An automated vision-based algorithm for out of context detection in images

R. Karthika*

Department of Electronics and Communication Engineering,
Amrita School of Engineering,
Amrita Vishwa Vidyapeetham,
Coimbatore 641112, India
Email: r_karthika@cb.amrita.edu
*Corresponding author

Latha Parameswaran

Department of Computer Science and Engineering,
Amrita School of Engineering,
Amrita Vishwa Vidyapeetham,
Coimbatore 641112, India
Email: p_latha@cb.amrita.edu

Abstract: Vehicular traffic on highways is a major concern relating to safety and security. Violation of traffic rules results in fatal incidents to a very large extent. In this work, an attempt has been made to detect violation of traffic rules namely vehicles in no parking and no stopping zones. Dataset consisting of cars in these zones has been used for experimentation. The proposed algorithm used histograms of oriented gradient (HOG) and Adaboost cascaded classifier for training. The traffic signs have been identified using Hough transform, Circlet transform and colour analysis. Experimental results are promising with an accuracy in the range of 90–97% with recognising no parking and no stopping sign.

Keywords: traffic sign; car detection: histograms of oriented gradient; HOG; circlet transform; Adaboost.

Reference to this paper should be made as follows: Karthika, R. and Parameswaran, L. (2018) 'An automated vision-based algorithm for out of context detection in images', *Int. J. Signal and Imaging Systems Engineering*, Vol. 11, No. 1, pp.1–8.

Biographical notes: R. Karthika is working as an Assistant Professor in the Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham. She is pursuing PhD in Computer Science Engineering at Amrita Vishwa Vidyapeetham. She received her BE in Electronics and Communication Engineering from the Madras University and MTech in Computer Vision and Image Processing from Amrita Vishwa Vidyapeetham. Her research interests include image processing, pattern recognition and computer vision.

Latha Parameswaran is working as a Professor in the Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham. She completed her Master's degree in Computer Applications from PSG College of Technology. Her PhD is from Bharatiyar University. Her areas of research include image processing, image mining, information security and theoretical computer science.

1 Introduction

Vehicular traffic has been growing during the past decade, demanding strict traffic rules and regulations. The traffic sign detection and recognition system enhance the safety by providing clear information to the drivers about parking, no parking, stop, speed limit etc. With cameras being deployed on every nook and corner of roads in cities, automatic detection of traffic rule violation will be useful. Object detection is a highly complex and a challenging task. In real world images, objects can form infinitely many combinations based on background variations, pose variations, background clutter, occlusions, illumination and intraclass variation.

In general, context refers to the information obtained from any source other than the appearance features of the object. Objects around us are very closely related to the scene. People easily recognise objects according to the scene information. It is proven that they find difficult to recognise objects, if objects are present in an unrelated environment. The objective of this work is to build a system that will detect traffic signs like no-parking and no-stopping signs in the presence of cars which is the contextual information. The purpose of such a system is to derive meaningful information whether a traffic rule is violated or not, which can be used in traffic surveillance systems.

An algorithm for detecting vehicles especially cars parked in a no parking, no stopping zones has been developed for a dataset created. The signs are specific to Indian traffic regulations. In this work, the following is the scenario considered for evaluation: the scene on a road where the vehicles are expected to be parked in parking zones and any vehicle (car) parked or stopped in front of No parking or a No stopping zone is referred to be out of context.

2 Literature survey

A number of different approaches are proposed to address the problem of object detection. Most object detection tasks involve the computation of keypoints as the first step. Keypoint detection is important for the descriptor (Harris and Stephens, 1988) invented a corner detector which was one of the robust keypoints detectors, based on the Eigenvalues to detect corner points. Scale invariant feature transform (SIFT) (Lowe, 2004), was a novel method to detect distinct invariant features which are used to perform matching under different viewpoints and different orientations. Viola and Jones (2001) proposed a machine learning based object detection with the integral image representation. The classifier is constructed by using a less number of unique features using Adaboost algorithm. To increase the speed of the detector and concentrate on the important regions of the image, cascade structure of the classifier is used. Dalal and Triggs (2005) implemented the histograms of oriented gradient (HOG) feature based linear support vector machine (SVM) classifier for pedestrian detection. Gabor filter based representation (Karthika and Parameswaran, 2016) for face detection is discussed. Han et al. (2006) proposed a system to detect people and vehicles which combines the stages of focus of attention generation using the depth map and SVM classifier. HOG can be used for the retrieval of the images (Sreelakshmi et al., 2016) segmenting the object from a cluttered scene. Agarwal and Maheshwari (2010) proposed a content-based image retrieval system using HOG feature descriptor and cooccurrence of Haar wavelet filters (Agarwal and Maheshwari, 2015).

Developing a traffic sign recognition system is a challenging task due to variable lighting condition, blurring effect, fading of traffic signs, partial obscured of traffic signs by trees and manmade objects. Wang et al. (2013) proposed a system capable of detecting violations, such as vehicle retrogress, running red lights and speeding in videos using background-updating techniques. Aliane et al. (2014) designed a system that recognises traffic signs from within the car, to provide the driver with warnings regarding traffic violations in the form of voice messages through speakers. A traffic sign is detected by integrating colour segmentation using RGB colour space, region detection and shape matching are discussed in Wali et al. (2015). The system identifies traffic signs in different lighting condition with the high speed. The traffic sign saliency region (TSSR) is constructed by using the saliency-based module in Xie et al. (2009). The gamma compression is applied as the pre-processing step for illumination in variation. In the top-down learning-based module, positive samples and negative samples are fed to the HOG to extract the feature vectors.

Ohgushi and Hamada (2009) bags of feature and SIFT are used for traffic sign recognition system. The region of interest is extracted based on colour information to reduce the SIFT computation cost. The Bag of features is classified by using the support vector machine. The comparison of three different approaches such as feature based, Adaboost based and SVM based are discussed in Li et al. (2010) for real-time traffic sign detection system. Among them Adaboost method achieved a high detection rate. Wang et al. (2012) proposed a traffic sign recognition system which is achieved by fuzzy C means clustering algorithm and content based image retrieval system.

While there are several feature-based approaches for traffic sign recognition, it is important to note that traffic signs contain vivid colours and standard geometric shape representation. In this work, No parking sign and No stopping sign which are having circular shape and line are taken. The circles can be detected using, the fast radial symmetry (FRS) detector which was first presented by Loy and Zelinsky (2003) which is an object detector based primarily on shape recognition technique. A qualitative evaluation of the detection results shows that the detector performs way better with sharp, non-shaky pictures, which makes complete sense as the detector works by looking for sharper gradients formed with a stronger contrast. Circlet transform (Chauris et al., 2011) used the image gradient and Fourier coefficients to detect the circle without the binary segmentation. Robust line detection started when Duda and Hart (1972) proposed an algorithm to detect the presence of lines and curves present in an image using angle-radius parameters as opposed to traditional slope-intercept parameters. Hatzidimos (2004) designed a traffic sign recognition system based on shape detection using Hough transform (Hough, 1962) and it was observed that the algorithm showed poor performance under low light conditions as the colour thresholds were set in RGB colour space rather than hue, saturation and value (HSV) colour space.

From the study of literature it may be observed that most authors have concentrated on objects present in a given image for the purpose of detection and retrieval; while this work focuses on identifying scenes which are out of context such as traffic rule violation.

3 Proposed work

The main objective of this work is to detect traffic signs such as no-parking and no-stopping signs along with cars. The first step is to detect and localise the presence of cars present in the image; the second step is to train the car detector using a cascaded classifier using Adaboost technique. Finally, recognise the presence of traffic signs present in the image using Hough Transform and Circlet transform.

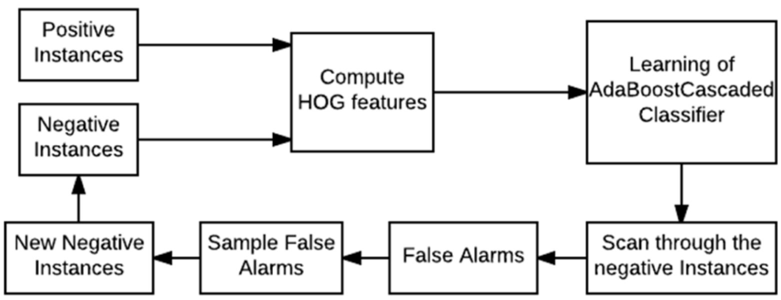
3.1 Detection of cars

The proposed algorithm involves the detection of car instances followed by recognition of traffic signs. The first step in the detection of cars is HOG feature extraction. HOG features have been extracted as they are robust to illumination variation and invariant to local geometric transformation (Tian et al., 2013). Extraction of HOG involves computing gradient magnitude and direction (Harris and Stephens, 1988), and Orientation binning. Each pixel present in a cell of an image casts a weighted vote based on its orientation value. Orientation bins are spread evenly over 0º to 180º into nine different histogram bins. The vote corresponding to each orientation value is either the gradient magnitude itself or is a function of gradient magnitude. After HOG feature extraction, Adaboost classifier eliminates irrelevant HOG features as a part of feature selection process. The implementation of such a classifier is similar to the Viola-Jones implementation (Viola and Jones, 2001) of boosted classifiers excepting that HOG features are used in the place of Haar features. The positive and negative samples for classifier training are taken from University of Illinois at Urbana-Champaign (UIUC) image database (Shivani and Aatif, 2004). HOG features for positive and negative instances are computed and fed to the first stage of Adaboost cascaded classifier. The negative instances for the next stage are generated by scanning through the negative instances and using the false positive classified instances for learning. Detection of cars follows the classifier training. In the detection of the target object, all the sub-windows of the input image are fed one by one. The sub-windows have to pass through all of the stages. Figure 1 shows the architecture of bootstrap algorithm.

The cascade of classifiers is used to increase detection rate with less computation time. Boosted classifiers are constructed by rejecting the majority of the negative subwindows and at the same time, it maintains that all positive samples are detected. The complete detection process is done using cascade structure. A positive result from the first classifier triggers the second classifier and it will trigger and evaluate the subsequent stages. A negative result leads to immediate rejection of the subwindow.

In the proposed work it is assumed that cars and no-parking sign/no-stopping sign cooccur. If either one of them is not present the scenario does not represent a traffic violation. Hence if the car detector does not detect the presence of a car, the algorithm will stop. Also after the detection of cars masking of all the instances of cars present in the image has been done to reduce computational complexity for traffic sign recognition.

Figure 1 Bootstrap training algorithm



3.2 Recognition of traffic signs

Methods that are based on colour detection techniques to classify traffic signs in an image take advantage of the fact that traffic signs are designed to be different from their surroundings. They have vivid colours and standard geometric shape to attract the driver. Figure 2 shows the architecture of this proposed system. The steps involving detection of traffic signs are as follows.

3.3 Circle detection

Edge detection is the first and foremost step for applying Hough transform. Canny edge detection has been applied to the input RGB colour image which has the car-localised portion masked. The masking is done carefully without masking the potential traffic signs which may be present around the detected cars. The circle detection is done using Circular Hough transform and circlet transform.

First Circular Hough transform is discussed here. It is applied after edge detection to match the circles present in the image. The minimum radius and maximum radius have been given as input parameters according to the size of the input image. The circle detector gives centre coordinates and radius of all the detected circles in the image.

The circle can be represented in parameter space. The equation of the circle is represented as

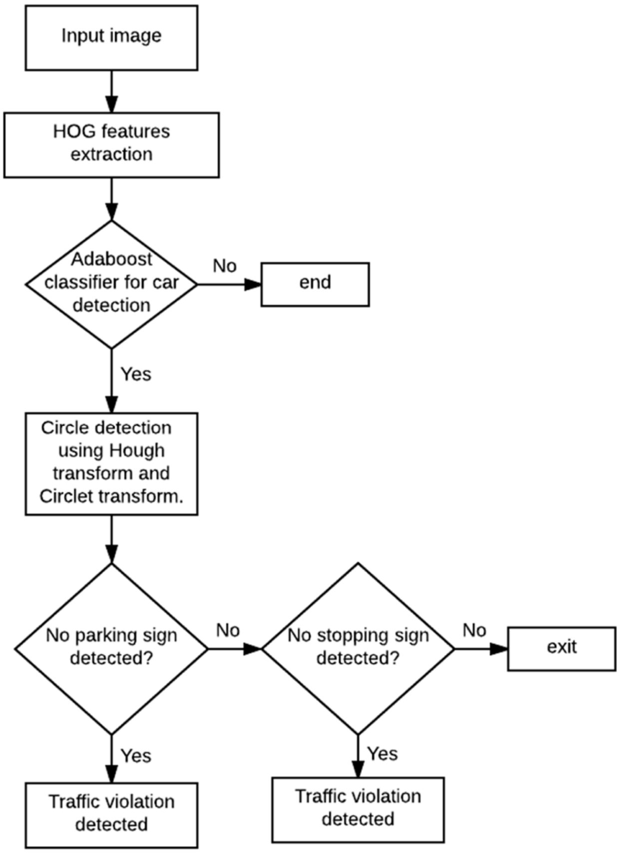
$$r^2 = (x-a)^2 + (y-b)^2 .$$

The parametric representation of the circle is

$$x = a + r \cos(\theta)$$
$$Y = b + r \sin(\theta)$$

a, b are the centre of the circle, r is the radius and θ is the angle.

Figure 2 Architecture of the proposed system



Circlet transform

Circlet transform has been designed to detect the circular shapes without the segmentation for robustness. Circlet Transform consists of decomposing the image circles with a range of radii and some particular width and series of Fast Fourier coefficients. This is called as circlet because it can be viewed as convolving a circle with a 1D wavelets to obtain to a 2D.

The circlet elements are found by a centre of circle (x_1, y_1) and f-certain width to the circlet and radius r. Based on the reference circlet $C(x,y)$ the circlet elements $C_{ele}(x,y)$ are computed by modifying the radius or the certain width of the circlet. The parameters σ (x_1, $y_{1,}$ r, f) fully characterise the circlet. The circlet function can be expressed as

$$C(x, y) = \Psi[2\pi f(r_1 - r)],$$

where $r_1 = \sqrt{(x - x_1)^2 + (y - y_1)^2}$ where Ψ is wavelet function. $C(x, y)$ is expressed as 2D Fourier domain for the practical approach. The image is decomposed into a sum of basic functions $C(x, y)$ which is similar to curvelet transform.

3.4 Applying colour thresholds

After obtaining the circles present in the image, the colour thresholds have been applied. Each circular disk is separated and the image is converted into an HSV image. In scenarios where colour content is important, an HSV model is preferred over the RGB colour space. HSV model describes the colour using Hue, Saturation and Value.

In that hue represents the colour type. Red colour and blue colour are then filtered using appropriate hue angle range. The output of a colour filter is a binary image having only the portion of the desired colour present inside the circular disk as white pixel and the rest of the image masked (black). The circular disks not containing the required amount of red and blue pixels are rejected. The number of pixels is decided according to the radius of the circular disk chosen.

3.5 Detecting lines

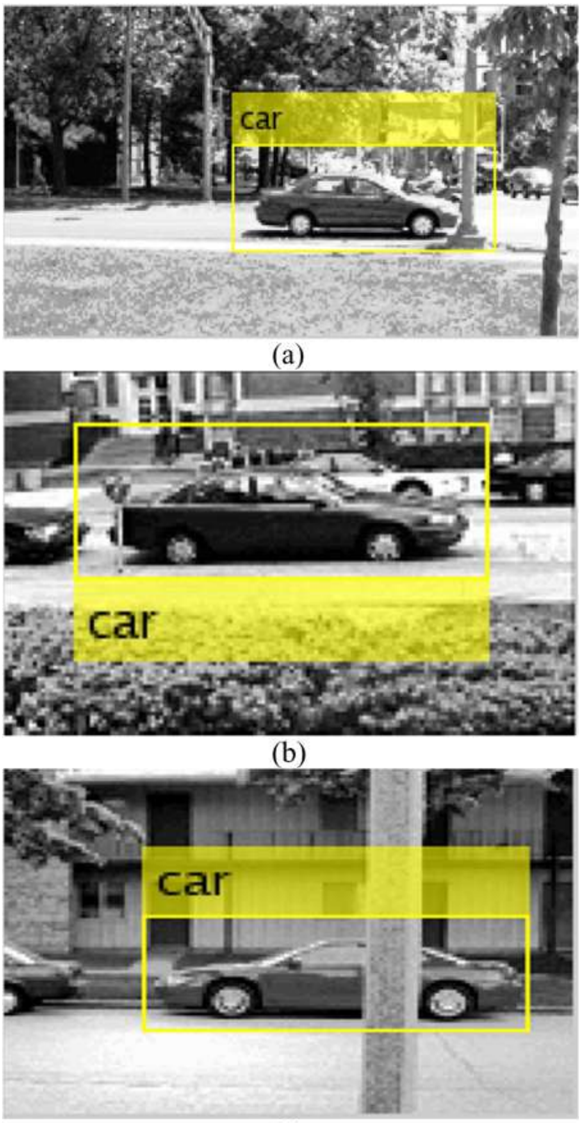
Since both no-parking sign and no-stopping sign contain slant red lines, Hough transform has been applied to fit the lines present inside the red filtered circular disk. The circular disks containing no lines are rejected. If there are lines present in the circle and the orientation of all the lines is only along 45º or –45º, the detected circle is classified as No-parking sign. If there are lines having orientations of 45º as well as –45º, the detected circle is classified as No-stopping sign.

4 Experimental results

A cascaded classifier using Adaboost and HOG features has been trained to detect cars present in an image. The training and testing samples were obtained from one of the standard car datasets, the UIUC image database (Shivani and Aatif, 2004). The dataset contains 550 positive samples and 500 negative samples. The test set I include 170 test images containing 200 cars in total (including images with multiple instances of cars). The test images were of same image sizes as in the training set.

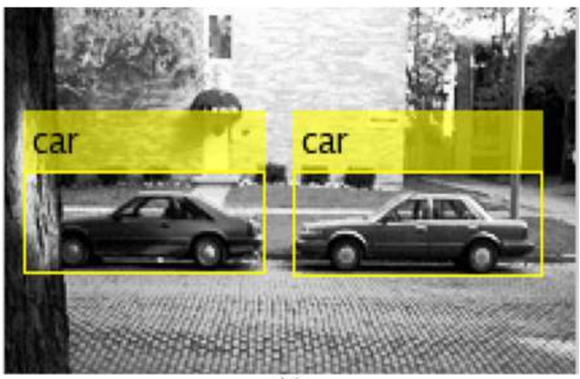
The test set II consists of 108 images containing 139 cars which have cars of different sizes ranging from 0.8 to 2 times the size of cars in the training images. The images are with different resolutions, multiple instances of cars in the image and the images with a different background. The detected images of the single car and multiple instances of the car from the testing dataset are shown in Figures 3 and 4. The performance of the car detector with single scale and multiscale is shown in Tables 1 and 2.

Figure 3 (a)–(c) Detection of single car in UIUC test dataset (see online version for colours)



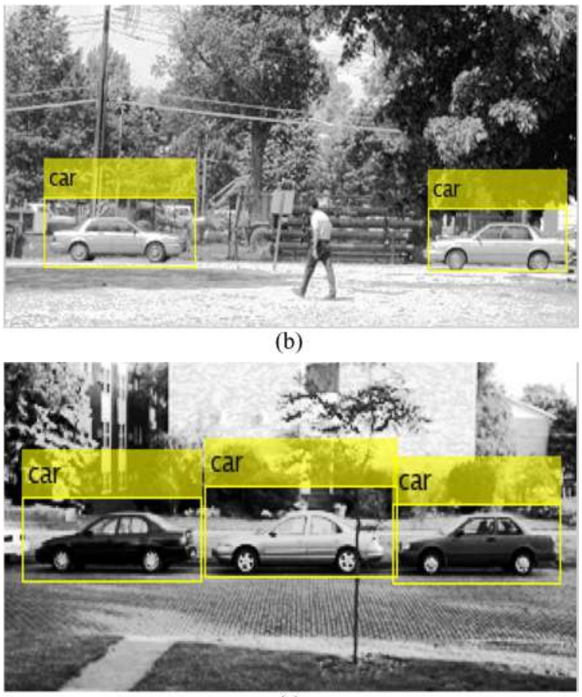
(c)

Figure 4 (a)–(c) Detection of multiple instances of cars in UIUC test dataset (see online version for colours)



(a)

Figure 4 (a)–(c) Detection of multiple instances of cars in UIUC test dataset (see online version for colours) (continued)



(c)

Table 1 Performance of single scale detection system on test Dataset I

Algorithm	*Number of true positive instances detected (TP)*	*Number of false positive instances detected (FP)*	*Recall R TP/200 (%)*	*Precision P TP/(TP + FP) (%)*
HOG	176	26	88	87.1
Neighbourhood suppression algorithm (Shivani and Aatif, 2004)	169	140	84.5	54.69

From Tables 1 and 2 the following observations are made:

In an object detection system, the objective is to know how many objects it detects and how often the detection false. The false positive rate is defined as the number of negative windows evaluated by the detection system. In single scale detector and multiscale detector (Shivani and Aatif, 2004) 100 × 40 windows was moved in steps of 5% window size. In neighbourhood suppression algorithm the relation features are eliminated. This leads to decrease in performance. It gives the information that additional information is captured by the relations. The interest points are found by using Forstner interest operator. After that image patches are highlighted then it determines the actual representation of an image. HOG captures edge or gradient patterns and invariance to illumination and shadowing. So it gives better results compared to the other algorithms.

Table 2 Performance of multiscale detection system using test Dataset II

Algorithm	*Number of true positive instances detected* (*TP*)	*Number of false positive instances detected* (*FP*)	*Recall R TP/139* (%)	*Precision TP/(TP + FP)* (%)
HOG	130	9	93.5	93.5
Neighbourhood suppression algorithm (Shivani and Aatif, 2004)	70	215	50.36	24.56
Repeated part elimination algorithm (Shivani and Aatif, 2004)	112	1216	80.58	8.43

For detecting traffic signs along with cars, a dataset was created and the performance was tested. There were a total of about 100 images of the same resolution. The images contained cars in the vicinity of No-parking sign, and cars in the vicinity of No-stopping sign. The No stopping sign and No parking sign used in the experiment is shown in Figures 5 and 6.The detected images indicating traffic violation under No parking sign and No stopping sign are shown in Figures 7 and 8.

Figure 5 No-stopping sign (see online version for colours)

Figure 6 No-parking sign (see online version for colours)

Figure 7 (a)–(c) Detection of traffic violation: cars parked under no-parking sign (see online version for colours)

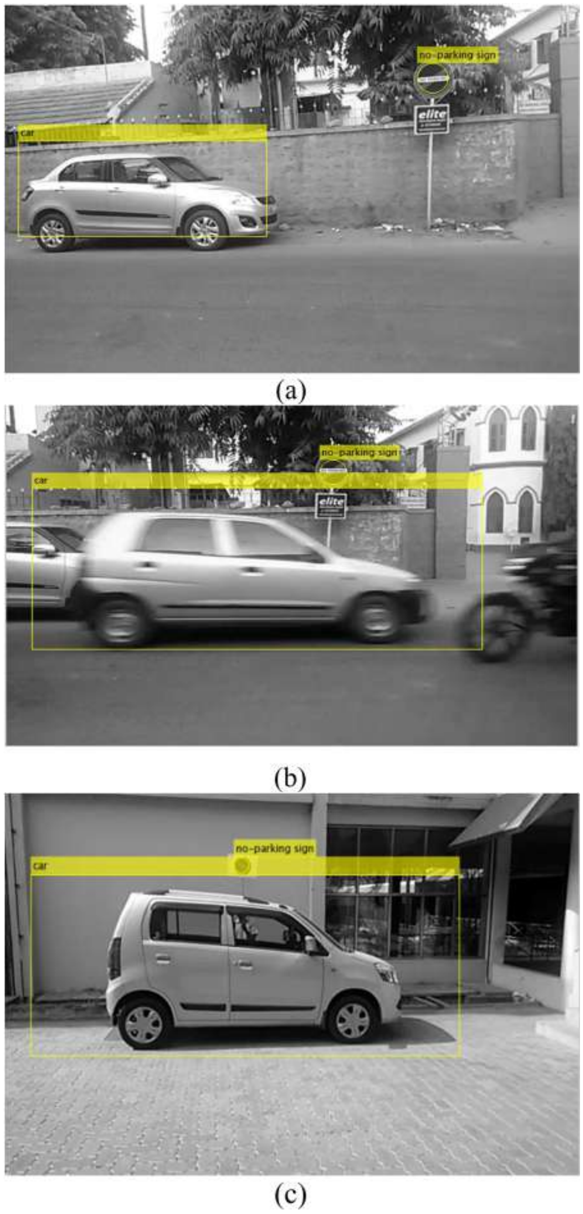


Figure 8 (a)–(c) Detection of traffic violation: cars parked under no-stopping sign (see online version for colours)

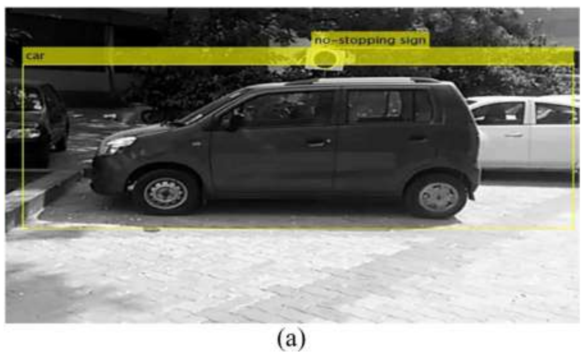
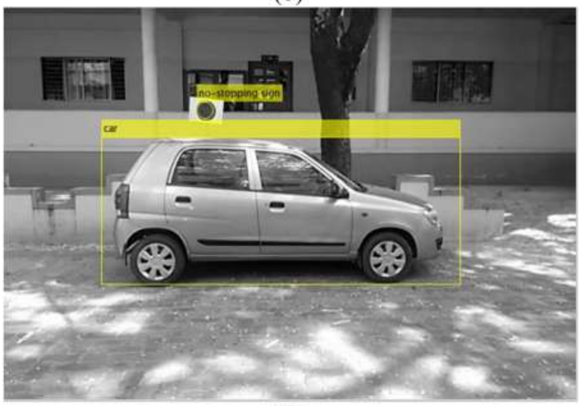


Figure 8 (a)–(c) Detection of traffic violation: cars parked under no-stopping sign (see online version for colours) (continued)

(b)



(c)

Table 3 Performance of traffic violation detection with no occlusion

Algorithm	*Class*	*Total positive instances*	*True positive*	*No prediction*
Hough transform	No-parking sign and cars (in percentage)	65 100%	59 90.76%	6 9.24%
	No-stopping sign and cars (in percentage)	35 100%	33 94.28%	2 5.72%
Circlet transform	No-parking sign and cars (in percentage)	65 100%	60 92.30%	5 7.7%
	No-stopping sign and cars (in percentage)	35 100%	34 97.14%	1 2.86%

From Table 3 the circlet transform is efficient because it is represented in the Fourier domain. By applying soft thresholding it detects a circle efficiently. Hough transform also perform well for the circle detection.

5 Conclusion

A system that can effectively detect specific traffic violations has been developed using object detection techniques. The algorithm used however is limited to identifying traffic violations only in parking areas. There are several other violations like No-entry sign, stop sign, Trucks not allowed the sign, etc., that can be detected. But these signs demand hard real-time object detection. The algorithm can be implemented in a real-time environment using suitable hardware. There is scope for improvement in various parameters like accuracy, occlusion handling, and background clutter. The speed of the detection process can be improved. The idea can be extended to detection of other types of vehicles including motorcycles.

References

Agarwal, M. and Maheshwari, R.P. (2010) ‘HOG feature and vocabulary tree for content-based image retrieval’, *International Journal of Signal and Imaging Systems Engineering*, Vol. 3, No. 4, pp.246–254.

Agarwal, M. and Maheshwari, R.P. (2015) ‘Co-occurrence of maximal Haar-like wavelet filters for CBIR’, *International Journal of Signal and Imaging Systems Engineering*, Vol. 8, No. 5, pp.316–330.

Aliane, N., Fernandez, J., Mata, M. and Bemposta, S. (2014) ‘A system for traffic violation detection’, *Sensors*, Vol. 14, pp.22113–22127.

Chauris, H., Karoui, I., Garreau, P., Wackernagel, H., Craneguy, P. and Bertino, L. (2011) ‘The circlet transform: a robust tool for detecting features with circular shapes’, *Computers and Geosciences*, Vol. 37, No. 3, pp.331–342.

Dalal, N. and Triggs, B. (2005) ‘Histograms of oriented gradients for human detection’, *Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005*, IEEE, IEEE Computer Society, June, Vol. 1, pp.886–893.

Duda, R.O. and Hart, P.E. (1972) ‘Use of the Hough transformation to detect lines and curves in pictures’, *Communications of the ACM*, Vol. 15, pp.11–15.

Han, F., Shan, Y., Cekander, R., Sawhney, H. S. and Kumar, R. (2006) ‘A two-stage approach to people and vehicle detection with hog-based SVM’, *Performance Metrics for Intelligent Systems 2006 Workshop*, August, pp.133–140.

Harris, C. and Stephens, M. (1988) ‘A combined corner and edge detector’, *Alvey Vision Conference*, pp.147–151.

Hatzidimos, J. (2004) ‘Automatic traffic sign recognition in digital images’, *Proceedings of the International Conference on Theory and Applications of Mathematics and Informatics-ICTAMI*, Thessaloniki, Greece, April, pp.174–184.

Hough, P.V.C. (1962) *Method and Means for Recognizing Complex Patterns*, U.S. Patent 3 069 654.

Karthika, R. and Parameswaran, L. (2016) ‘Study of gabor wavelet for face recognition invariant to pose and orientation’, *Proceedings of the International Conference on Soft Computing Systems, Advances in Intelligent Systems and Computing*, Vol. 397, pp.501–509.

Li, Y., Pankanti, S. and Guan, W. (2010) ‘Real-time traffic sign detection: an evaluation study’, *20th International Conference on Pattern Recognition* (*ICPR), 2010*, IEEE, August, pp.3033–3036.

Lowe, D.G. (2004)‘Distinctive image features from scale-invariant key points’, *International Journal of Computer Vision*, Vol.60, No.2, pp.91–110.

Loy, G. and Zelinsky, A. (2003) Fast radial symmetry for detecting points of interest', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. *25*, No. 8, pp.959–973.

Ohgushi, K. and Hamada, N. (2009) Traffic sign recognition by bags of features', *10th Conference on TENCON 2009-2009 IEEE Region*, IEEE, January, pp.1–6.

Shivani, A. and Aatif, A. (2004) 'Learning to detect objects in images via a sparse, part-based representation', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, pp.1475–1490.

Sreelakshmi, S., Vijai, A. and Senthilkumar, T. (2016) 'Detection and segmentation of cluttered objects from texture cluttered scene', *Advances in Intelligent Systems and Computing*, Vol. 398, pp.249–257.

Tian, S., Lu, S., Su, B. and Tan, C.L. (2013) 'Scene text recognition using co-occurrence of histogram of oriented gradients', *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*, IEEE, August, pp.912–916.

Viola, P. and Jones, M. (2001) 'Rapid object detection using a boosted cascade of simple features', *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* Vol. 1, pp.1–511.

Wali, S.B., Hannan, M.A., Abdullah, S., Hussain, A. and Samad, S.A. (2015) 'Shape matching and color segmentation based traffic sign detection system', *Threshold*, Vol. 90, p.255.

Wang, W., Wei, C.H., Zhang, L. and Wang, X. (2012) 'Traffic-signs recognition system based on multi-features', *2012 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications* (*CIMSA*), IEEE, July, pp.120–123.

Wang, X., Meng, L.M., Zhang, B., Lu, J. and Du, K.L. (2013) 'A video-based traffic violation detection system', *Proceedings 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC)*, IEEE, December, pp.1191–1194.

Xie, Y., Liu, L.F., Li, C.H. and Qu, Y.Y. (2009) 'Unifying visual saliency with HOG feature learning for traffic sign detection', *Intelligent Vehicles Symposium, 2009 IEEE*, IEEE, June, pp.24–29.