SciencePG
Science Publishing Group

Vehicle Detection and Tracking Based on GMM and Enhanced Camshift Algorithm

Kaiyang Zhong[*], **Zhaoyang Zhang, Zhengyu Zhao**

Department of Software Engineering, Xiamen University, Xiamen, China

Email address:
972887209@qq.com (Kaiyang Zhong), 296753977@qq.com (Zhaoyang Zhang), 24320152202869@stu.xmu.edu.cn (Zhengyu Zhao)
[*]Corresponding author

To cite this article:
Kaiyang Zhong, Zhaoyang Zhang, Zhengyu Zhao. Vehicle Detection and Tracking Based on GMM and Enhanced Camshift Algorithm. *Journal of Electrical and Electronic Engineering*. Vol. 6, No. 2, 2018, pp. 40-45. doi: 10.11648/j.jeee.20180602.11

Abstract: Vehicle detection and tracking is an important part of the intelligent transportation system. With the rapid development of computer vision, video based vehicle detection and tracking technology has become a hot topic. In this paper, on the foundation of the present work, an enhanced detection tracking algorithm is proposed based on the popular Gauss mixture model(GMM) and Camshift. First, GMM is used to extract the foreground, and then the morphological operations is carried out to enhance the image, so that to remove the random noises. Finally, enhanced Camshift is designed to track the vehicle which is discussed in detail below. The experimental results demonstrate that the tracking accuracy can be improved.

Keywords: Vehicle Detection, Vehicle Tracking, GMM, Camshift

1. Introduction

Nowadays, with the sharply increasing of city transportation burden, it is urgent to build a real-time and efficient ITS (Intelligent transportation system)[1] which is applied to traffic surveillance, traffic control and road traffic information collection. As the primary level of ITS, the detection and tracking of on-road vehicles timely and effectively become an important task, because it determines the performance of the following upper-level task-vehicle behaviour analysis [2]. The integrated information of the vehicle features gathered in the lower level are imported into the highest level to model and predict the typical on-road behavior and goals of other vehicles on the road [3]. Based on this, many researchers devote to the research on this hot topic. For decades of years, many algorithms and frameworks are proposed [4-11]. Due to the development of computer vision technology and the popularity of the camera, the vision-based detection and tracking have became the most popular research branch. Regard to the detection, there are two technical frameworks: appearance-feature and motion-based. For the former, the appearance, such as the size, shape and color etc is extracted by some typical descriptors including HOG, SURF, gabor, Harris and optical flow etc [4-8]. Considering the dynamic of vehicle, motion-based algorithms are given in [9-11]. During the tracking stage, the verified vehicles are tracked in subsequent images by various tracking filters, such as the typical Kalman Filter [12], particle filter [13] and mean-shift [14]. In recent years, deep learning [15-17] is also applied to vehicle detection. But the process of training the model takes too long and has high requirements for hardware. For balancing the performance and cost, it comes to Camshift. Based on mean-shift filter, Camshift [18] is proposed to deal with the target deformation and occlusion. For better tracking accuracy in the complex environment, the enhanced Camshift method is proposed because of its good performance in solving the adhesion vehicle tracking.

The rest of the paper is organized as follows: Section 2 describes the vehicle detection and tracking methods in detail. Section 3 shows the experimental results of the proposed method and conclusions of the work is given in Section 4.

2. Proposed Method

Concentrated on designing a valid vehicle detection and tracking algorithm, in the paper, the extra noises are removed by morphology operator after the background suppression. Additional, the GMM and enhanced Camshift method are

discussed ~~to~~ enhance the accuracy and efficiency of vehicle tracking. The flowchart of the proposed method is given in Figure 1. According to the flowchart, the work is carried out by five steps which is inputting image sequence, background modeling by GMM, noise removal by morphology, blob analysis and tracking based on cameshift one by one. In this section, a detailed description of the algorithm is given in the following.
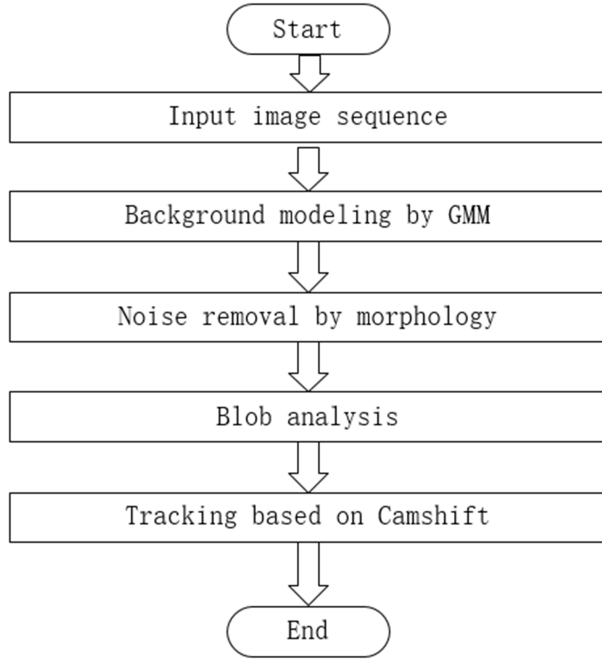
```
        ┌─────────────┐
        │    Start     │
        └─────────────┘
               ⇩
   ┌───────────────────────────┐
   │    Input image sequence    │
   └───────────────────────────┘
               ⇩
   ┌───────────────────────────┐
   │ Background modeling by GMM  │
   └───────────────────────────┘
               ⇩
   ┌───────────────────────────┐
   │ Noise removal by morphology │
   └───────────────────────────┘
               ⇩
   ┌───────────────────────────┐
   │       Blob analysis         │
   └───────────────────────────┘
               ⇩
   ┌───────────────────────────┐
   │  Tracking based on Camshift │
   └───────────────────────────┘
               ⇩
        ┌─────────────┐
        │     End      │
        └─────────────┘
```

Figure 1. *The flowchart of proposed method.*

Step 1. Input image sequence

The image sequences are collected by the camera setting up on the overbridge and decomposed into frames. The frames can vary from 1 to n so they can be accepted by the system one by one. The above step can be expressed as follows:

$$V = \{F_i\}, i = 1,...n \qquad (1)$$

Where, V is the frames collection and n represents the total number of frames in the sequence.

Step 2. Background modeling by GMM

GMM is a density model composed of several components of Gauss function. It models each pixel as a mixture of two or more Gaussians temporally with online updated. These distributions are estimated as either a stable background process or short-term foreground process by evaluating its stability. If the pixel distribution is stable above threshold, then it is classified as background pixel. This method is suitable for background extraction process because of its good performance in multiple target detection and changes of light. Gauss mixture model is widely used in background modeling. In the proposed method, 40 frames of the video without cars are selected to model the Gauss background.
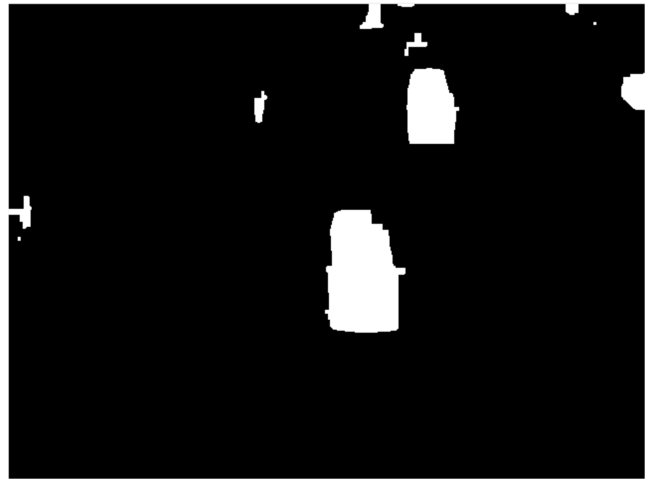
Step 3. Noise removal by morphology

After step2, the background is achieved by GMM and then the foreground is segmented by subtracting the background from the original image. However, there still exist some stubborn noises which is caused by changing light and other disturbances. In order to remove the unwanted noise as much as possible, the morphological opening and closing operators are utilized to suppress the noise, for it has good performance in eliminating the noise with sharp edges caused by traffic lights, sign posts etc. Further it can fill the holes due to the subtraction operation and forming multiple single-connected components. As shown in Figure 2 (a) (b), it is clear that after the morphological operator the foreground is composed of several connected regions without holes.
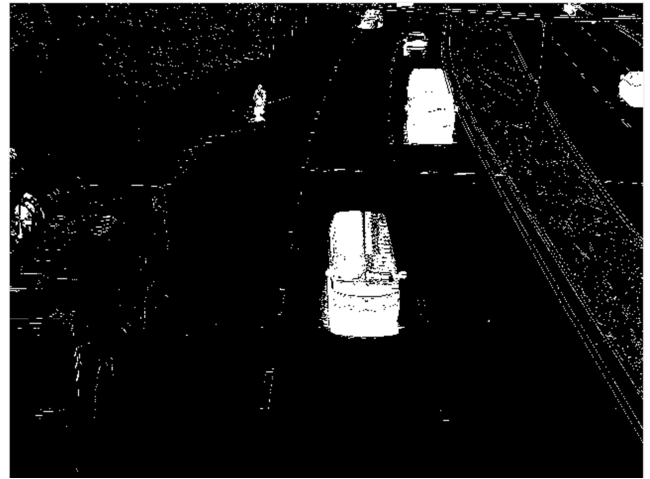
(a) The image with random noise

(b) Result of noise removal by morphological operator

Figure 2. *Result of noise reduction0*

Step 4. Blob analysis

After the noise removal, the foreground objects is achieved by extracting the connected component. In this section, the Blob analysis method [19] is used to detect cars. Blob analysis is used to analysis the connected region of the same pixel in the image. In the experiment, Every connected component is treated as a vehicle marked by a bounding box.

The camera is fixed on the bridge so the area of vehicles have a certain range in the image. To extract the candidate

vehicle, the connected component whose size doesn't meet with the range limitation will be removed. Large amounts of data are collected for the Gauss distribution fitting to get the satisfied threshold of vehicles' area in the image. Finally 400 is chosen as the minimum area of the connected component. Only the satisfied connected component are assigned a bounding box tag. Figure 3 shows the result of blob analysis.
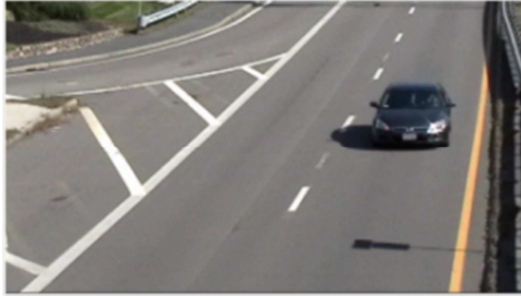
(a) Original image

(b) Image after blob analysis

***Figure 3.** Result of blob analysis.*

Step 5. Tracking based on Camshift
a. Camshift algorithm

Camshift is an improved version of the mean shift algorithm for video object in one dimensional histogram which is applied to find the peak of the probability density histogram backprojection. Its basic idea is that all frames of the video image are performed MeanShift operation, and the operation result of the previous frame (the center and size of the Search Window) is taken as the initial value of the Search Window in the next frame with the MeanShift algorithm.

Camshift has been widely used in face recognition. Due to its excellent performance in the face recognition area, it is creatively applied to road vehicle detection in this paper. The principle of Camshift can be summed up as the following steps:

Step 1: Select the ROI which contains the vehicle object.
Step 2: Convert the frame to HSV model.
Step 3: Generate the object model using the ROI image.
Step 4: Use histogram back projection to create probability density image.
Step 5: Use Mean-shift to find the centroid of the object.
Step 6: The new position is treated as the new ROI for the next frame.

b. Enhanced Camshift method

The original Camshift could not be applied to vehicle tracking directly because it's too sensitive to noise and adhesion of the vehicles. So in this section, some constraints on the original Camshift method are introduced to enhance the tracking performance.

From the start frame of one video, each detected vehicle target is identified by an identity(ID) which could uniquely identify a car as shown in Figure 4.
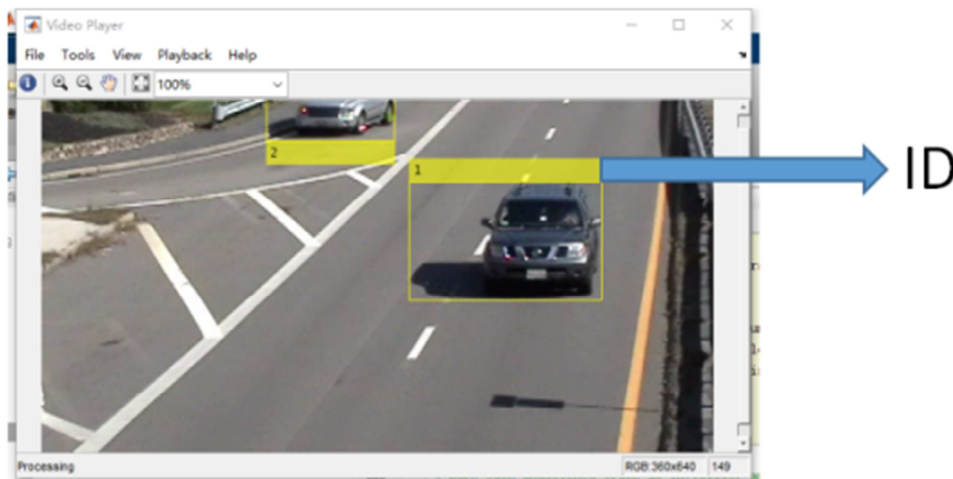

***Figure 4.** Tracking ID.*

The algorithm uses cost matrix to record the position of cars and predicted cars. For cost matrix, the number of column is equal to the number of bounding box in the current frame. The number of row is equal to the number of cars in the previous frame. The cost matrix records the distance between the centroid of each car predicted by last frame and the one in the current frame. It can be expressed as formula(2). I is equal to the number of column. J is equal to the number of row.

$$Cost(i,j) = [(track(i).centroid.x - \det ection(j).centroid.x)^2 + (track(i).centroid.y - \det ection(j).centroid.y)^2]^{\frac{1}{2}} \quad (2)$$

After generating the cost matrix, it comes to the enhanced

part. If the distance in the cost matrix is less than 20, it's considered that the car appears again in the next frame. The location information of this car is updated with the new location. The car in the current frame is called to be matched. If the target in the current frame doesn't have a car to match, then its location would be stored and it would be treated as a new car which just appears. If the car in the last frame doesn't have a target to match, without considering it having disappeared at once, the system watches for 20 frames and uses the predicted location to update its location. If still no car could match it in 20 frames, it's finally considered that the car disappeared. But if a car can't have a match when it appears within the time of 8 frames, it's just considered that the object was noise and just let it disappear. All of the parameters are obtained by doing a lot of experiments and comparing their performance. The performance of the enhanced Camshift method will be discussed in detail in the following section.

3. Experimental Result

The performance of the proposed method is compared with the camshift algorithm without the enhanced details above. The experiments have been performed in MATLAB version 2016a on an Intel(R) Core(TM)i7-4790CPU@3.60GHz machine with 4 GB RAM. A digital camera is set up on the bridge whose height is same with the traffic light in China. The videos which were used to test can be broadly categorized into three types. (I) Video of road with only cars, (II) Video of road with different kinds of vehicles like bicycle and bus which is not the detection target. (III) Video of road

with vehicle adhesion. The performance of the proposed method is discussed in the three scenes in detail.

3.1. Performance Testing in Various Scenarios

Scene 1:
As the table 1 shows, the input video contains 3 cars which is moving under the camera. Firstly, the GMM is used to detect the foreground. Then the morphology processing is applied to reduce the noise. After that, it's visually clear that the foreground generated by the proposed method does not suffer with random noise. Finally the enhanced camshift algorithm is applied to track each car.
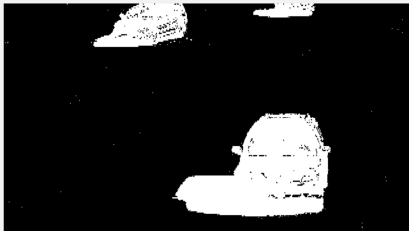
Scene 2:
The processing steps are mainly as same as the scene 1. The different point is that a maximal threshold of the vehicle area is added to filter the big target like bus which is not the detection target. Also the bicycle and motor car are filtered by the minimum threshold.

Scene 3:
The processing steps are still mainly as same as the scene 1, but the road environment is more complex. As shown in the Table 1 Scene 3, on the right lane two vehicles seem to stick together with a street light blocking part of one car behind. In this extreme conditions, the car is having nearly disappeared, so the system gives a mark of 'predicted' to this car whose location information is predicted by the system. This method has been proved performing well under the complex rode environments like vehicle adhesion and occlusion. The Table 1 shows the results of the experiments in the 3 scenes.

***Table 1.** The results of experiment.*

	Scene 1	Scene 2	Scene 3
Original images			
Blob analysis			
Noise reduction			

	Scene 1	Scene 2	Scene 3
Tracking process			
Results			
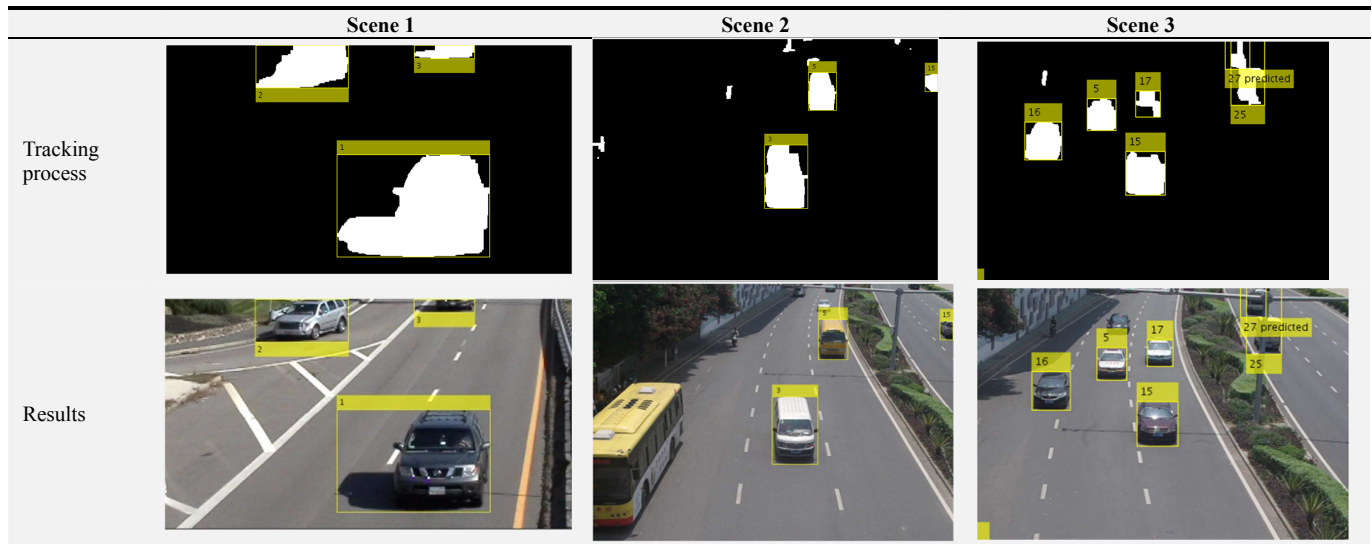
It's clearly to see from the results of the experiments that it's easy to catch the outline of the car after Blob analysis and noise reduction when there are only cars on the road. So the object is detected and tracked without much problem. In the scene 2, there is a big bus in the picture which is out of the range of the detection target. So during the process of Blob analysis, the bus is filtered because its area is out of the prescribed scope. In the scene 3, the road environment is more complex with the vehicle adhesion and the blob analysis can not give a clear divide of different cars. Because of the good performance of enhanced Camshift, each car could be tracked even when vehicle adhesion happens.

3.2. Performance Comparision with Different Methods

After giving the qualitative results of the experiments above, the quantitative analysis of proposed method along with the original camshift algorithm are presented. In order to compare the evaluation of two methods well, a numerical accuracy is computed by using the parameters. The parameters are defined as follows[20]:

True Positive (TP): the number of correctly extracted true vehicle;

False Positive (FP): the number of incorrectly extracted false vehicles;

False Negative (FN): the number of omitted vehicles. Based on these parameters, the measures are computed as follows:

$$Correctness = \left(\frac{TP}{TP+FP} \right) \times 100 \qquad (3)$$

$$Completeness = \left(\frac{TP}{TP+FN} \right) \times 100 \qquad (4)$$

$$Quality = \left(\frac{TP}{TP+FP+FN} \right) \times 100 \qquad (5)$$

Table 2 and *Table* 3 show the correctness(%), completeness(%) and overall quality(%) of the proposed method and the original Camshift method.

***Table 2.** Efficiency of proposed method.*

	TP	FP	FN	Correctness (%)	Completeness (%)	Quality (%)
Scene 1	4	0	0	100	100	100
Scene 2	3	0	1	100	75	75
Scene 3	7	2	1	77.78	87.5	70

***Table 3.** Efficiency of original camshift method.*

	TP	FP	FN	Correctness (%)	Completeness (%)	Quality (%)
Scene 1	4	1	1	80	80	66.67
Scene 2	7	2	3	77.78	70	58.33
Scene 3	5	3	2	62.5	71.43	50

From the tables above, it's easy to see that enhanced Camshift method performs better than original method in different kinds of scenes especially the complex road environment.

4. Conclusion

In this article, an enhanced vehicle detection and tracking method is proposed. The GMM and enhanced Camshift provide a better solution in solving the problem of vehicle detection and tracking in complex road environment. The proposed method have addressed the complex situations in the experiment and generated better quality of results. The experimental result shows that the proposed method has certain advantages in the city road environment for vehicle

detection and good robustness in the complex road environment. The experimental result contributes to intelligent transportation system. In the future, we will continue to optimize the algorithm to realize the detection and tracking of vehicles in more complex road environment.

Acknowledgements

References

[1] "Intelligent Transportation Systems Joint Program Office", United States Department of Transportation, vol. 10, November 2016.

[2] S. Sivaraman, and M. M. Trivedi, "Looking at vehicles on the road: A survey of vision-based vehicle detection, tracking, and behavior analysis," IEEE Transactions on Intelligent Transportation Systems, vol. 14, no. 4, pp. 1773-1795, 2013.

[3] X. Wang, "Intelligent multicamera video surveillance: A review," Pattern recognition letters, vol. 34, no. 1, pp. 3- 19, 2013.

[4] M. A. Manzoor, Y. Morgan, "Vehicle Make and Model Classification System using Bag of SIFT Features", 7th IEEE Annual Conference on Computing and Communication Workshop and Conference (CCWC), vol. 02, pp. 572-577, March 2017.

[5] S. M. Elkerdawi, R. Sayed, and M. ElHelw, "Real-time vehicle detection and tracking using Haar-like features and compressive tracking," in 1st Iberian Robotics Conference, Jan. 2014, pp. 381-390. Springer International Publishing.

[6] Honghong Yang, Shiru Qu, "Real-time vehicle detection and counting in complex traffic scenes using background subtraction model with low-rank decomposition", IET Intelligent Transport Systems, vol. 12, pp. 75-85, January 2018.

[7] M. Atibi, I. Atouf, M. Boussaa, A. Bennis, "Real-time detection of vehicle using the haar-like features and artificial neuron networks", Proc. Computer Science, vol. 73, pp. 24-31, 2015.

[8] Gao Lei, "Based on the optical flow in the dynamic scene of the vehicle detection and tracking algorithm [D]", University of Science and Technology of China, 2014.

[9] Junpeng Zhang, Xiuping Jia, Jiankun Hu, "Motion Flow Clustering for Moving Vehicle Detection from Satellite High Definition Video", 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA), 29 Nov.-1 Dec. 2017.

[10] N. Dalal, and B. Triggs, "Histograms of oriented gradients for human detection,"in proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, Jun. 2005, pp. 886-893.

[11] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in Computer vision (ECCV), Jan. 2006, pp. 404-417. Springer Berlin Heidelberg.

[12] Y. Du, and F. Yuan, "Real-time vehicle tracking by Kalman filtering and Gabor decomposition," in 1st International Conference on Information Science and Engineering (ICISE), Dec. 2009, pp. 1386-1390.

[13] H. T. Niknejad, A. Takeuchi, S. Mita, and D. McAllester, "On-road multivehicle tracking using deformable object model and particle filter with improved likelihood estimation," IEEE Transactions on Intelligent Transportation Systems, vol. 13, no. 2, pp. 748-758, 2012.

[14] L. Wei, X. Xudong, W. Jianhua, Z. Yi, and H. Jianming, "A SIFT-based mean shift algorithm for moving vehicle tracking," in Proc. IEEE Intelligent Vehicles Symposium, Jun. 2014, pp. 762-767.

[15] Kaijing Shi, Hong Bao, Nan Ma,"Forward Vehicle Detection Based on Incremental Learning and Fast R-CNN" in 2017 13th International Conference on Computational Intelligence and Security (CIS).

[16] Alireza Asvadi, Luis Garrote, Cristiano Premebida, Paulo Peixoto, Urbano J. Nunes,"DepthCN: Vehicle detection using 3D-LIDAR and ConvNet" in 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), 16-19 Oct. 2017.

[17] Flaviu Ionut Vancea, Arthur Daniel Costea, Sergiu Nedevschi, "Vehicle taillight detection and tracking using deep learning and thresholding for candidate generation", 2017 13th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP), 7-9 Sept. 2017.

[18] Z. Wang, and K. Hong, "A new method for robust object tracking system based on scale invariant feature transform and camshift," in Proc. 2012 ACM Research in Applied Computation Symposium, Oct. 2012, pp. 132-136.

[19] Rhen Anjerome Bedruz, Edwin Sybingco, Argel Bandala, Ana Riza Quiros, Aaron Christian Uy, Elmer Dadios, "Real-time vehicle detection and tracking using a mean-shift based blob analysis and tracking approach", 2017IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), 1-3 Dec. 2017.

[20] K. V. Arya, Shailendra Tiwari, Saurabh Behwal, "Real-time Vehicle Detection and Tracking", Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2016 13th International Conference on.