Wireless Personal Communications (2020) 114:2121–2133
https://doi.org/10.1007/s11277-020-07468-y

A Comprehensive Survey on Autonomous Driving Cars: A Perspective View

S. Devi[1] · P. Malarvezhi[1] · R. Dayana[1] · K. Vadivukkarasi[1]

Published online: 15 May 2020

Abstract
Over the past decades Machine Learning and Deep Learning algorithm played a vital part in the development of Autonomous Vehicle. It is indeed for the perception system to examine the environment around the vehicle and identify the objects such as pedestrian, vehicle and traffic signals, etc. Using this information, control system module can take necessary action to control the vehicle in terms of braking, speed, lane change or steering, etc. This paper focuses on the survey of machine learning algorithms and techniques applied in the design of autonomous driving system over a decade. Performance of each algorithm was analyzed in terms of prediction time and accuracy have been documented and compared.

Keywords Machine learning · Deep learning · Autonomous driving · Object detection · Navigation

1 Introduction

In the recent years, Automobile Industries compete with each other to launch the first fully autonomous vehicle. In future we will see lot of self-driving cars around the world. Many companies like GM, Ford, Toyota, Tesla, etc. are taking test drives in recent years.

Many Automobile industries have invested for developing Autonomous Vehicle. GM paid out $581 million to obtain cruise automation in the year 2016. It was testing its vehicles in various cities across California, Arizona and Michigan. In 2017 Ford spent $1 billion into AI-start up Argo AI [1]. In 2019 Ford's Argo AI had put $15 million for forming an autonomous vehicle research Centre. In [2] in 2015 Toyota invested $1 billion to develop autonomous vehicle. Volvo joint venture with Uber [3] spent $300 million to develop next generation self-driving cars. Hyundai also financed $30 million in self-driving car developer Aurora and invested $1.7 billion targeting for the highway driving in 2020 and urban driving in 2030. BWM with Daimler spent $250 million to work in the development of self-driving cars (BWM iNEXT) [4]. Many other automakers [5, 6] like Tesla, Fiat-Chrysler, Renault Nissan partnered with Waymo, Honda, Waymo LLC and other non-Automotive companies like Amazon, Apple, Baidu,

✉ P. Malarvezhi
malarvip@srmist.edu.in

[1] Department of ECE, SRMIST, Chennai, India

Cisco, Microsoft are working on developing full autonomous vehicle. Figure 1 shows that automobile industries have done lot of investment to develop fully Autonomous Vehicle (AV).

The AV should provide more comfort and safety to the people than human drivers. In a recent report, Business Insider predicted that as many as 10 million self-driving cars would be on the road in early 2020. But this doesn't seem likely. Autonomous vehicle require both the right legal and technological framework.

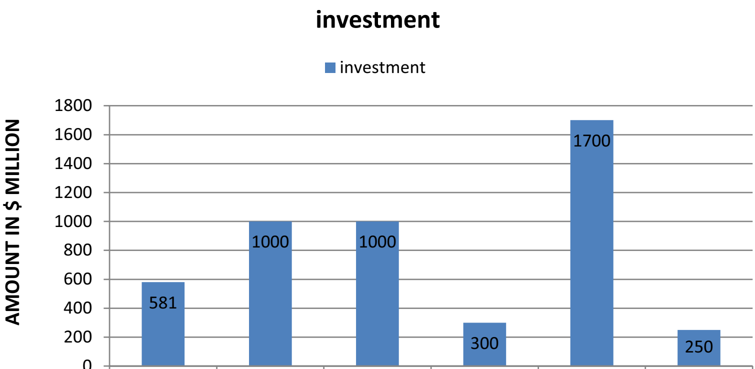
Autonomous vehicles are equipped with a variety of sensors (Lidar, camera-360° & GPS, etc.) to perceive their surroundings. An accurate environment perception is needed to operate safely.

The task of perception system realizes the detection of the surrounding objects that can be captured by the sensors. Depending upon this prediction it gives control to the AV for smooth driving. Therefore visual perception is crucial for autonomous driving.

AV requires the system to be vigorous to various environmental conditions (weather, road conditions, illumination). Here machine learning algorithms are employed for accurate object identification to tackle many challenging tasks. These praiseworthy algorithms should provide smooth mobility for autonomous driving control through low processing time, cost and high prediction accuracy.

Perception system plays an essential role in the autonomous vehicle. Figure 2 shows the representative working flow of Autonomous vehicle.

The rest of the paper is summarized as follows; Sect. 2 defines about generic object detection framework. Section 3 lists the vehicle detection algorithms used in different environmental conditions. Section 3.2 provides information regarding pedestrian detection and trajectory detection. Section 3.4 discuses about some basis frame work needed for autonomous application. In the end, Conclusions are drawn.



	GM	Ford	Toyota	Volvo with Uber	Hyundai	BMW
investment	581	1000	1000	300	1700	250

Fig. 1 Automobile industry investment for fully autonomous vehicle

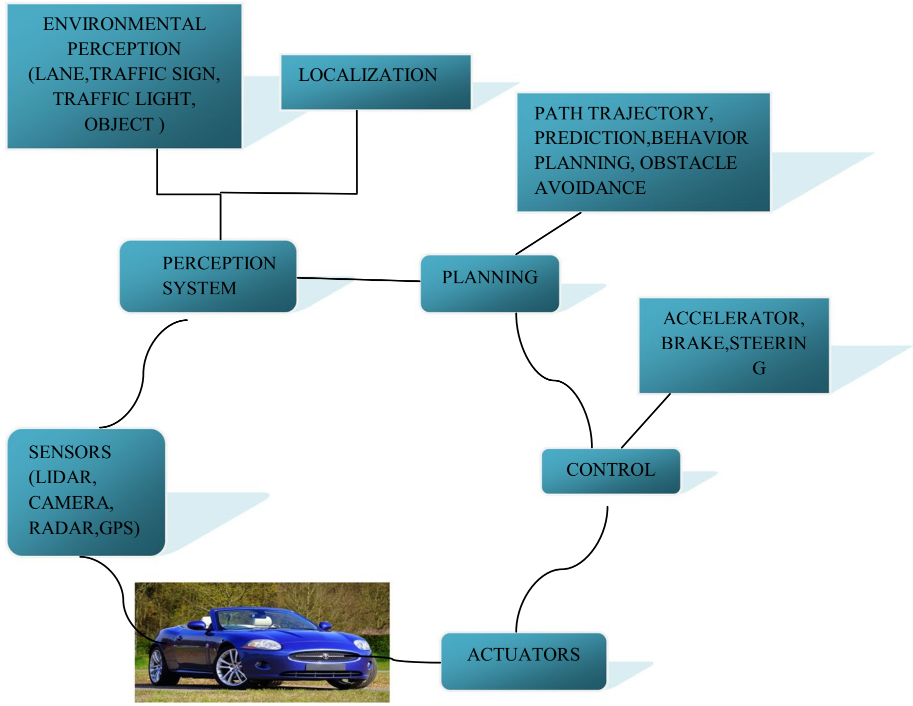


Fig. 2 Autonomous vehicle working flow

2 Objection Detection

Real time object detection is an important research problem for computer vision of the real world. In driving environment, Real time visual object recognition is still very challenging with speed and accuracy to meet real world environment.

Ciberlin et al. [7] proposed an object detection and tracking mechanism which gets necessary information for autonomous vehicle while driving by using front view camera, in which object detectors built by using Viola–Jones algorithm [8] and You Only Look Once (YOLOv3) [9] is used for object recognition. There are Median Flow [10] and correlation tracking methods [11] implemented for tracking object. In this paper Vehicle, Pedestrian, Traffic Light and Traffic Sign Detectors are implemented using Viola–Jones algorithm. To develop an object detector, own data set is collected. The drawback of this technique is the predetermined dimension of detected object [8]. Object detection model implemented using YOLOv3 algorithm are Cars, Traffic Lights, Pedestrians and Stop Signs. The drawback of this algorithm is that “it does not give a good performance on larger size objects” [12]. The evaluation of object detection algorithm shows that processing speed of object detection model based on YOLOv3 provides enhanced result, and also improvement in outcome of precision and recall. The evaluation of object tracking is done based on used test videos using MOTA (Multiple Object Tracking Accuracy). The result shows that it achieved similar accuracy (0.86 & 0.87) in both tracking methods. When compared with correlation tracking method (36.7), median flow tracking method (66.6) offers a better processing speed. Author says that, “The proposed system is not reliable enough to be used in autonomous vehicle due to obtained accuracy and processing speed. It also does not meet real time constraints of automotive application”. Therefore it is essential to train object detectors with additional data to increase its precision and recall.

In real time driving environment, Visual Object Detection is a very crucial problem due to the difficulties from illumination variation, perspective distortion, object occlusion, scale of object variation, inter-object structural variations and different weather condition (rainy, snowy, etc.,). Recently Convolutional Neural Networks (CNN) [13, 14] provides best achievement on object detection in terms of accuracy and speed. To address this crucial problem (large object scale variation, object occlusion, bad light condition), Wei et al. [15] proposed three enhancements on a Multi-Scale CNN model (MS-CNN) [16]. First enhancement, deconvolution of CNN feature is applied at smaller feature output scales, which is further fused with features at larger feature output scales, to provide richer context for object detection at individual feature output scale. It can address the large object scale variation challenge. Second, soft non maximal suppression (NMS) [17, 18] was applied through object proposals at different feature output scales to identify objects in occluded condition. Third, the aspect ratio of objects is diverse. For accurate object identification and localization, setting anchor box with proper measure of aspect ratio of object. It is achieved by using OPBB (Object Proposal Building Block).

The suggested framework was verified by experimental results over KITTI dataset with improved or comparable detection which displayed significant object detection performance improvement over traditional CNN method. The test result shows that the detection performance is improved by this method than MS-CNN from several aspects. That are False Proposals are reduced, more small objects are detected, many bounding boxes for detected object can be avoided. Table 1 describes achieved precision and recall of object detection model.

Table 1 Achieved precision and recall of object detection model

Literature	Methodology used	Achieved precision and recall
Ciberlin et al. [7]	Viola–Jones algorithm and YOLOv3 algorithm for object detection Median Flow tracking and correlation method for object tracking	Achieved precision and recall for detection model are: cars is 0.88 and 0.86 Pedestrians is 0.9 and 0.92 Traffic light is 0.80 and 0.84 Stop sign is 0.9 and 0.89 Processing speed is 1.2 Processing speed of tracking methods: median flow is 66.6 and correlation tracker is 36.7 Accuracy of object tracking methods: median flow is 0.86 and correlation tracker is 0.87
Wei et al. [13]	Enhanced MS-CNN model	Achieved average precision Pedestrian detection Easy—85.12, moderate—74.52, hard—69.35 Car detection Easy—90.49, moderate—89.64, hard—77.95 Time taken—0.24 s

3 Vehicle Detection

In [19] to detect vehicles at different environmental lighting condition (day, dusk and dark), for better resource utilization and to increase speed and accuracy, Hemmati et al. developed an adaptive vehicle detection system, in which Partial Reconfiguration (PR) controller [20, 21] is employed on Zynq SoC, which is triggered for reconfiguration process by the exterior signal which indicates the changes in light intensity in the environment. This has achieved reliable resource utilization during vehicle detection. It takes 20 ms for the reconfiguration process. The module for pedestrian detection is implemented in the hardware system. It automatically finds the pedestrian during reconfiguration process; this is not changed during the vehicle detection. Histogram of Oriented Gradients (HOG) [22] and Support Vector Machine (SVM) [23] detection methods are applied to detect vehicles during day and dusk (low light) condition. Two different datasets are selected to train day and dusk models. UPM vehicle dataset [24] is used for day conditions and SYSU night time vehicle dataset [25] is used for dusk conditions. The experimental test result shows that there is considerable improvement in the accuracy of detection. Deep Belief Network (DBN) is used to detect vehicles in dark environment conditions by extracting features of vehicle taillight [26, 27]. A subset of SYSU dataset is used for testing this model and an accuracy of 95% is obtained. Finally the proposed method is trained for real time object detection in different lighting conditions and is capable of running at 125 MHz at the rate of 50 fps.

To detect the vehicle in high closing off environment, Zhang et al. [28] developed vehicle detector in two stages based on Faster R-CNN [29] (Region-based Convolutional Neural Network). In the first stage, the part aware the Region Proposal Network (RPN) [30] concentrate on global and local feature of vehicle and localize the vehicle in the highly occlusion environment. In the second stage, the part aware non-maximum suppression (NMS) [31] used to locate the vehicle when high overlap occurs between vehicles under high IoU (Intersection-over-Union). The proposed detector is evaluated on KITTI dataset [32] and compared with existing methods [33–35]. Datasets are classified as Easy, Moderate and Hard on the basis of difficulties among occlusion, truncation and object size. Achieved Recall on vehicle detection performance with IoU of 0.7 is Easy-90.21, moderate-89.01, and hard-80.72. Test result shows that best performance is obtained on the KITTI dataset for different IoU thresholds compared with Faster R-CNN and traditional methods. It achieves better performance with part-aware NMS compared with old NMS. Finally the proposed method proves that improvement in the performance of accuracy is obtained when the vehicles are highly occluded.

When driving at night, due to poor illumination identification of the objects are very tough. Camera can't grasp well in the dark. To address this problem Mayr et al. [36] suggested machine learning method Deep Neural Network (DNN) to automatically estimate the pitch angle of the headlight from camera images using physical relationships between headlight settings and road characteristics. The DNN approach achieves best performance compared with hand-crafted baseline method. This method is challenged by obstacles in front of the vehicle, external light source and slope.

3.1 Extraction of Accurate Vehicle Trajectory

Data set of vehicle trajectories is presented for very particular circumstances such as driving in highways and having small proportion of trajectories. Therefore there is need of dataset for vehicle trajectories in global manner. To address this problem, in [37] introduced a framework, it automatically create vehicle trajectories [38–41] in accurate on the basis of information gathered on fixed monocular traffic camera. The camera is placed in high traffic condition where a high interaction occurs to collect real time accurate traffic data. Aim of this framework is to generate dataset. It should satisfy the requirements to learn driving behaviour effectively, to provide states of vehicle accurately and the capability to scale, therefore the dataset imitates the change of real time situations. The evaluation result shows that the obtained accuracy is 78.6%. Table 2 presents the achieved accuracy of various algorithms used for vehicle detection.

3.2 Pedestrian Detection

While moving with the vehicle pedestrian detection is often challenging, because pedestrians are normally very small on images due to the distance and image resolution and unpredictable movement of pedestrian.

In this paper [42], two cameras are used for pedestrian identification, stereo vision and thermal cameras [43, 44] are drawn together on a vehicle for consistency of data collection. The disparity data from stereo vision camera and thermal data from thermal camera are collected and aligned using trifocal tensor [45] as per the report of the point registration. In this framework, HOG with SVM [46, 47] or CCF (Convolutional Channel Features) with AdaBoost is used for feature extraction and classification and is executed individually for each data source before the decision fusion stage. The information is collected from one or various classifiers to detect the pedestrian in the decision fusion stage. The evaluation of proposed system proved that CCF considerably outstrips the HOG. The achieved log average miss rate is 9%.

Pedestrian Trajectory Foretelling is a very challenging problem for autonomous vehicle. Styles et al. [48] had addressed a deep learning approach Dynamic Trajectory Predictor

Table 2 Achieved accuracy of algorithms used in vehicle detection model

Literature	Methodology used	Achieved accuracy
Hemmati et al. [18]	DBN is used to detect vehicles in the presence of taillights for dark environment HOG + SVM model used in the day and dusk condition Partial configuration controller on Zynq SoC	Achieved accuracy 86% in dusk environment and 95% in dark environment
Zhang et al. [27]	Adaptive faster CNN Part aware RPN & part aware NMS	Achieved recall on KITTI dataset with IoU 0.7 Occlusion easy 90.21, moderate 89.01 and hard 80.72 Time taken—2.1
Mayr et al. [36]	Deep learning based methods DNN	Achieved accuracy 57.90 and MAE 0.05
Clausse et al. [37]	Trajectory extraction framework	Achieved tracking accuracy is 78.6%

(DTP) for estimating pedestrian trajectory [49] using single mounted camera and developed machine annotation scheme [50, 51]. It automatically trains the DTP model using large dataset without human footnote that is in the absence of labelled data. JAAD dataset and two metrics mean squared error (MSE) and displacement error (DE@t) at timestamps up to 15 is used to evaluate performance of DTP. The achieved MSE is 610 ± 21 and DE@15 is 34.6 ± 0.5 for 9 optical flow frames before pre training with machine annotation scheme the performance of DTP is evaluated after pre-training with machine annotation scheme [50] using YOLOv3 and Faster R-CNN object detectors. The effect of pre training on BDD-10 K with YOLOv3, MSE is 539 ± 13 and DE@15 is 32.7 ± 0.4. Using this good performance for the forecasting of pedestrian trajectory is achieved. Table 3 displays achieved miss rate and MSE of pedestrian detection model.

3.3 Efficient Driving Scene Image Creation

Identifying single object in front of the vehicle is easy. But gathering data about various circumstances to realize this scenario as well as to recognize the object is very difficult. In this case, it is essential to be familiar with the location itself instead of identifying only one object. For this purpose Choi et al. [52] proposed a framework is Generative Adversarial Network (GAN) [53–55]. It can effectively create a training image with the context and can learn the location and size relationships of objects successfully. GAN is trained through the feature map generated from the object in the situations recognized by YOLO model [56]. After this the situation images are created from the feature map data learned by GAN. The test result shows that GAN generate a real situation scenario effectively.

3.4 Basis for Autonomous Vehicle Navigation

An autonomous vehicle requires the proper vehicle control. To address this in [57] Pandey et al. developed a DNN (Deep Neural Network) framework [58–60] that organize two control parameters like vehicle speed and the angle for steering for smooth navigation. The proposed framework was designed using convolutional network with eight layers, with two gated recurrent and four dense layers were trained and tested on GPU based system. The test result shows that achieved mean squared error (MSE) are 1.79 and 2.69% for speed and inclination respectively with 38 ms for real time response delay. Table 4 shows that achieved precision and MSE of framework used for capture driving environment and vehicle control.

4 Conclusion

Now a days all automobile industries focusing on developing fully autonomous vehicle (level 5, no driver is required). An accurate environment perception and realization is essential for an AV to operate safely. Machine learning and deep learning algorithms are most commonly used in autonomous vehicles for perception, decision-making and autonomous navigation. This paper addresses all the major aspects of an object detection frame work in the past decade. For each aspect, the state-of-art research works are discussed in detail. Thus the paper presents a concise summary of all available algorithms in object detection for AV platform such as yolov3, Viola–Jones algorithm, enhanced CNN model,

Table 3 Achieved miss rate and MSE of pedestrian detection model

Literature	Methodology used	Achieved miss rate and MSE
Chen et al. [42]	Combine stereo vision camera with a thermal camera used for pedestrian detection Trifocal tensor—align data from multiple camera Both HOG and CCF detection methods used	Achieved average miss rate 9% log without disparity information using CCF method Achieved 13% log average miss rate with disparity information using CCF method
Styles et al. [46]	Dynamic trajectory predictor (DTP) and machine annotation scheme	Achieved MSE and DE@15 for DTP model using JAAD dataset are 610 ± 21 and 34.6 ± 0.5 Using machine annotation scheme achieved MSE and DE@15 on BDD-10 K dataset are 539 ± 13 and 32.7 ± 0.4

Table 4 Achieved precision and MSE

Literature	Methodology used	Achieved precision and MSE
Choi et al. [50]	Generative adversarial network YOLO used for object detection process	Mean average precision using GAN method is 0.83
Pandey et al. [54]	Deep neural network framework designed	Achieved MSE are 1.79 and 2.69% for speed and inclination respectively with 38 ms for real time response delay

DBN, HOG with SVM and CCF, DNN etc. Each algorithm achieves better performance under different driving conditions. From the above discussion, the Enhanced MS-CNN model offers better performance over traditional methods used for object detection in autonomous vehicle application. Also its accuracy and speed are good in different driving conditions such as different lighting condition, object occlusion and different scale of objects. Hence from this survey, future researchers in this domain can grasp the multidimensional point on best algorithm for a real time efficient and scalable object detection frame work with less computational cost, processing time and more accuracy.

References

1. https://media.ford.com/content/fordmedia/fna/us/en/news/2017/02/10/ford-invests-in-argo-ai-new-artificial-intelligence-company.html.
2. https://pressroom.toyota.com/toyota+establish+artificial+intelligence+research+development+company.
3. https://www.media.volvocars.com/us/en-us/media/pressreleases/194795/volvo-cars-and-uber-join-forces-to-develop-autonomous-driving-cars.
4. https://www.press.bmwgroup.com/global/article/detail/T0258269EN/bmw-group-driving-the-transformation-of-individual-mobility-with-its-strategy-number-one-next.
5. https://www.cbinsights.com/research/autonomous-driverless-vehicles-corporations-list/.
6. https://emerj.com/ai-adoption-timelines/Self-Driving-Car-Timeline-themselves-top-11-automeakers.
7. Ciberlin, J., Grbic, R., Teslic, N., & Pilipovic, M. (2019). Object detection and object tracking in front of the vehicle using front view camera. In *IEEE conference on zooming innovation in consumer technologies conference (ZINC)*, Novi Sad.
8. Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *Computer vision and pattern recognition*, USA.
9. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified real time object detection. In *The IEEE conference on computer vision and pattern recognition* (pp. 779–788).
10. Kalal, Z., Mikolajczyk, K., & Matas, J. (2010). Forward–backward error: Automatic detection of tracking failures. In *International conference on pattern recognition*.
11. Danelljan, M., Häger, G., Khan, F. S., & Felsberg, M. (2014). Accurate scale estimation for robust visual tracking. In *Proceedings of the British machine vision conference BMVC*.
12. Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. arXiv:1804.02767.
13. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Procedings of NIPS* (pp. 1097–1105).
14. Kebria, P. M., Khosravi, A., Salaken, S. M., & Nahavandi, S. (2020). Deep imitation learning for autonomous vehicles based on convolutional neural networks. *IEEE/CAA Journal of Automatica Sinica, 7*(1), 82–95.
15. Wei, J., He, J., Zhou, Y., Chen, K., Tang, Z., & Xiong, Z. (2019). Enhanced object detection with deep convolutional neural networks for. advanced driving assistance. *IEEE Transactions on Intelligent Transportation Systems, 99,* 1–12.

16. Cai, Z., Fan, Q., Feris, R. S., & Vasconcelos, N. (2016). A unified multi-scale deep convolutional neural network for fast object detection. In *Computer vision—ECCV* (pp. 354–370). New York, NY, USA: Springer.
17. Bodia, N., Singh, B., Chellappa, R., & Davis, L. S. (2017). Improving object detection with one line of code.
18. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems* (pp. 91–99).
19. Hemmati, M., Biglari-Abhari, M., & Niar, S. (2019). Adaptive vehicle detection for real-time autonomous driving system. In *IEEE conference on design, automation & test in Europe conference & exhibition*.
20. Partial Reconfiguration User Guide. (2017). Xilinx Inc.
21. Vipin, K., & Fahmy, S. A. (2014). ZyCAP: Efficient partial reconfiguration management on the Xilinx Zynq. *IEEE Embedded Systems Letters, 6*(3), 41–44.
22. Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)* (Vol. 1, pp. 886–893).
23. Vapnik, V., Golowich, S. E., & Smola, A. (1996). Support vector method for function approximation regression estimation and signal processing. In *Advances in neural information processing systems* (Vol. 9, pp. 281–287). MIT Press.
24. Arróspide, J., Salgado, L., & Nieto, M. (2012). Video analysis-based vehicle detection and tracking using an MCMC sampling framework. *EURASIP Journal on Advances in Signal Processing, 2012*(1), 2. https://doi.org/10.1186/1687-6180-2012-2.
25. Chen, L., Hu, X., Xu, T., Kuang, H., & Li, Q. (2017). Turn signal detection during nighttime by cnn detector and perceptual hashing tracking. *IEEE Transactions on Intelligent Transportation Systems, 99,* 1–12.
26. Chien, C. L., Hang, H. M., Tseng, D. C., & Chen, Y. S. (2016). An image based overexposed taillight detection method for frontal vehicle detection in night vision. In *2016 Asia-Pacific signal and information processing association annual summit and conference (APSIPA)* (pp. 1–9).
27. O'Malley, R., Jones, E., & Glavin, M. (2010). Rear-lamp vehicle detection and tracking in low-exposure color video for night conditions. *IEEE Transactions on Intelligent Transportation Systems, 11*(2), 453–462.
28. Zhang, W., Zheng, Y., Gao, Q., & Mi, Z. (2019). Part-aware region proposal for vehicle detection in high occlusion environment. *IEEE Access, 7,* 100383–100393.
29. Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 39*(6), 1137–1149.
30. Cao, Z., Simon, T., Wei, S. -E., & Sheikh, Y. (2017). Realtime multi-person 2D pose estimation using part affinity fields. In *Proceedings of IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 1302–1310).
31. Liu, S., Lu, C., & Jia, J. (2015). Box aggregation for proposal decimation: Last mile of object detection. In *Proceedings of IEEE international conference on computer vision (ICCV)* (pp. 2569–2577).
32. Li, B., Wu, T., & Zhu, S. -C. (2014). Integrating context and occlusion for car detection by hierarchical and-or model. In *Computer vision—ECCV*. Cham, Switzerland: Springer.
33. Mousavian, A., Anguelov, D., Košecká, J., & Flynn, J. (2017). 3D bounding box estimation using deep learning and geometry. In *Proceedings of IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 5632–5640).
34. Xiang, Y., Choi, W., Lin, Y., & Savarese, S. (2017). Subcategory-aware convolutional neural networks for object proposals and detection. In *Proceedings of IEEE winter conference on applications of computer vision (WACV)* (pp. 924–933).
35. Chabot, F., Chaouch, M., Rabarisoa, J., Teulière, C., & Chateau, T. (2017). Deep MANTA: A coarse-to-fine many-task network for joint 2D and 3D vehicle analysis from monocular image. In *Proceedings of IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 1827–1836).
36. Mayr, J., Giracoglu, C., Unger, C., & Tombari, F. (2019). Headlight range estimation for autonomous driving using deep neural networks. In *IEEE conference on intelligent vehicles symposium (IV)*.
37. Clausse, A., Benslimane, S., & de La Fortelle, A. (2019). Large-scale extraction of accurate vehicle trajectories for driving behavior learning. In *IEEE conference on intelligent vehicles symposium (IV)*.
38. Ren, X., Wang, D., Laskey, M., & Goldberg, K. (2018). Learning traffic behaviors by extracting vehicle trajectories from online video streams. In *14th IEEE international conference on automation science and engineering CASE 2018* (pp. 1276–1283).

39. Simon, D. (2006). *Optimal state estimation: Kalman H infinity and nonlinear approaches*. New York, NY: Wiley-Interscience.
40. Behbahani, F., Shiarlis, K., Chen, X., Kurin, V., Kasewa, S., Stirbu, C., Gomes, J., Paul, S., Oliehoek, F. A., Messias, J. V., & Whiteson, S. (2018). Learning from demonstration in the wild. arXiv:1811.03516.
41. Bochinski, E., Eiselein, V., & Sikora, T. (2017). High-speed tracking-by-detection without using image information. In *International workshop on traffic and street surveillance for safety and security at IEEE AVSS 2017*.
42. Chen, Z., & Huang, X. (2019). Pedestrian detection for autonomous vehicle using multi-spectral cameras. *IEEE Transaction on Intelligent Vehicles, 4,* 211–219.
43. Gade, R., & Moeslund, T. B. (2014). Thermal cameras and applications: A survey. *Machine Vision and Applications, 25*(1), 245–262. https://doi.org/10.1007/s00138-013-0570-5.
44. Krotosky, S. J., & Trivedi, M. M. (2007). On color- infrared- and multimodal-stereo approaches to pedestrian detection. *IEEE Transactions on Intelligent Transportation Systems, 8*(4), 619–629.
45. Hartley, R., & Zisserman, A. (2003). *Multiple view geometry in computer vision*. Cambridge: Cambridge University Press.
46. Jeon, H. -M., Nguyen, V. D., & Jeon, J. W. (2019). Pedestrian detection based on deep learning. In *Conference of the IEEE industrial electronics society*.
47. Hbaieb, A., Rezgui, J., & Chaari, L. (2019). Pedestrian detection for autonomous driving within cooperative communication system. In *IEEE wireless communications and networking conference (WCNC)*.
48. Styles, O., Ross, A., & Sanchez, V. (2019). Forecasting pedestrian trajectory with machine-annotated training data. In *IEEE conference on intelligent vehicles symposium (IV)*.
49. Ilg, E., Mayer, N., Saikia, T., Keuper, M., Dosovitskiy, A., & Brox, T. (2017). Flownet 2.0: Evolution of optical flow estimation with deep networks. In *CVPR*.
50. Wojke, N., Bewley, A., & Paulus, D. (2017). Simple online and realtime tracking with a deep association metric. In *ICIP*.
51. Yagi, T., Mangalam, K., Yonetani, R., & Sato, Y. (2018). Future person localization in first-person videos. In *CVPR*.
52. Choi, S. -Y., Jeong, H. -J., Park, K. -S., & Ha, Y. -G. (2019). Efficient driving scene image creation using deep neural network. In *2019 IEEE international conference on big data and smart computing (BigComp)*.
53. Goodfellow, L. J., et al. (2014). Generative adversarial networks. arXiv preprint arXiv:1406.2661.
54. Alec, R., et al. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv:1511.06434.
55. Odena, A. (2016). Semi-supervised learning with generative adversarial networks. arXiv preprint arXiv:1606.01583.
56. Masmoudi, M., Ghazzai, H., Frikha, M., & Massoud, Y. (2019). Object detection learning techniques for autonomous vehicle applications. In *IEEE international conference on vehicular electronics and safety (ICVES)*.
57. Pandey, D., & Niwaria, K. (2019). A novel single front camera based simpler approach for autonomous taxi navigation for smart city deployments. In *6th international conference on signal processing and integrated networks (SPIN)*, IEEE.
58. Hillel, A. B., Lerner, R., Levi, D., & Raz, G. (2014). Recent progress in road and lane detection: A survey. *Machine Vision and Applications, 25*(3), 727–745.
59. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521,* 436–444.
60. Xu, H., Bojarski, M., et al. (2017). End-to-end learning of driving models from large-scale video datasets. Available: openaccess.thecvf.com.

S. Devi received her Master of Engineering degree in Computer Science and Engineering from Anna University, Tamilnadu, India, in 2012 and Bachelor of Engineering degree in Computer Science and Engineering from Anna University, Tamilnadu, India, in 2005. Her main research interests are computer vision, AI, Machine learning and deep learning.

P. Malarvezhi received her Ph.D. degree from SRM Institute of Science and Technology, India (2018), Master of Engineering degree in Communication System from SRM University, Tamilnadu, India in 2011 and Bachelor of Engineering degree in Electronics and Communication Engineering from Madras University, Tamilnadu, India in 2004. Her main research interests are wireless communication system design with main emphasis on next generation communication systems, IOT, smart communication with AI, ML and DL platforms.

R. Dayana received her Ph.D. degree from SRM Institute of Science and Technology, India (2018), Master of Engineering degree in Communication System from SRM University, Tamilnadu, India in 2011 and Bachelor of Engineering degree in Electronics and Communication Engineering from Madurai Kamaraj University, Tamilnadu, India, in 2004. Her main research interests are in wireless communication system design with main emphasis on multiple-input-multiple-output (MIMO) system, cognitive radio networks, Universal Filtered Multicarrier (UFMC), green communication and 5G mobile communications.

K. Vadivukkarasi received her Ph.D. degree from SRM Institute of Science and Technology, India (2016), Master of Engineering degree in Embedded System Technology from Anna University, Tamilnadu, India in 2007 and Bachelor of Engineering degree in Electronics and Communication Engineering from Madras University, Tamilnadu, India, in 2001. She is working an assistant professor in the Department of Electronics and Communication Engineering, SRM University, India. Her main research interests are in wireless sensor networks, Embedded system design, Embedded networking and Internet of Things.