Original Research Article

Concurrent Engineering: Research and Applications
2022, Vol. 0(0) 1–11

DOI: 10.1177/1063293X211069193
journals.sagepub.com/home/cer
SAGE

A cost-effective computer vision-based vehicle detection system

Altaf Alam[1], Zainul Abdin Jaffery[1] and Himanshu Sharma[2]

Abstract

Vehicle detection plays an important role in the development of an autonomous driving system. Fast processing and accurate detection are two major aspects of generating the autonomous vehicle detection system. This paper proposes a novel computer vision-based cost-effective vehicle detection system. Here, a Gentle Adaptive Boosting algorithm is trained with Haar-like features to generate the hypothesis of vehicles. Haar-like feature generates hypotheses very fast but may detect false vehicle candidates. The support vector machine algorithm is trained with the histogram of oriented gradient features to filter out the generated false hypothesis. The histogram of oriented gradients descriptor utilizes the shape and outlines of the vehicles, hence detects vehicles more accurately. Haar-Likes features and histogram of oriented gradients features are organized to accomplish the aspects of autonomous driving. The performance of the proposed vehicle detector is evaluated for day time and night time captured images and compared with three different existing vehicle detectors. The average precision of the proposed system for day time captured image is 0.97 and for night time captured image is 0.94. The proposed system requires 15 times less training time as compared to the existing technique for the same number of image data and on the same CPU.

Keywords

vehicle detection system, Haar-like feature, histogram of oriented gradients features, Adaboost classifier, support vector machine, autonomous driving system

Introduction

A vehicle that can sense the surrounding environment and drive itself without human driver is known as an autonomous vehicle. Autonomous driving vehicle needs to perform road detection, divider detection, traffic sign, traffic light signal detection, pedestrian detection, and detection of another vehicle on road to drive the vehicles safely. According to the global status report on road safety by the World Health Organization (WHO), approximately 1.35 million people die every year due to road accidents in the world (World Health Organization, 2018). A major reason for road accidents is the fatiguing nature of drivers due to over workload. A driver assistance system or a fully autonomous vehicle may reduce the collision occurrence by taking the workload of drivers. The autonomous vehicle utilizes different types of sensors such as Radio Detection and Ranging (Radar), Light Detection and Ranging (Lidar), Sound Navigation Ranging (Sonar), Global Positioning System (GPS), and complex algorithms to execute the driving task. Recently, vision sensors have become popular as a sensing device for autonomous vehicle. Vision sensors monitor the surrounding condition similar to the human driver and are cost-effective. This research work proposed a vision-based system for vehicle detection on road. Many vision-based vehicle detection systems have been proposed by the researchers (Jazayeri et al., 2011; Nguyen et al., 2013; Toulminet et al., 2006; Zielke et al., 1992;Zhang, 2018) in the last few decades. The available literature categorizes the vehicle detection work into three classes (Bertozzi and Broggi, 1998) such as Appearance-based, stereo vision-based, and features based vehicle detection system. Appearance-based vehicle detection techniques utilize some specific information like shadow underneath the vehicle (Jazayeri et al., 2011), horizontal and vertical edges (Nguyen et al., 2013), contour symmetry (Toulminet et al., 2006; Zielke et al., 1992), and color information (Zhang, 2018) of vehicle. Appearance information based vehicle detector detects vehicle for some condition but it did

[1]Department of Electrical Engineering, Jamia Millia Islamia, New Delhi, India
[2]Electronics and Communication, KIET Group of Institutions, Ghaziabad, India

Corresponding author:
Zainul Abdin Jaffery, Department of Electrical Engineering, Jamia Millia Islamia, Jamia Nagar, New Delhi 110025, India.
Email: zjaffery@jmi.ac.in

not detect vehicle for the various condition due to the dynamic nature of the outdoor environment. The shadow of the vehicle varies with light variation, and the size and position of the vehicle vary with the motion that causes faulty detection. Edge information based vehicle detection system countenances the issue of perfect threshold selection for vehicle candidates due to the alteration in distance (Unzueta et al., 2012). Contour symmetry information based vehicle detector depends on complex rules that utilize the histogram of contour, horizontal, and vertical line symmetry of the vehicle (Liu1 et al., 2013). A stereo vision is another method utilized for vehicle detection (Cai et al., 2018; Sivaraman and Trivedi, 2013). The stereo vision system accounted for two cameras that collect the surrounding environment information. The optical flow also has resources to detect the vehicle; however, it detects vehicles accurately for a stationary background only (Urmson and Bagnell, 2008). In recent years, computer vision has achieved great advancement in feature extraction, feature description, and object recognition (Alam and Jaffery, 2020). The computer vision-based vehicle detection utilizes appearance, optical flow, and color information for hypothesis generation and robust features such as Scale-Invariant Feature Transform (SIFT), Speeded up Robust Features (SURF), Histogram of Oriented Gradients (HOG), etc. for hypothesis verification (Sahin and Ozkan, 2014). A trained classifier can generate hypotheses more accurately rather than random prediction. Thirty three stages of cascade classifiers were trained with Haar-like features to detect the vehicle in (Tang et al., 2017). Whereas, a neural network has been trained with Haar-like features to detect the vehicle by (Ai et al., 2017). Haar-like feature-based vehicle detection is fast but less accurate, so the Support Vector Machine (SVM) classifier was trained with the HOG descriptor in (Wu et al., 2014). Apart from machine learning-based detection, the deep neural network is also a very effective method for vision-based vehicle detection (Schilling et al., 2018). Convolution neural networks (CNNs) (Tarmizi and Aziz, 2018), region-based convolution neural network (RCNN) (Deng et al., 2017), and You Only Look Once (YOLO) (Nguyen et al., 2019) are some common deep learning approach for vehicle detection. A deep learning-based vehicle detector requires a graphic processing unit (GPU), a huge amount of images, and large processing time (Schilling et al., 2018). GPU reduces the training time requirement, but at the same time, it enhances the cost of the product. Machine learning-based vehicle detection neither requires GPU nor a huge amount of image. If a power full GPU is not available and the number of images is limited, in that case, machine learning approach is a better choice for vehicle detection. The performance of a machine learning-based vehicle detection system depends only on the utilized features and training algorithms. Hence, optimize feature and good algorithms selection can create an accurate, less complex, and less costly system. A single feature-based vehicle detector may not be more accurate; however, blend of multiple features can detect the vehicle accurately. Therefore, this work combines two different features that made the system robust against the dynamic nature of the outdoor environment, cost-effective, accurate, and appropriate for real-time processing. Haar-like features with the AdaBoost classifier (Tang et al., 2017) and HOG descriptors with SVM classifier (Matsumoto, 2012) techniques are utilized for vehicle detection. Haar-like features use the integral image to extract the features, so its processing is fast but it uses brighter and darker region information around the object rather than the shape information; hence, its detection is less accurate. Contrarily, the HOG features-based vehicle detection utilizes shape information of the object and subsequently detection is accurate. Although HOG features-based vehicle detection is accurate, HOG features extraction and vehicle similarity matching in the whole image are a time-consuming process and confers multiple bounding boxes around a single-vehicle candidate (Madhogaria et al., 2015). Haar-like feature-based vehicle detection is fast but less accurate while HOG features-based vehicle detection is accurate but slow that makes them complementary for each other. A meticulous utilization of these features can create a convincing vehicle detector. In this work, Haar-like features conferred the bounding box around a region in the image for a vehicle, and HOG features verify those regions as vehicle and non-vehicle. Accuracy of the Haar-like features-based vehicle detector is enhanced by combining them with HOG features. Conversely, multiple bounding boxes detections of HOG based detection are transformed into single bounding box detection by verifying the bounding box region instead of the vehicle detection in the whole image.

In the proposed vehicle detector, the Haar-Like feature and HOG features are correlated with each other that achieved fast processing and accurate detection. Moreover, the performance of vehicle detection is evaluated for the proposed vehicle detector, Haar cascade, HOG-SVM, and YOLO-V3 detector. Results concluded that the overall performance of the proposed system is lucrative without any high-cost GPU utilization and without large training time.

The rest of the paper is organized as follows. Data collection, Feature extraction, and training detail of hypothesis generation classifier and hypothesis verification classifier are described in the proposed vehicle detection systems section. Simulation implementation, detection result, and evaluation of the system performance are discussed in the simulation and results section. The conclusions section concludes the work by highlighting the achieved results and effectiveness of the system.

Proposed vehicle detection systems

The process flow diagram of the proposed vehicle detection system is shown in Figure 1. The proposed system divides the acquired image into sub-regions and dispatches those regions to the AdaBoost classifier. The AdaBoost classifier generates the hypothesis by conferring the bounding box

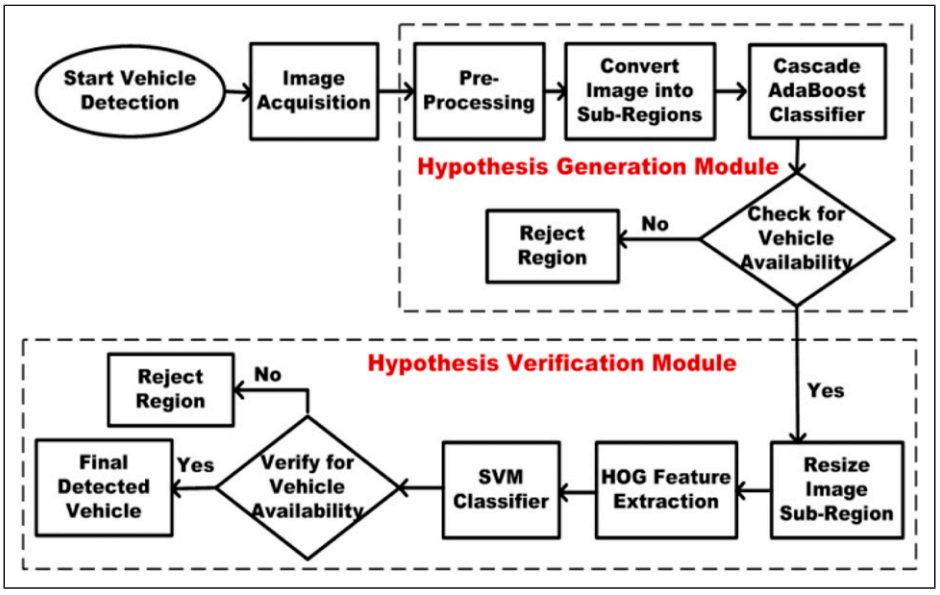

Figure 1. Process flow diagram of the proposed vehicle detector.

around the region in the image. Generated hypothesis conveyed to the SVM classifier that assigned a label as vehicle or non-vehicle to the hypothesis. The system considered vehicle availability in the region based on the decision of the SVM classifier.

Classifier creation for hypothesis generation

The AdaBoost algorithm trained and created a classifier to hypothesize the vehicle. In the boosting process, 20 stages of the classifier were cascaded and aggregated the results for the hypothesis. The combination of multiple classifiers with the selection of training feature sets at every iteration and assigning the right amount of weight in final voting can get better accuracy. Hence, this work ensembles 20 classifiers and created a strong classifier for hypothesizing the vehicle. Training data, training features, and training algorithm detail are given as follows.

Training data. Classifier training required both positive as well as negative data set. A custom data set from the real-world images is collected manually and cropped out to the vehicle and non-vehicle region for training. A total of 2600 images are collected for the training data set. Out of 2600 images, 1300 images are of vehicle and 1300 images are of non-vehicle data. The vehicle training image sample is shown in Figure 2(a) and the non-vehicle training image sample is shown in Figure 2(b).

Feature extraction. Direct use of the pixel information in classifier creation makes the system more complex, time-consuming, and less accurate. The feature extraction process kept only relevant information and dropout the irrelevant ones. Hence, it reduces the redundancy and dimensionality of image information. Haar-like features were extracted and utilized for training. Haar-like features related masks are shown in Figure 3(a) and (b) shows the Haar-like feature mask sliding on the vehicle. Table 1 summarizes information about the extracted features from the vehicle and non-vehicle training images. These features were utilized for the training process. The total extracted features from the training images are 581029.

AdaBoost algorithm training. Twenty weak classifiers are cascaded to make a strong classifier. Every weak, classifier has an optimized threshold function h (x, f, p, θ), which provides the limitation to misclassification. Mathematical expression for weak classifier is given by equation (1).

$$h(x_i,f,p,\theta) = \begin{Bmatrix} 1 & if\ pf(x) < p\theta \\ 0 & otherwise \end{Bmatrix} \tag{1}$$

where x_i represents the sample used at the current stage of training, f represents features used in that stage, θ is a threshold value for that features at that stage, and p is a polarity that indicates the direction of the inequality sign. Extracted features from training images are (x_1,y_1), (x_2,y_2), ——————— (x_n,y_n), where x_i represents a sample of extracted features for $i = (1, 2, 3 - - - n)$ and y_i is a class label. If $y_i = 1$, it means that it is a positive sample (vehicle), and if $y_i = -1$, it means that it is a negative sample (non-vehicle). Initialize weights $(w_1,i) = 1/2P$ and $1/2N$ for $y_i = 1\ and - 1$, respectively, where P and N are the numbers of positive and negative samples used in training. Twenty stages of the classifier are trained to hypothesize the vehicle location. Each stage of training creates a classifier and updates its weight. The stage of classifier has been represented as t. In this work, $t = 1, 2, 3 - - - - - - 20$. Weight update and strong classifier selection process are adopted

Figure 2. (a-b) Generated training data. (a) Vehicle data (b) Non-vehicle data.

from (Viola and Jones, 2001). Mathematical representation of strong classifier is represented by equation (2)

$$H(x) = \begin{cases} 1 & \sum_{t=1}^{T} \alpha_i h_t(x) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha_t \\ 0 & otherwise \end{cases} \quad (2)$$

$$\alpha_t = log \frac{1}{\beta_i} \quad (3)$$

where α_t represents applied weight to the strong classifier determined by AdaBoost training.

Training parameters of the classifier such as utilized algorithm, the number of stages, minimum hit rate, maximum false alarm rate, weight trimming, the minimum number of positive features per cluster, and required false alarm rate information at each leaf are given in Table 2. Out of 20 stages, an example of one stage training is shown in Table 3. Where, N is the current stage of training, seven samples of training data is used in this stage. The percentage of the sample (%SMP) is different for every feature, the threshold value for each feature, hit rate (HR), False Alarm (FA), and EXP. ERR is the error information used in this stage.

Classifier creation for hypothesis verification

Haar-like features utilize brighter and darker region information rather than the shape of the vehicle; hence, AdaBoost classifier generated some false hypothesis for vehicle

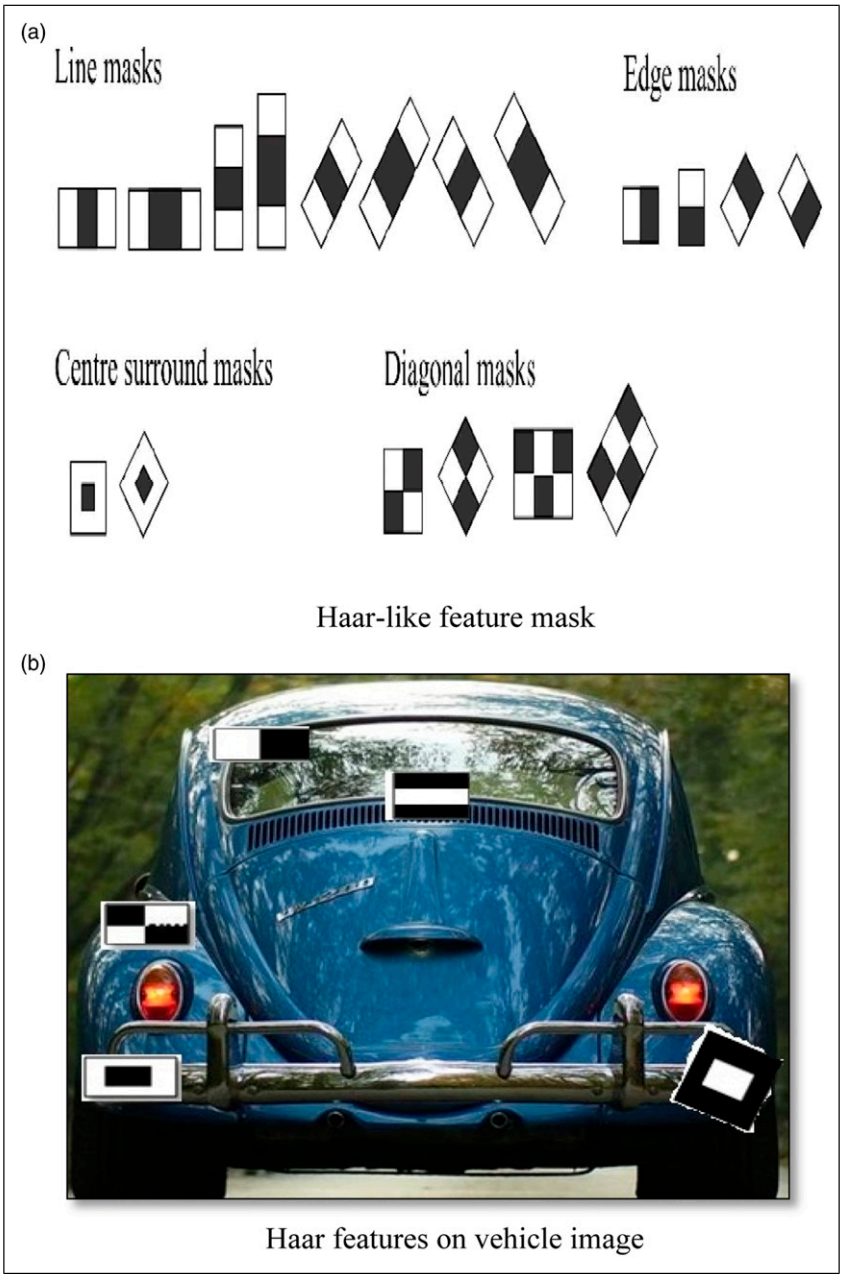

Figure 3. (a-b). Feature extraction for hypothesis generation: (a) Haar-like feature mask. (b) Haar features on vehicle image.

Table 1. Haar-like feature extraction details.

Vehicle image	Non-vehicle image	Image size	Total extracted features
1300	1300	35x35	581029

Table 2. Specified parameters for created classifier.

Applied algorithm	Number of stages	Minimum hit rate	Maximum false alarm	Trimming weight	Positive sample per cluster	Required false alarm
GAB	20	0.995	0 500	0.950	500	$9.5367e10^{-7}$

candidate. Furthermore, to improve the accuracy, hypothesis verification step is added. Hypothesis verification on the same features can convey faulty results again; therefore HOG features were utilized for the verification. SVM algorithm is trained with HOG features and employed as a verification classifier.

Histogram of oriented gradients Feature extraction. The HOG feature extraction process is adopted from (Wang and Zhang, 2017) to extract the features. Mathematical expressions for calculating the gradient and orientation from images are given by equations (4) and (5).

$$M(x,y) = \sqrt{(G_x)^2 + (G_y)^2} \tag{4}$$

$$\theta = tan^{-1}\frac{G_x}{G_y} \tag{5}$$

where $M(x,y)$ represents gradient magnitude of the image, G_x and G_y represent the derivative of an image concerning x and y, and θ represents calculated orientation information of gradient. The utilized parameter details for the HOG descriptor are given in Table 4.

Hypothesis verification classifier training. The support vector algorithm is trained with extracted HOG features to create the verification classifier. SVM classifier extracts the support vector from the training data set and draws a hyperplane to separate the vehicle and non-vehicle candidates (Jung and Kim, 2014). SVM draws many hyperplanes to separate the class but considered the only hyperplane which maximizes the margin between the classes. The mathematical expressions of hyperplanes to separate the vehicle and non-vehicle candidates are shown by equations (6) and (7).

$$W \times x_i - b \geq 1 \quad \text{if } \theta_i = 1 \tag{6}$$

$$W \times x_i - b \leq 1 \quad \text{if } \theta_i = -1 \tag{7}$$

where w_i denotes the weight of hyperplane, θ_i denotes the classes, and x_i is a feature vector. The value of θ_i might be 1 or -1 according to the similarity with a vehicle and non-vehicle class. Image region exists at the right side hyperplane and equation (6) indicated that those regions are the vehicle class, so classifier assigned label 1 to the region. Image region candidate stays at the left side of the Equation plane (7) and indicated that those regions are non-vehicle class, therefore, classifier assigns label -1 to the region.

Distance between the vehicle and non-vehicle class hyperplanes is $2/w_i$ and is known as distance margin.

The distance margin should be maximum between the planes for good accuracy. To maximize the distance margin, w_i should minimum. A combined equation for proper classification of vehicle and non-vehicle is given by equation (8).

$$w_i \min for\ \theta_i\ w \times x_i - b \geq 1 \quad \forall_i = 1,2,3 - - - - - n \tag{8}$$

Equation (8) is an equation of the hyperplane which separates the vehicle and non-vehicle candidates effectively. If any candidates lie above this plane, classifier classifies it as a vehicle candidate while if the candidate lies below the hyperplane, classifier classifies this as non-vehicle candidates.

Simulation and results

The proposed system for vehicle detection is implemented on 3.7 GHz Intel Core i7 with 4 GB primary memory as a processor and MATLAB 2017 as a processing platform. Images were acquired by utilizing a high quality $\frac{1}{4}$ CMOS sensor with resolution 1600 x 1200 under different light condition. A total of 3000 images were collected. Out of 3000 images, 1500 images were collected during daylight condition and 1500 images were collected during night time. The acquired image is divided into sub-regions and passed these regions from created AdaBoost classifier. The created classifier has 20 stages; hence, image sub-regions

Table 3. Training example for one stage.

N	1	2	3	4	5	6	7
%SMP	100	100	89	100	85	94	85
ST.THR	−0.621	−0.846	−1.082	−1.318	−1.680	−1.024	−1.149
HR	1.000	1.000	1.000	1.000.	1.000.	1.000.	1.000.
FA	1.000.	1.000.	1.000.	1.000.	1.000.	0.682	0.391
EXP.ERR	0.222	0.198	0.224	0.137.	0.132	0.132	0.089

Table 4. HOG feature specification.

Image size	Cell size	Block size	Block overlap	Number of bins
64x128	[8,8]	[2,2]	[1,1]	9

are verified in every stage, if any stage of classifier rejects the sub-region, then the system did not deliver those regions into the next stage. AdaBoost classifier conferred bounding boxes for vehicles hypothesis that epitomized the image region as shown in Figure 4(a-b). Figure 4(a) shows the hypothesized vehicle candidates for day time captured image, and Figure 4(b) shows the hypothesized vehicle candidate for night time captured image. Suppose, the AdaBoost classifier has conferred n bounding box regions in a single image. The system extracted HOG features from each region one by one and the SVM classifier verified those regions gradually. The final detected vehicle after verification of the SVM classifier is shown in Figure 4(c-d). Figure 4(c) shows the detected vehicle on day time captured image, and the detected vehicle on night time captured image is shown in Figure 4(d). The system contemplated the detected vehicle by the detector as True Positive (TP), False Negative (FN), and False Positive (FP) detection. Based on TP, FN, and FP detection by the system, the true positive rate (TPR), positive predictive value (PPV), and false

Figure 4. (a-d) Vehicle detection by the proposed system under different light condition: (a) Sample of generated hypothesis by Haar cascade detector on day time captured images, (b) Sample of generated hypothesis by Haar cascade detector on night time captured images, (c) Sample of the detected vehicle after hypothesis verification on day time captured images, (d) Sample of the detected vehicle after hypothesis verification on night time captured image.

discovery rate (FDR) performance parameters of the detectors are calculated. Mathematical expression for TPR, PPV, and FDR calculation is given by equations (9–11).

$$TPR\% = \frac{TP}{TP + FN} \times 100 \tag{9}$$

$$PPV\% = \frac{TP}{TP + FP} \times 100 \tag{10}$$

$$FDR\% = \frac{FP}{TP + FP} \times 100 \tag{11}$$

Performance evaluation

The performance of the vehicle detectors is evaluated based on the above discussed TPR, PPV, and FDR performance parameters. Images were captured under day and night light condition; hence, the detector assessment is carried out for both types of images individually. Table 5 shows the value of TP, FP, and FN detection, and the TPR, PPV, and FDR of each detector for day time captured images, while Table 6 shows the value of TP, FP, and FN detection, and the TPR, PPV, and FDR of each detector for night time captured images. TPR of the proposed system is 97.65%, PPV of the proposed system is 97.08%, and the FDR of the proposed system is 2.91% for day time captured image. Whereas TPR of the proposed system is 95.87%, PPV of the proposed system is 95.08%, and the FDR of the proposed system is 4.89% for night time captured images. Results demonstrated in Table 5 and Table 6 reported that the performance of the proposed vehicle detector is intending towards the YOLO-V3 detector and superior to the Haar cascade and HOG-SVM detector. The proposed system performed lesser than the YOLO-V3 detector, but it is simple, less time consuming, and can be implemented on CPU.

The performance of the detectors furthermore evaluated based on precision, recall, required training time, and taken time in prediction. Table 7 shows the information about the Haar cascade, HOG-SVM, YOLO-V3, and proposed detectors required training time, taken time in detection per image, and average precision. Table 7 notifies that the average precision value of the proposed value is less than the YOLO-V3 detector, but it is more than the Haar cascade and HOG-SVM. The average precision value of YOLO-V3 is

Table 5. Performance evaluations for day time captured image.

Parameters algorithms	True positive detection	False positive detection	False positive detection	True positive rate (%)	False discovery rate (%)	Positive predictive value (%)
Haar cascade	1403	97	99	93.59	6.59	93.43
HOG-SVM	1435	65	72	95.69	4.77	95.23
YOLO-V3	1480	20	26	98.70	1.72	98.3
Proposed	1465	35	44	97.65	2.91	97.08

Table 6. Performance evaluations for night time captured image.

Parameters algorithms	True positive detection	False positive detection	False positive detection	True positive rate (%)	False discovery rate (%)	Positive predictive value (%)
Haar cascade	1236	264	255	82.38	17.10	80.29
HOG-SVM	1309	191	204	87.23	13.48	86.53
YOLO-V3	1443	57	60	96.20	3.99	96.01
Proposed	1438	62	74	95.87	4.89	95.08

Table 7. Training time, detection time, and average precision of the detectors.

Parameters algorithms	Required training time	Detection time per image, sec	Average precision for day image	Average precision for night image
Haar cascade	5 min	0.28	0.92	0.83
HOG-SVM	10 min	0.85	0.94	0.88
YOLO-V3	3 h	0.40	0.98	0.95
Proposed	12 min	0.35	0.97	0.94

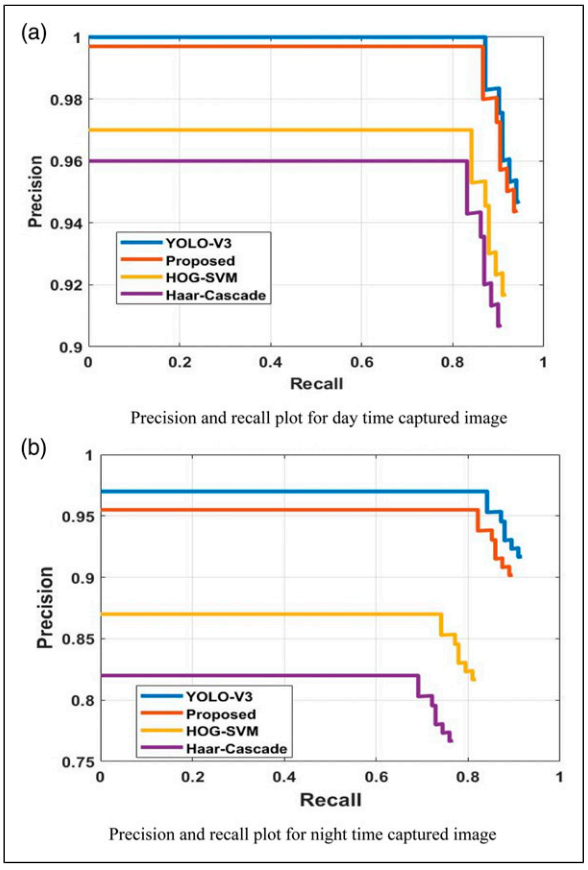

Figure 5. (a-b) Performance evaluation plots for different light condition: (a) Precision and recall plot for day time captured image, (b)Precision and recall plot for night time captured image.

superior to the proposed system; however, simultaneously it requires almost 15 times greater training time than the proposed method. The inference time of the proposed system is larger than the Haar cascade detector, yet the precision of the proposed system is better than Haar cascade and HOG-SVM. Precision and recall plots of the various detectors for day time and night time captured image are shown in Figures 5(a) and (b). Precision and recall plots conclude that the performance of the system is better than the Haar cascade and HOG-SVM but less than that of the YOLO-V3.

Conclusions

In this work, an environmentally adaptive, cost-effective, and simple algorithm is proposed for vehicle detection. The developed algorithm can detect vehicles under different lighting conditions such as day and night time. Two different features Haar-like and HOG features are combined in such a way that their individual quality supports each other to create an accurate and fast vehicle detector. The AdaBoost algorithm was trained with Haar-like features which generate the vehicle hypothesis, and the SVM algorithm was trained with HOG features that verify the generated hypothesis. The proposed vehicle detector performance is evaluated and compared with three different detectors based on TPR, PPV, and FDR parameters. The average precision of the proposed system is 0.97 for day time captured image and 0.94 for night time captured image. The proposed system takes 15 times less training time as compared to the YOLO-V3 vehicle detector. Further, the proposed system gives better performance as compared to conventional methods.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work is supported by a scholarship provided by the University Grants Commission (UGC), Govt. of India under MANF fellowship.

ORCID iDs

Altaf Alam https://orcid.org/0000-0002-7307-1374
Zainul Abdin Jaffery https://orcid.org/0000-0002-3502-6599

References

Ai J, Tian R, Luo Q, et al. (2017) Multi-scale rotation-invariant Haar-like feature integrated CNN-based ship detection algorithm of multiple-target environment in SAR imagery. *IEEE Transactionson Geosciences and Remote Sensing* 57(12): 10070–10087.

Alam A and Jaffery ZA (2020) Indian traffic sign detection, and recognition. *International Journal of Intelligent Transportations Research* 18(1): 98–112.

Bertozzi M and Broggi A (1998) GOLD: a parallel real-time stereo vision system for generic obstacle and lane detection. *IEEE Transactions on Image Processing* 7(1): 62–81.

Cai S, Huang Y, Ye B, et al. (2018) Dynamic illumination optical flow computing for sensing multiple mobile robots from a drone. *IEEE Transactions on Systems Man and Cybernetics* 48(8): 1370–1382.

Deng Z, Sun H, Zhou S, et al. (2017) Toward fast and accurate vehicle detection in aerial images using coupled region-based convolutional neural networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 10(8): 3652–3664.

Jazayeri A, Cai H, Zheng JY, et al. (2011) Vehicle detection and tracking in car video based on motion model. *IEEE Transactions on Intelligent Transportation Systems* 12(2): 583–595.

Jung HG and Kim G (2014) Support vector number reduction: survey and experimental evaluations. *IEEE Transactions on Intelligent Transportation Systems* 15(2): 463–476.

Liu1 S, Huang1 Y and Zhang1 R (2013) On-road vehicle recognition using the symmetry property and snake models. *International Journal of Advanced Robotic Systems* 10(12): 1–9.

Madhogaria S, Baggenstoss PM and Schikora M (2015) Car detection by fusion of HOG and causal MRF. *IEEE Transactions on Aerospace and Electronic Systems* 51(1): 575–590.

Matsumoto M (2012) SVM-based object detection using self-quotient ε-filter and histograms of oriented gradients. *Studies in Computational Intelligence* 399(1): 277–285.

Nguyen VD, Nguyen T and Nguyen D (2013) A fast evolutionary algorithm for real-time vehicle detection. *IEEE Transaction on Vehicular Technology* 62(6): 2453–2468.

Nguyen DT, Nguyen TN, Kim H, et al. (2019) A high-throughput and power-efficient FPGA implementation of YOLO CNN for object detection. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 27(8): 1861–1873.

Sahin I and Ozkan K (2014) A comparative evaluation of well - known feature detectors and descriptors. *International Journal of Applied Mathematics, Electronics, and Computers* 3(1): 1–6.

Schilling H, Bulatov D, Niessner R, et al. (2018) Detection of vehicles in multi sensor data via multi branch convolutional neural networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11(11): 4299–4316.

Sivaraman S and Trivedi M (2013) Looking at vehicles on the road: a survey of vision-based vehicle detection, tracking, and behaviour analysis. *IEEE Transactions on Intelligent Transportation Systems* 14(4): 1773–1795.

Tarmizi IA and Aziz A (2018) Vehicle detection using convolutional neural network for autonomous vehicles. In: International Conference on Intelligent and Advanced System (ICIAS), Kuala Lumpur, Malaysia, 13-14 August 2018, pp. 1–5. IEEE.

Tang Y, Zhang C and Gu R (2017) Vehicle detection and recognition for intelligent traffic surveillance system. *Multimedia Tools Application* 76(1): 5817–5832.

Toulminet G, Bertozzi M, Mousset S, et al. (2006) Vehicle detection by means of stereo vision-based obstacles features extraction and monocular pattern analysis. *IEEE Transactions on Image Processing* 15(8): 2364–2375.

Urmson CA and Bagnell J (2008) Autonomous driving in urban environments: boss and the urban challenge. *Journal of Field Robotics* 25(1): 425–466.

Unzueta L, Nieto M, Cortes A, et al. (2012) Adaptive multicue background subtraction for robust vehicle counting and classification. IEEE Transactions on Intelligent Transportation Systems 3(2): 527–540.

Viola P and Jones M (2001) Rapid object detection using a boosted cascade of simple features. In: Proceedings of the IEEE computer society conference on computer vision and pattern recognition, Kauai, HI, 8–14 December 2001, pp. 511–518. IEEE.

Wang Q and Zhang G (2017) Ore image edge detection using HOG-index dictionary learning approach. *IEEE Journal of Engineering* 9(1): 542–543.

World Health Organization (2018) *Global Status Report on Road Safety*. Geneva, Switzerland, World Health Organization, pp. 1–244.

Wu B, Kao C, Jen C, et al. (2014) A relative-discriminative-histogram-of-oriented-gradients-based particle filter approach to vehicle occlusion handling and tracking. *IEEE Transactions on Industrial Electronics* 61(8): 4228–4237.

Zhang D (2018) Vehicle target detection methods based on color fusion deformable part model. *J Wireless Com Network* 94(1): 1–6.

Zielke T, Brauckmann M and von Seelen W (1992) Intensity and edge-based symmetry detection applied to car-following. *Computer Vision Lecture Notes in Computer Science* 588(1): 865–873.

Author biographies

Altaf Alam obtained his bachelor's degree in electronics and communication from Integral university Lucknow (U.P) India in 2010 and his master's degree in control and instrumentation from Jamia Millia Islamia New Delhi India in 2014. He completed Ph.D. from Jamia Millia Islamia, New Delhi, India in 2020.

Zainul Abdin Jaffery is at present Professor and Head of the Department of Electrical Engineering, Jamia Millia Islamia (a Central university Govt. of India). He obtained his PhD degree from Jamia Millia Islamia (a central Govt. of India University) in 2004. Prof. Jaffery has published about 80 research papers in the area of Electronics and Electrical engineering in refereed international/national journals and conferences. His research area includes Digital signal processing, Digital image processing, and their applications in power engineering and electronics engineering. He is reviewer of many reputed journals. He is senior member of IEEE (USA).

Himanshu Sharma is presently working as Asst. Professor, in KIET Group of Institutions, Delhi-NCR, Ghaziabad, U.P., India. He is a PhD from Jamia Millia Islamia in the field of Machine Learning & Internet of Things. His M.Tech. & B.Tech. are from Dr. APJ Abdul kalam Technical University, Lucknow. His research Area includes, Machine Learning & Internet of Things.