredictV, and this proposed work is divided into several steps, and these are:

Step 1 (Frame Extraction and Cleaning): This step is the most common step of this approach, followed by each researcher for extracting sample data from the video. In this work, for further steps, the video frame has been extracted and stored. To improve the recognition system's performance, data cleaning must be considered the most crucial step in which an anisotropic filter [25] is used to enhance an image sample's features. Then for background subtraction, the SubSense [26] method is used. For object enhancement and data cleaning, some morphological operations like dilation, opening, and erosion are also utilized for data cleaning in this particular work.

Step 2 (Vehicle Detection): In this step, the Region prop method uses cleaned data to detect the frames' object, and it is commonly used to detect the objects on the given input image [27]. This method has also provided the number of objects which are used for different purposes

Step3 (Position Tracking): This step is an essential step for the proposed approach as it provides the value for predicting the objects on the frames. This step consists of various substeps, as explained below:

A. VEHICLE DETECTION FOR THE FIRST FRAME

The blob detection method has been used to detect and plot the first frame object. Fig 1(a) shows the detected objects on the first frame out of n number of frames.

B. POSITION CHANGE CALCULATION

For every object on the frame, the change in the positions and the objects' height and width are calculated in this step. This step is repeated for every frame, and then the average of the calculated values will be used for the next steps. Fig 1 (b) shows the change in two vehicles' positions with time 't' from frames 1 to 2. Then the change of points from the vehicle on one frame to the vehicle on the next frame till the end of the frame, i.e., Δx, Δy, Δw, Δh, is calculated as:

$$\left.\begin{array}{c} \Delta x = x_2 - x_1 \\ \Delta y = y_2 - y_1 \\ \Delta w = w_2 - w_1 \\ \Delta h = h_2 - h_1 \\ \text{For Instance, Let, } x_1 = nx_2 = n_1 \text{ So,} \\ \Delta x = x_2 - x_1 => n_1 - n \end{array}\right\} \tag{1}$$

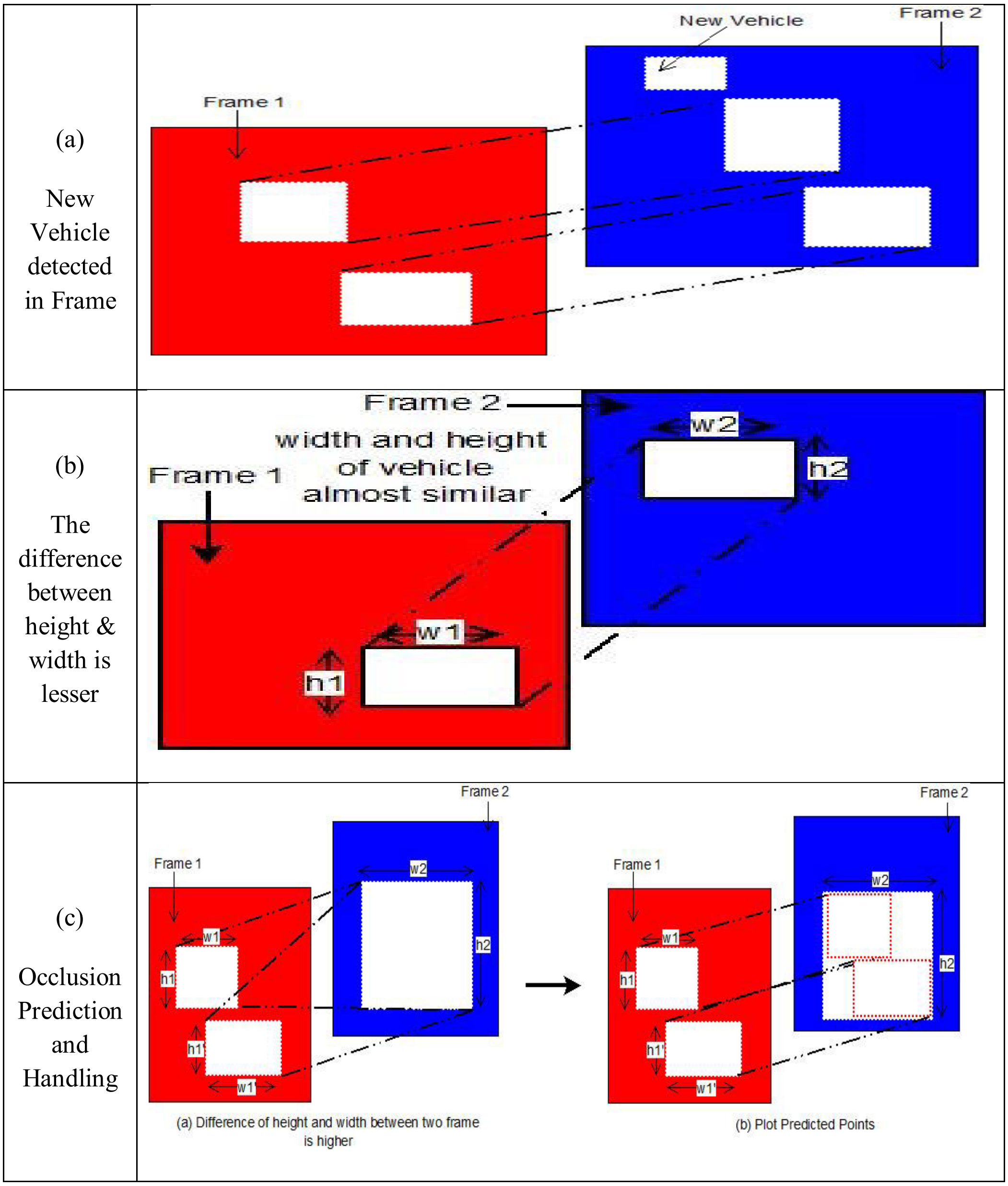


FIGURE 2. Vehicle prediction phases.

Similarly, the values will be calculated for the x-axis, where points are $x_1, x_2, x_3, \ldots\ldots\ldots\ldots\ldots\ldots.., x_n$, and find the average value, and similarly, calculations will be done for the y-axis, height, and width.

C. AVERAGE POSITION CHANGE CALCULATION

The average change value for all the points can be calculated and used for calculating Average Position Change in percentage. This step will help predict the objects accurately than the average value because the average value can vary from vehicle to vehicle. Still, percentage change will opt and predict the value as per the current vehicle. For Instance, Let, Average change in x is m, so, Percentage calculation will be done using the following equation:

$$\left.\begin{aligned} x_1 * k &= m \\ k &= \frac{m}{x_1} \end{aligned}\right\} \tag{2}$$

K is the percentage value for increment, so, by using this method, the average change for all x-axis, y-axis, height, and width can be calculated.

The pseudo-code for the calculation of percentage change in the vehicle and its position is as given below:

Step 4 (Vehicle Prediction): This step is the final step of the suggested technique that uses step 3 to predict the frames' actual position using previously detected objects. The following equation is used to predict the new objects on

Algorithm 1 AVG_CHANGE

INPUT: FRAMES(I) WHERE I=1, 2, 3, ….., N

```
Begin
  For each frame, i=1 to n
    Capture the frame, im← frame(i)
    Perform Vehicle Detection, val=blob_detect(im)
    For each vehicle on the frame, i=1 to length(val)
      Check the detected points, pt = val
      if i>1
        calculate the change in values, Δpt = pt − pt1
        p(i − 1) = Δpt
        pt1=pt
      else
        pt1 = pt
      End if
    End For
  End for
  Calculate Average Position Change, Avg=average(p)
  Calculate Percentage Change, perc = avg/pt1
  Return(perc)
  Stop
```

Algorithm 2 PREDICT V

INPUT: FRAMES(I) WHERE I=1, 2, 3, ….., N

```
Begin
Call Algorithm 1, Perc=avg_change();
  For each frame, i=1 to n
  Capture a frame, Im← frame(i)
    If(i>1)
      Perform Vehicle Detection,
      M=blob_detect(i)
     Calculate number of Vehicles,
      num1=length(M)
     Perform the following steps for prediction of
      vehicle's location:
      If(num1 == num)
       K←m(i-1) +perc(m(i-1))
        Plot(k)
      Elseif num1 >num
        Plot (m)
      Elseif num1<num
        diff=m(i)-m(i-1)
        if diff>Th
          k←m(i-1)+perc(m(i-1))
          plot k
        else
          plot m
        end if
      end if
    else
      m=blob_detect(i)
      plot(m)
    end if
  End For
//Function for Plotting
Function plot(Val)
      For j=1 to length(Val)
        Plot rectangle using j points
      End For
End Function
Stop
```

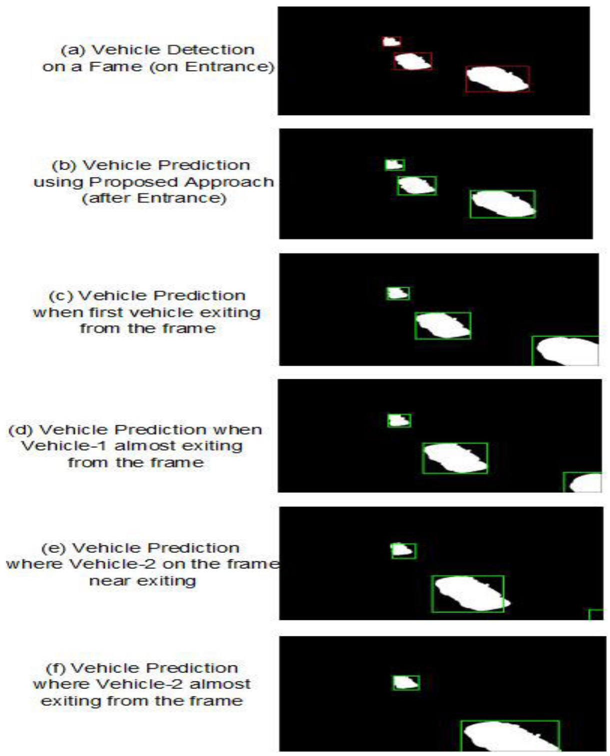


FIGURE 3. Vehicle predictions using predictv (sherbrooke video of urban tracker video dataset).

the frame:

$$x_{new} = x_{old} + k(x_{old}) \tag{3}$$

The above equation will be used to calculate changes in all frames for each vehicle and all x-axis, y-axis, height, and width.

For Instance, the prediction of the x-axis points for n frames are:

$$\left.\begin{array}{l} x_2 = x_1 + \%\Delta x(x_1) \\ x_2 = x_2 + \%\Delta x(x_2) \\ \vdots \\ x_n = x_n + \%\Delta x(x_n) \end{array}\right\} \tag{4}$$

This change will be calculated for every frame's object and x, y-axis, height, and width.

The proposed approach performed well if the number of vehicles in all frames is equal, but it will not be a possible or actual factor in real-time as the number of vehicles can vary even in seconds. So, it is one of the challenges. To cope with this problem and check the number of vehicles in the frame, Blob detection defined in step-2 is used, and PredictV is

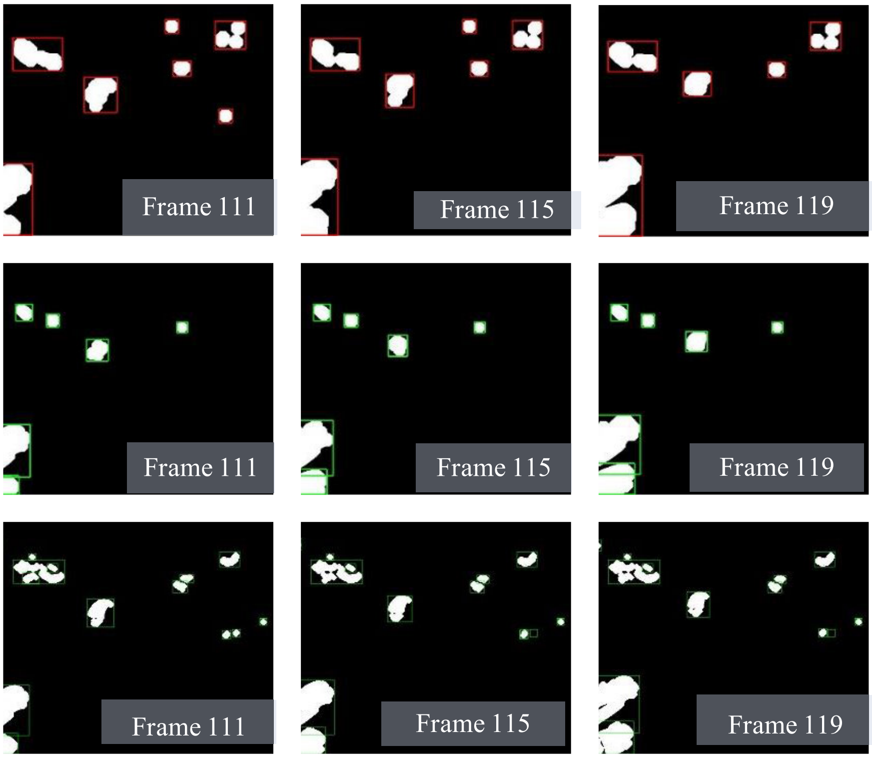


FIGURE 4. (a) Detected and (b) Predicted frames using proposed approach (c) Predicted frame using kalman filter [17] [sequence_1a Ko-PER dataset].

extended towards this problem's solution. The issue described above is further subdivided into two cases:

Case1→ If no. of vehicles in a frame (i) is greater than the frame (i-1), then simply step-2 is called, and blobs on that frame are plotted accordingly. Fig 2(a) clear that frame 1 has 2 objects, whereas in Frame 2, there are 3 objects, so here to detect the new vehicle on the frame, the blob detection scheme will be used. It means for every new object, step-2 will be called.

Case 2→ If no. of vehicles in a frame (i) is lesser than the frame (i-1), then the change in height and width of the detected frame will be compared. While dealing with height and width, it has been noticed that the difference is higher or lesser. So, the following sub-cases will be considered to evaluate the same.

Sub Case-1: If the difference is lesser, then the frame's vehicle is considered a new frame and plotted using step 2.

Sub Case-2: If the difference is higher, it is treated as an occluded frame, and predicted points would be plotted using step 4.

Fig 2(b) represents the process if the difference is lesser, whereas, with a high difference, the process is as depicted in Fig 2(c). From the above figure, it is clear that the proposed PredictV helps to detect and predict the vehicles accurately, even in occlusion, and provides accurate results. The pseudo-code of the proposed PredictV approach is as given below:

IV. SIMULATION RESULTS

The MATLAB simulator is used to simulated the proposed approach. Six datasets, Sherbrooke and Rouen (Urban Tracker Video Dataset [28]) and Sequence1a,1b,1c, and 1d (Ko-PER dataset [27]), have been used for extracting samples. The extracted frame's size is different for different data samples, and all the frames are stored in '.jpg' format. The different qualitative and quantitative parameters are used to evaluate the performance, which is given below:

A. QUALITATIVE ANALYSIS

To predict the vehicles accurately from their entry to exit, this particular PredictV approach helps. So the quality of the suggested method can be checked based on the predictions, usually the entry and exit of the vehicles. It is shown in the figure where the vehicle on the first frame (fig 3 (a)) is detected using the blob detection method. In comparison, other frames detect the vehicle based on the Proposed PredictV approach and the Sherbrooke video of Urban Tracker Video Dataset [23]. Fig 3 (b) depicts the predicted vehicles after the entrance and so on. Fig 3 (c) and (d) show the exiting of vehicle-1 from the frame, and it is predicted accurately.

Similarly, Vehicle -2 exiting is also predicted accurately, as shown in Fig 3 (e) and (f). Therefore, it is clear that the vehicles in frames are predicted accurately by the proposed solution for each frame from the figure. (Urban Tracker Video Dataset [28]-Sherbrooke)

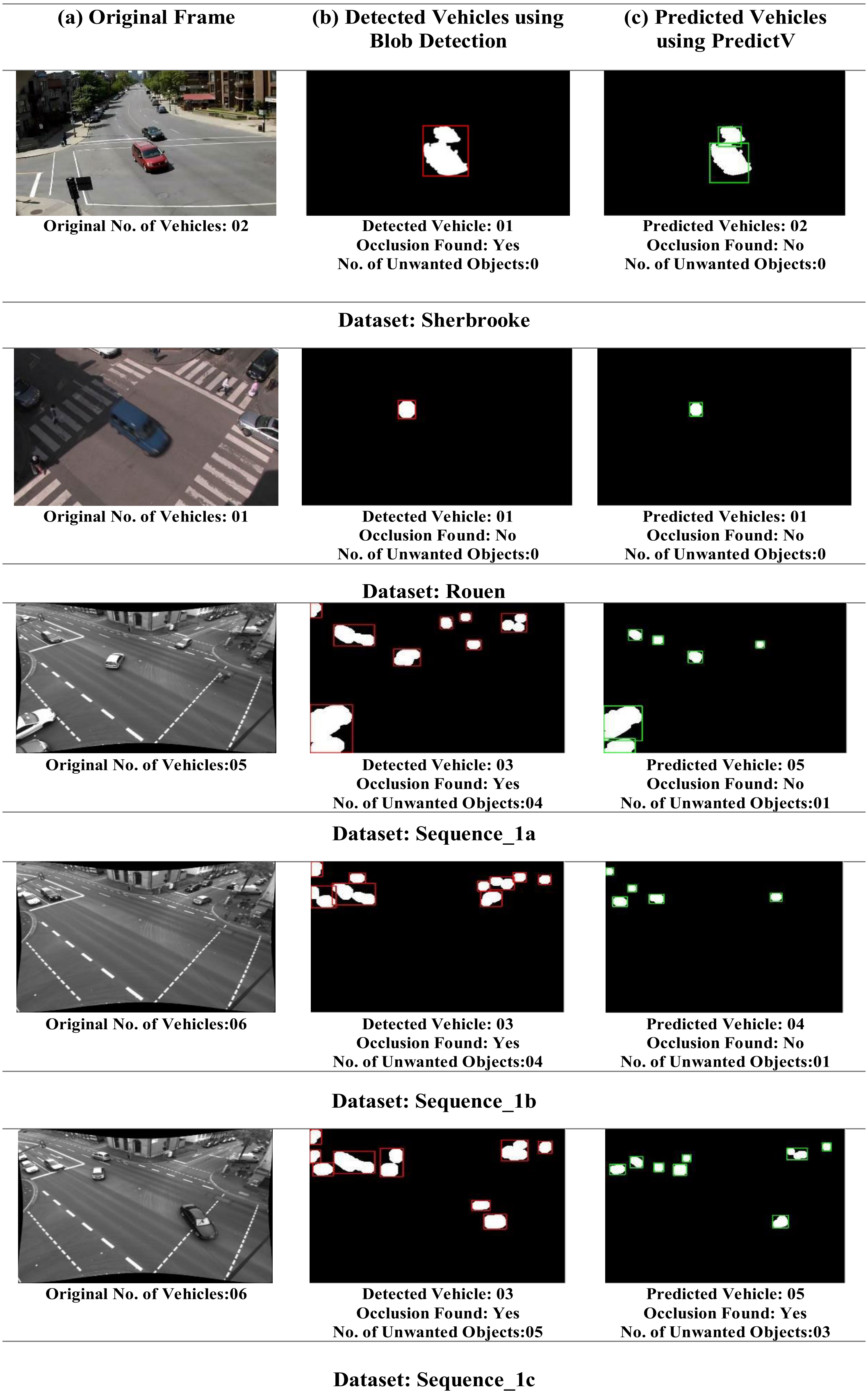


FIGURE 5. Detection and prediction results.

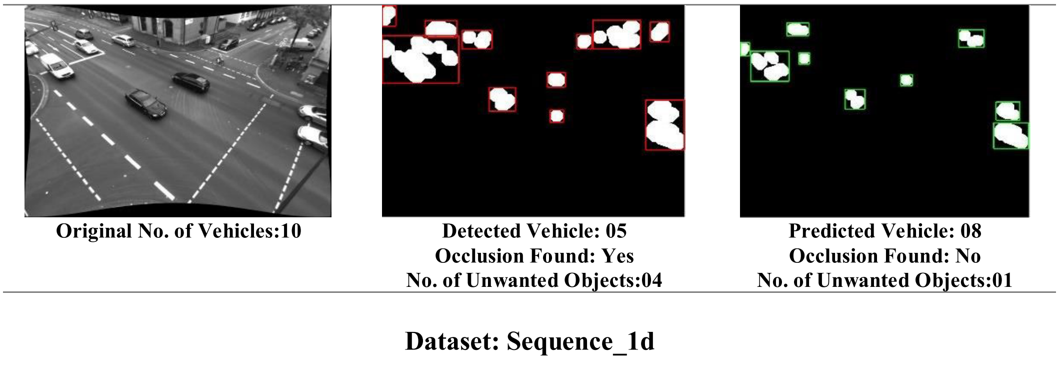


FIGURE 5. *(Continued.)* **Detection and prediction results.**

The other position tracking results have been depicted in fig 4. Sequence_1a (Ko-PER dataset) and frames are extracted using existing blob detection and the proposed PredictV approach in this figure. The figure portrays the occlusion detected on frames 111, 115, and 119 using the existing approach shown in fig 4(a) and 4(c) and is resolved by the prediction method as shown in Figure 4(b). So, the efficiency of the proposed approach in terms of occlusion resolver has been proved via these results.

The quality can also be measured based on the detection/prediction of vehicles. Fig 5 deliberates the detected and predicted vehicles from different datasets using blob detection and the proposed PredictV Approach.

The above results clarify that the prediction-based proposed approach PredictV predicts the vehicles very close to the original number of vehicles on the frames with minimum occlusion and unwanted objects. In contrast, occlusion is almost in every frame of the detection-based approach. This demonstrates the potency of the proposed technique over the existing technique.

B. QUANTITATIVE ANALYSIS

The proposed approach's performance is quantitatively measured based on different parameters like vehicle detection accuracy, occlusion detection, and others frames. For each dataset, parameters are calculated, and the performance is measured for both the existing and proposed approaches. Some of the performance parameters are as demonstrated in table 1.

1) DETECTION ACCURACY

The proposed approach's performance can also be evaluated based on the number of vehicles detected. In this work, six different datasets were used that contain nearly 46876 vehicles on 7152 frames. The proposed approach predicted 37439 vehicles accurately, whereas the existing blob-based detection method detects only 28453 vehicles. It means the performance of the proposed approach is better to detect the vehicles from the frames. The percentage of vehicle detection for different datasets is as depicted in Fig 6.

The above results show that the proposed approach is better than the blob detection by 25.4% and Kalman Filter [17] by 20.6%, as per the predicted results.

2) OCCLUSION

As the prediction-based approach is proposed to reduce the occlusion and detect the vehicle accurately. The approach's performance is also evaluated based on Occluded Frames and is as shown in fig 7.

It has been clear from the above results that the proposed approach has lesser occlusion as compare to the Blob detection technique and is reduced by 60.05%. The results are also improved from the Kalman filter-based prediction [17] approach and are improved by 47.3%.

3) UNWANTED OBJECT DETECTION

Though the proposed approach is developed to detect vehicles like cars, trucks, and other heavy objects, some unwanted objects are also detected. The dataset also consists of other objects like humans, prams, and many other objects. The blob detection method detects 3878 unwanted objects in 7152 frames of different datasets, whereas the proposed scheme performed well and detected only 1550 unwanted objects. The other Kalman Filter based Prediction Approach [17] is also designed to predict the vehicles, but it still faces this challenge, and this approach detects around 1642 unwanted objects. Again it is higher than the proposed technique. The details for each dataset are given in table 2.

These unwanted objects are detected because the proposed scheme uses the blob detection method when the number of vehicles on the frame increases or decreases, as defined in the previous section. Overall, the proposed approach detected 60% lesser unwanted objects than the blob detection method and revealed its virtue.

4) ERROR RATE

The error rate is defined as a difference between the value predicted by the proposed scheme and the vehicle's original position. Here, the value for every 5th frame has been checked

TABLE 1. Quantitative analysis based on different parameters.

Dataset	Video	Number of Frames Extracted	Detection Accuracy (in %)			Occlusion Detection (in %)		
			Blob Detection	Kalman Filter	PredictV	Blob Detection	Kalman Filter	PredictV
Urban Tracker	Sherbrooke	2000	58.35	73.82	79.61	51.60	37.80	21.05
	Rouen	315	48.09	70.08	72.13	0.00	0.00	**0.00**
Ko-PER	Sequence1a	1210	67.07	74.91	80.32	44.63	21.57	17.02
	Sequence1b	1213	57.79	75.13	80.47	41.96	20.77	18.55
	Sequence1c	1209	61.99	77.10	**82.02**	44.00	18.28	**15.72**
	Sequence1d	1205	58.64	72.37	77.46	40.66	19.09	16.68

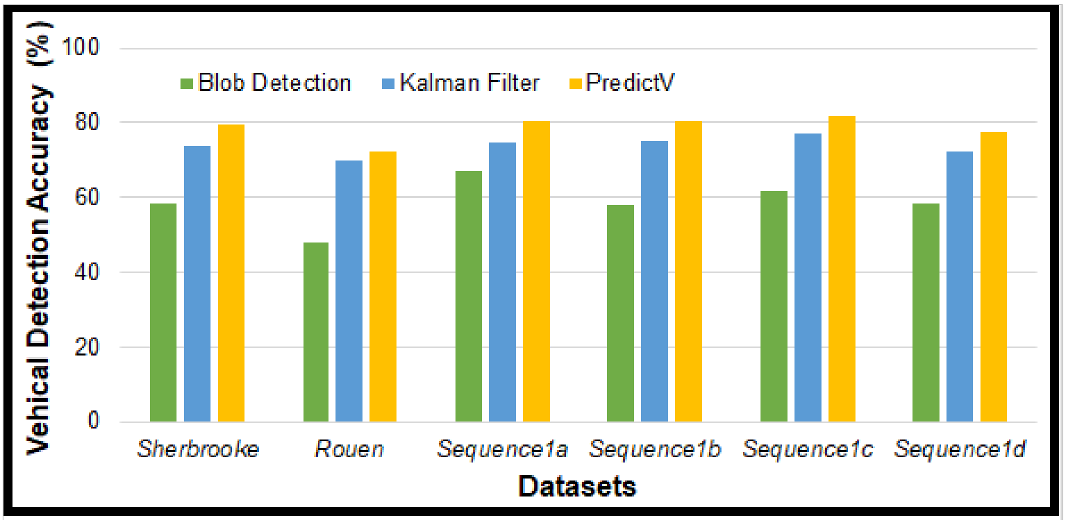


FIGURE 6. Vehicle detection accuracy.

for Sherbrooke [Urban Tracker Dataset] [28], and analyze that the average error rate for each vehicle on the frame is 0.34. The average error rate for each x-axis, y-axis, width, and height has been calculated, and the results are as shown in the table given below.

Fig 8 (a) shows that the error rate can be different for vehicles present in the frame, and even it is different for each point of the rectangle plotted on the frame. This error rate can be further use to evaluate the accuracy of the proposed scheme.

So, the accuracy in predicting the vehicle position of the proposed prediction method is calculated by,

$$Accuracy = 100 - Error\ Rate \tag{5}$$

The accuracy calculated for each x-axis, y-axis, height, and width is as given below:

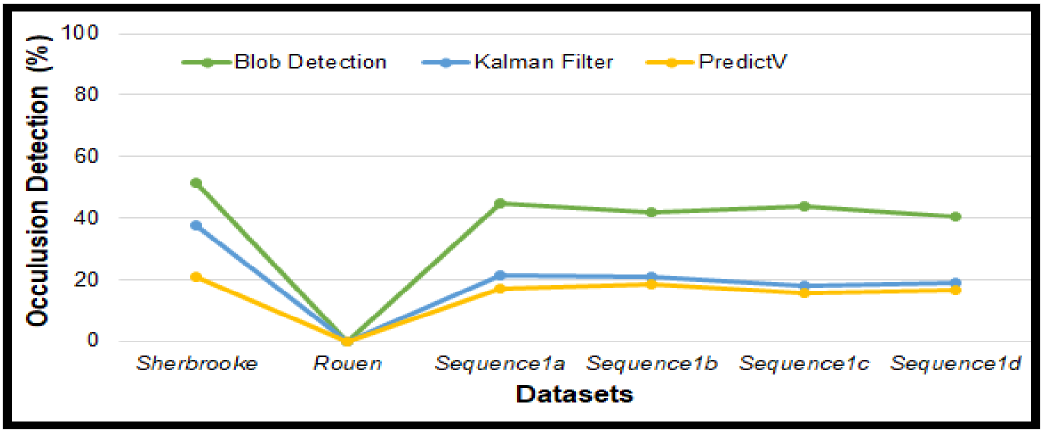


FIGURE 7. Percentage of detected occlusion.

TABLE 2. Number of unwanted objects.

Sr. No.	Video	Number of Frames Extracted	No. of Unwanted Objects		
			Blob Detection	Kalman Filter [17]	PredictV
1	Sherbrooke	2000	971	442	392
2	Rouen	315	75	39	32
3	Sequence1a	1210	713	308	287
4	Sequence1b	1213	690	291	276
5	Sequence1c	1209	702	275	262
6	Sequence1d	1205	727	287	301

TABLE 3. Average error rate (in %) [sherbrooke].

	x-axis	*y-axis*	*width*	*height*
Vehicle 1	0.33	0.27	0.38	0.32
Vehicle 2	0.43	0.29	0.37	0.37
Vehicle 3	0.27	0.26	0.46	0.33

TABLE 4. Accuracy (in %) [sherbrooke].

	x-axis	*y-axis*	*width*	*height*
Vehicle 1	99.67	99.73	99.62	99.68
Vehicle 2	99.57	99.71	99.63	99.63
Vehicle 3	99.73	99.74	99.54	99.67

When evaluated based on original points, the predicted points' accuracy then found that this prediction method accurately predicts most of the vehicles, and overall, 99% accuracy was achieved. For evaluating this, position points for every 10^{th} frame of the database have been compared with the predicted points.

5) MOT METRIC EVALUATION

The CLEAR MOT metrics are calculated to evaluate the proposed approach. MOTA determines the accuracy of multiple object tracking based on the miss rate, false positives, and ID changes. The MOTP method is used to calculate the precision of multiple object tracking based on the total error in the predicted location of the matched items. The following are the definitions of MOTP and MOTA:

$$MOTA = 1 - \frac{\sum_{i,t} e_t^i}{\sum_t m_t}$$

Here e_t^i is the error in estimation of the position of target i at time t and mt is the number of matched targets found at

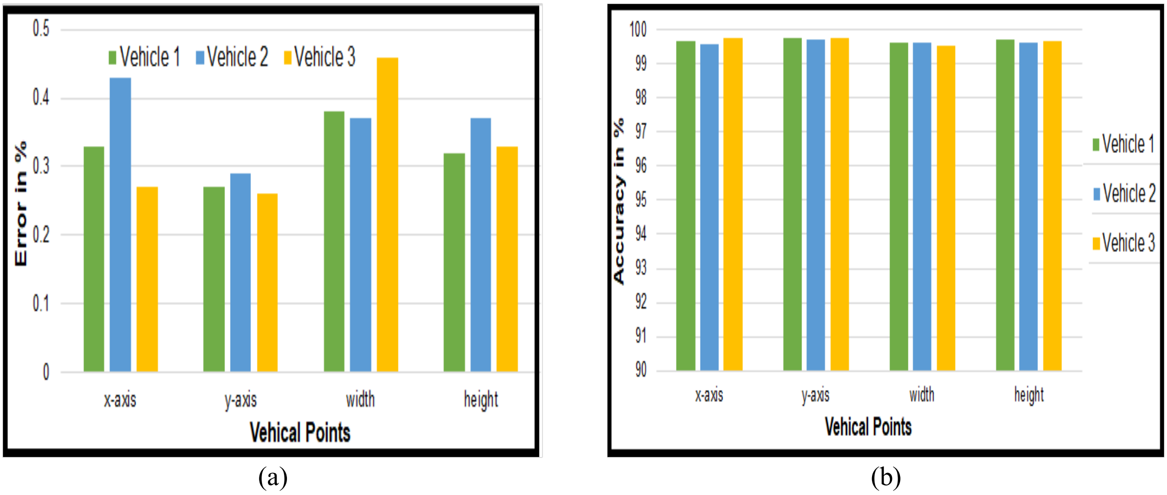


FIGURE 8. (a) Average error rate and (b) accuracy to detect each vehicle points on the Frames [sherbrooke- urban tracker dataset].

TABLE 5. MOT metric evaluation.

Dataset	Video	Number of Frames Extracted	MOTA			MOTP		
			Blob Detection	Kalman Filter [29]	PredictV	Blob Detection	Kalman Filter [29]	PredictV
Urban Tracker	Sherbrooke	2000	0.4705	0.6868	0.7504	37.27	15.42	10.87
Ko-PER	Sequence1a	1210	0.5930	0.7155	0.7719	30.13	18.36	13.91
	Sequence1b	1213	0.5006	0.7187	0.7738	36.71	21.17	18.29
	Sequence1c	1209	0.5444	0.7414	0.7920	31.34	19.28	14.82
	Sequence1d	1205	0.5186	0.6969	0.7465	29.60	16.72	13.37

time t.

$$MOTP = 1 - \frac{\sum_{i,t} (ms_t + fp_t + mme_t)}{\sum_t o_t}$$

Here, mst, fpt, mmet and ot are the number of misses, false positives, mismatches, and ground truth, respectively, at time t. The results of the proposed PredictV compare to some traditional methods is given in the following table.

The above table demonstrate the results of MOTA and MOTP. These results ensures the capabilities of the proposed approach. The proposed approach achieved highest MOTA value and better MOTP values for both datasets.

V. CONCLUSION AND FUTURE SCOPE

The goal of this proposed PredictV approach is to predict the vehicles on the frame to avoid occlusion. For proposing this scheme, the existing blob detection mechanism is utilized. All points of the vehicles on the frames are predicted based on the previous vehicle and its position. It means the percentage change is calculated for the same and used for prediction. Some problems like the change in the number of vehicles from frame to frame also arise in this work, which is then solved by extending this proposed scheme. This proposed scheme's performance is calculated based on various parameters and metrics, and it has been noticed that the proposed scheme reveals its virtue. The parameters like the accuracy of entry and exit of the vehicle and accuracy to resolve occlusion are adequate. Also, as per the calculation, the proposed scheme predicted 78.66% of vehicles from all 7152 frames of Urban Tracker and Ko-PER datasets accurately, whereas blob detection detects only 58% of vehicles. On the other hand, Kalman filter-based approach is also implemented on the same dataset, and it has been found that it detects 73% of vehicles, which is again 6% lesser than the proposed approach. This proposed scheme's other objective is to avoid occlusion, but a small percentage of occlusion is still found. The calculations return that 17% of frames are occluded in the proposed method, whereas the rate is higher than43%in blob detection and 24% in Kalman Filter-based approach. Some unwanted objects are also found on the frame, but it does not affect the proposed scheme's performance much. The proposed scheme's effectiveness has also been evaluated based on the predicted points' error rate and accuracy. For this, the frames' original points have been compared with predicted points for every 10th frame of the Sherbrooke video of the Urban Tracker dataset. Overall, 99% accuracy has been achieved with the predicted points, proving the proposed scheme's effectiveness. The quality of the proposed approach measured in terms of MOTP and MOTA, also define its effectiveness. In the future, the work can be enhanced to reduce unwanted object detection. The work can also be extended using real-time datasets to evaluate the proposed scheme's accuracy.

COMPLIANCE WITH ETHICAL STANDARDS

Ethical approval: The authors have considered publically available datasets for performing the experiments in the considered work. The source of these datasets are:

1. https://www.jpjodoin.com/urbantracker/dataset.html
2. https://www.researchgate.net/publication/287042907_The_KoPER_intersection_laserscanner_and_video_dataset

Informed Consent: All the authors are agreed for this submission.

***Shikha Tuteja*:** Writing - original draft, Conceptualization, Methodology, Implementation.

***Shashi Poddar*:** Experimental Work, Testing, Writing - review & editing.

***Divya Agrawal*:** Writing, Review & editing.

***Vinod Karar*:** Testing, Review & editing final draft.

REFERENCES

[1] C. Y. Wang, H. Liu, H. Q. Zhang, M. Hu, and X. Li, "Anti-occluding tracking algorithm based on grey prediction and mean-shift," *Control Eng. China*, vol. 24, 2017.

[2] Y. Jia, W. Min-yuan, C. Shu-zhen, and Z. Qing-lin, "Anti-occlusion tracking algorithm based on mean shift and fragments," *Opt. Precis. Eng.*, vol. 18, pp. 1413–1419, Jan. 2010.

[3] G. Liang, W. Niu, B. Li, Q. Li, and P. Zhao, "Structural sparse tracking method based on importance weighting," *Chin. J. Sens. Actuators*, vol. 31, pp. 61–67, Jul. 2018.

[4] H. Kawano, H. Orii, K. Shiraishi, and H. Maeda, "A SIFT feature-based template matching method for detecting and counting objects in life space," *Appl. Mech. Mater.*, vol. 36, pp. 413–421, Oct. 2010, doi: 10.4028/www.scientific.net/AMM.36.413.

[5] B.-F. Wu, C.-C. Kao, C.-L. Jen, Y.-F. Li, Y.-H. Chen, and J.-H. Juang, "A relative-discriminative-histogram-of-oriented-gradients-based particle filter approach to vehicle occlusion handling and tracking," *IEEE Trans. Ind. Electron.*, vol. 61, no. 8, pp. 4228–4237, Aug. 2014, doi: 10.1109/TIE.2013.2284131.

[6] J.-F. Chen, C.-C. Wang, and C.-F. Chou, "Multiple target tracking in occlusion area with interacting object models in urban environments," *Robot. Auto. Syst.*, vol. 103, pp. 68–82, May 2018, doi: 10.1016/j.robot.2018.02.004.

[7] H. N. Phan, L. H. Pham, D. N.-N. Tran, and S. V.-U. Ha, "Occlusion vehicle detection algorithm in crowded scene for traffic surveillance system," in *Proc. Int. Conf. Syst. Sci. Eng. (ICSSE)*, Jul. 2017, pp. 215–220, doi: 10.1109/ICSSE.2017.8030868.

[8] P. Valiere, L. Khoudour, A. Crouzil, and D. N. Truong Cong, "Robust vehicle counting with severe shadows and occlusions," in *Proc. Int. Conf. Image Process. Theory, Tools Appl. (IPTA)*, Nov. 2015, pp. 99–104, doi: 10.1109/IPTA.2015.7367105.

[9] J. L. Chang, L. F. Wang, G. F. Meng, S. Xiang, and C. Pan, "Vision-based occlusion handling and vehicle classification for traffic surveillance systems," *IEEE Intell. Transp. Syst. Mag.*, vol. 10, no. 2, pp. 80–92, Feb. 2018, doi: 10.1109/MITS.2018.2806619.

[10] C. Wang, J. Delport, and Y. Wang, "Lateral motion prediction of on-road preceding vehicles: A data-driven approach," *Sensors*, vol. 19, no. 9, p. 2111, May 2019, doi: 10.3390/s19092111.

[11] J. Nilsson, M. Brännström, E. Coelingh, and J. Fredriksson, "Lane change maneuvers for automated vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 5, pp. 1087–1096, May 2017, doi: 10.1109/TITS.2016.2597966.

[12] W. Hanwool, Y. Ji, H. Kono, Y. Tamura, Y. Kuroda, T. Sugano, Y. Yamamoto, A. Yamashita, and H. Asama, "Lane-change detection based on vehicle-trajectory prediction," *IEEE Robot. Autom. Lett.*, vol. 2, no. 2, pp. 1109–1116, Apr. 2017, doi: 10.1109/LRA.2017.2660543.

[13] D. Isele, R. Rahimi, A. Cosgun, K. Subramanian, and K. Fujimura, "Navigating occluded intersections with autonomous vehicles using deep reinforcement learning," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 2034–2039, doi: 10.1109/ICRA.2018.8461233.

[14] M. Lee, K. Jo, and M. Sunwoo, "Collision risk assessment for possible collision vehicle in occluded area based on precise map," in *Proc. IEEE 20th Int. Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2017, pp. 1–6, doi: 10.1109/ITSC.2017.8317666.

[15] S. Hoermann, M. Bach, and K. Dietmayer, "Dynamic occupancy grid prediction for urban autonomous driving: A deep learning approach with fully automatic labeling," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 2056–2063, doi: 10.1109/ICRA.2018.8460874.

[16] X. Fu, W. Luo, C. Xu, and X. Zhao, "Short-term traffic speed prediction method for urban road sections based on wavelet transform and gated recurrent unit," *Math. Problems Eng.*, vol. 2020, May 2020, Art. no. 3697625, doi: 10.1155/2020/3697625.

[17] Z. Li, Y. Chen, and Z. Yin, "Vehicle tracking fusing the prior information of Kalman filter under occlusion conditions," *SN Appl. Sci.*, vol. 1, no. 8, p. 822, Aug. 2019, doi: 10.1007/s42452-019-0852-2.

[18] E. Krieger, P. Sidike, T. Aspiras, and V. K. Asari, "Vehicle tracking under occlusion conditions using directional ringlet intensity feature transform," in *Proc. Nat. Aerosp. Electron. Conf. (NAECON)*, Jun. 2015, pp. 70–74, doi: 10.1109/NAECON.2015.7443041.

[19] L.-Y. Ong, S.-H. Lau, and V.-C. Koo, "Performance of invariant feature descriptors with adaptive prediction in occlusion handling," in *Proc. 3rd Int. Conf. Control, Autom. Robot. (ICCAR)*, Apr. 2017, pp. 385–388, doi: 10.1109/ICCAR.2017.7942723.

[20] X. Qin, S. Dai, and F. Li, "Target tracking algorithm based on improved kernel correlation filter," *Meas Control Technol.*, vol. 36, pp. 3–6, Apr. 2017.

[21] B. Xiaoan, Z. Xiujuan, W. Qiang, H. Lingling, and G. Jiangsheng, "Anti occlusion target tracking algorithm based on KCF and SIFT feature," *Comput. Meas. Control.*, vol. 26, pp. 148–152, Jul. 2018.

[22] M. Qing, V.-D. Hoang, and K.-H. Jo, "Localization and tracking of same color vehicle under occlusion problem," in *Proc. 9th France-Jpn. 7th Europe-Asia Congr. Mechatronics (MECATRONICS)/13th Int. Workshop Res. Educ. Mechatronics (REM)*, Nov. 2012, pp. 245–250, doi: 10.1109/MECATRONICS.2012.6451016.

[23] N. Zhao, Y. Xia, C. Xu, X. Shi, and Y. Liu, "APPOS: An adaptive partial occlusion segmentation method for multiple vehicles tracking," *J. Vis. Commun. Image Represent.*, vol. 37, pp. 25–31, May 2016, doi: 10.1016/j.jvcir.2015.04.011.

[24] S.-G. Shin, D.-R. Ahn, and H.-K. Lee, "Occlusion handling and track management method of high-level sensor fusion for robust pedestrian tracking," in *Proc. IEEE Int. Conf. Multisensor Fusion Integr. Intell. Syst. (MFI)*, Nov. 2017, pp. 233–238, doi: 10.1109/MFI.2017.8170434.

[25] M. H. O. Rashid, M. A. Mamun, M. A. Hossain, and M. P. Uddin, "Brain tumor detection using anisotropic filtering, SVM classifier and morphological operation from MR images," in *Proc. Int. Conf. Comput., Commun., Chem., Mater. Electron. Eng. (IC4ME2)*, Feb. 2018, pp. 1–4, doi: 10.1109/IC4ME2.2018.8465613.

[26] P. L. St-Charles, G. A. Bilodeau, and R. Bergevin, "SuBSENSE: A universal change detection method with local adaptive sensitivity," *IEEE Trans. Image Process.*, vol. 24, no. 1, pp. 359–373, Jan. 2015, doi: 10.1109/TIP.2014.2378053.

[27] E. Strigel, D. Meissner, F. Seeliger, B. Wilking, and K. Dietmayer, "The ko-PER intersection laserscanner and video dataset," in *Proc. 17th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2014, pp. 1900–1901, doi: 10.1109/ITSC.2014.6957976.

[28] J. Jodoin, G. Bilodeau, and N. Saunier, "Urban tracker: Multiple object tracking in urban mixed traffic," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, Mar. 2014, pp. 885–892.

SHIKHA TUTEJA received the M.Tech. degree in electronics and communication engineering from the Guru Nanak Dev Engineering College, Ludhiana, India, in 2014. She is currently pursuing the Ph.D. degree with AcSIR—Central Scientific Instruments Organisation (CSIO), Chandigarh, where she is working on occlusion and feature-based traffic handling. She is also working as a Senior Project Fellow with CSIO. In CSIR—CSIO, she is presently dealing with a project based on fast track translation (FTT), Project Visual Landing for Naval Operations. She has also worked on CSIR—CSIO on a Defense Project ''MLP-0046-design and development of enhanced vision system for military surface transport vehicles'' from September 2018 to March 2020. As her past experience, she has worked as an Assistant Professor with the Rayat and Bahra Institute of Engineering and Biotechnology, from 2007 to 2016. From 2016 to 2017, she worked with the Chandigarh Group of Colleges, Landran. She has filed five patents in various fields. She has published two SCI articles and five has been under communicated. She has published one book chapter with Taylor and Francis (CRC Press). She has published many national and international articles in different journals. Her research interests include occlusion handling, feature descriptors, computer vision, machine learning, and deep learning. She received the Best Paper Award in the Jammia Teachers Association Multidisciplinary International Conference on JTACON in 2020. She is also an associate editor of SCI-indexed journals.

SHASHI PODDAR received the B.Tech. degree from SRM University, Chennai, in 2010, and the M.Tech. and Ph.D. degrees from the Academy of Scientific and Innovative Research, Chandigarh, India, in 2012 and 2017, respectively. He is currently enrolled as a Scientist with the Central Scientific Instruments Organisation (CSIR). His research interests include optimal state estimation, computer vision, and system development.

DIVYA AGRAWAL received the B.Tech. degree in biomedical engineering from the National Institute of Technology, Raipur, and the M.Tech. degree in advanced instrumentation engineering and the Ph.D. degree in engineering from AcSIR—CSIO, Chandigarh. She is currently working as a Principal Scientist with the Department of Imaging Avionics and Display Systems, CSIR—CSIO. She is also engaged as an Honorary Assistant Professor with AcSIR—CSIO. Her research interests include avionics instrumentation, opto-electronics, imaging, and situation awareness. She is an Associate Member of IETE and a member of IAENG. She was awarded the Gold Medal for her B.Tech. degree. She was conferred with the prestigious CSIR Young Scientist Award (Engineering) in 2019. She also received the CSIR's Quick Hire Scientist Trainee Scholarship for duration of two years 2009–2011.

VINOD KARAR received the B.E. degree in electronics from NIT Bhopal, India, and the M.E. degree in electronics from the Punjab Engineering College Chandigarh, India, and the Ph.D. degree in electrical and instrumentation engineering from Thapar University, Patiala, India.

He is currently working as the Chief Scientist with the Central Scientific Instruments Organisation (CSIR), and a Professor and a Coordinator with AcSIR—CSIO, Chandigarh. He has to his credit more than 145 publications in peer-reviewed journals and conference proceedings, more than 75 conference presentations as the author/coauthor, eight book chapters, over 150 technical reports and documents, and two patents. He has successfully completed eight major projects involving multiple agencies while eight major projects are in advanced stage of execution. He also has to his credit more than 15 technologies commercialized or developed till TRL-7 or more. He has guided more than 30 postgraduate students and more than 100 undergraduate students, eight students have completed Ph.D. course under him while six Ph.D. students are in progress. He has deep understanding of airworthy design and certifications for avionic equipment and systems. His research interests include electronics and instrumentation, opto-electronics, optical fabrication techniques, sensors, avionics and cockpit instrumentation, image processing, machine learning and artificial intelligence, ergonomics, human aspects, and situation awareness.

Dr. Karar was a recipient of several awards including, the CSIR Technology Award 2020: CSIR Most Significant Technology of Plan Period for Head up display and associated test equipment for multiple aircraft platforms, the Defence Gold Skoch Award 2020: NVG Compatible Helicopter Visual Landing Aid Solutions for Indian Naval Ships Enabling Night Operations during Critical Missions, the CSIR Technology Award Certificate of Merit in 2018, the Gold Skoch Transformational Innovation Award in 2017 for Aviation Cockpit Display Validation Platform, the SIATI Award of Excellence for Indigenization of Aerospace Technology: Head Up Display for LCA-Tejas in 2011, the 34th IETE Hariramji Toshniwal Award for Outstanding Contribution in Development of Head Up Display Technology for Light Combat Aircraft in 2011, the Award of Merit in Leadership Development Program by CSIR-HRDC in 2009, and 12 best paper awards as the author/coauthor.