2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI 2017)

# An Efficient Automatic Traffic Sign Detection and Recognition Method for Smartphones

Po-Cheng Shih, Chi-Yi Tsai* and Chun-Fei Hsu
Department of Electrical and Computer Engineering, Tamkang University,
151 Ying-chuan Road, Danshui District,
New Taipei City 251, Taiwan R.O.C.
*E-mail: chiyi_tsai@mail.tku.edu.tw

***Abstract*—The usage of automatic driving assistance systems (ADAS) has become more and more popular in recent years. In the design of ADAS, traffic sign detection (TSD) and traffic sign recognition (TSR) are two important functions and have been widely studied in the literature. This paper addresses the design of a vision-based TSD and recognition (TSDR) system, which is computationally efficient and can run on a common smartphone with real time performance. To achieve this, a novel TSD algorithm is proposed based on the Maximally Stable Extremal Regions (MSER) algorithm to accurately detect all traffic sign candidates in real time. Then, the feature vector of each candidate region is extracted via the Histogram of Oriented Gradient (HOG) algorithm. To recognize the traffic sign, the proposed TSR algorithm is designed by combining a Linear Support Vector Machine (LSVM) classifier with a voting process to improve the traffic sign recognition rate in real-world environments. The proposed TSDR system had been implemented on iOS embedded platform and can operate at an average speed of 30 fps for processing 640 × 480 video streams. Moreover, experimental results show that the proposed system achieves about 96% precision rate for recognition on the German Traffic Sign Recognition Benchmark (GTSRB). Therefore, the proposed TSDR algorithm has a potential to be used in realistic products.**

***Keywords- ADAS; MSER; HOG; LSVM; Traffic Sign Detection; Traffic Sign Recognition***

## I. INTRODUCTION

Helping drivers to avoid car accident is an essential task of the design of intelligent ADAS. Because many drivers sometimes do not pay attention to traffic signs standing on the roadside, the development of accurate TSD and TSR functions in ADAS becomes a rapidly growing demand. There are several distinguishing methods that can be used to detect and recognize traffic signs. These methods usually are designed in some specific colors and shapes, with the high contrast line enclosed the boundaries. This feature selection is benefit for the TSD and TSR design. Moreover, the effect of geometric and rotational distortion is usually limited because traffic signs are generally upright and facing the camera.

Sorting by shapes and colors, traffic signs can be separated in different groups. However, there are a lot of external interferences that obstruct TSDR. The common external interferences include fog, raining, and sandstorm, etc. In addition, some traffic signs are not well maintained so that they are really dirty, and some are rusty. These make TSDR becoming a difficult task in the real work environment.

For the past few years, the explosive growth of the smartphones has greatly changed people's life. Smartphones are indispensable to people in the 21st century, which motivates us to design and implement an efficient TSDR algorithm for smartphones. To achieve this, this paper presents a computationally efficient TSDR method that is suitable to be implemented on smartphones for drivers to easily obtain the information of traffic signs appeared in front of the vehicle. The proposed TSDR method also consists of a TSD and a TSR stage. In the TSD stage, some existing systems use shape-based and color-based sign detection algorithms to extract traffic signs in the road image. In [1], Paclîk et al. proposed a color-based TSD method that converts the Red, Blue, Green (RGB) image to Hue, Saturation, Intensity (HSI) color space in order to robustly find out traffic signs in the road image. In [2], Kuo and Lin also converted the image from RGB to HSI and extracted the red traffic signs by image thresholding approach. In [3], Estevez and Kehtarnavaz employed a color segmentation method for detecting a small subset of traffic signs which contain red components. In [4], Loy used fast radial symmetry transform to design a general regular polygon detector for extracting traffic signs in the image. In [5], Paulo and Correia applied the Harris corner detector to detect triangular and rectangular signs. In [6], Gavrila implemented a shape detection method using a distance-transform based template matching algorithm.

In the TSR stage, there are a variety of systems to recognize traffic signs. In [7], Sermanet et al. used a set of multiscale Convolutional Neural Networks (CNNs) for TSR. In [8], Wang et al. used a set of SVMs to compose a hierarchical classification structure. In [9], Stallkamp et al. used linear discriminant analysis approach to build the classifier for recognizing traffic signs. In [10], Zaklouta et al. used k-dimensional (k-d) tree and random forests to build the TSR classifier. However, there are only a few papers which address the implementation of the TSDR algorithm on smartphones [11, 12].

The rest of this paper is organized as follows. In Section 2, we present the system overview and detail operations of the proposed TSDR method. Experiment results are provided in Section 3, and we make a conclusion and future works in Section 4.

This work was supported by the Ministry of Science and Technology of Taiwan, ROC under grant MOST 106-2221-E-032-006, 106-2813-C-032-009-E and 105-2628-E-032-001-MY3.

978-1-5386-1937-7/17/$31.00 ©2017 IEEE

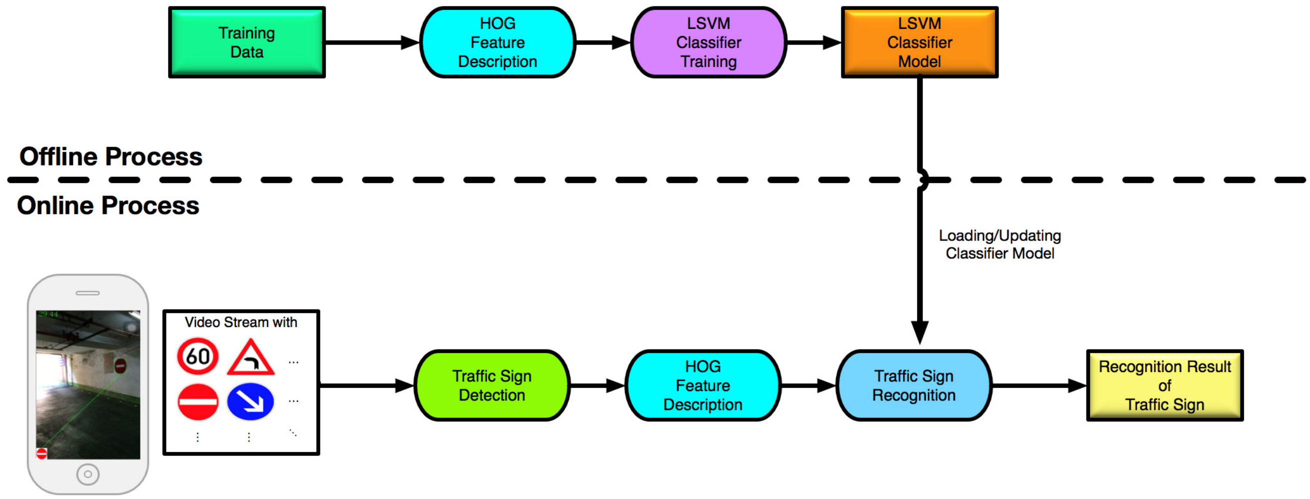


Figure 1. Block diagram of the proposed TSDR method.

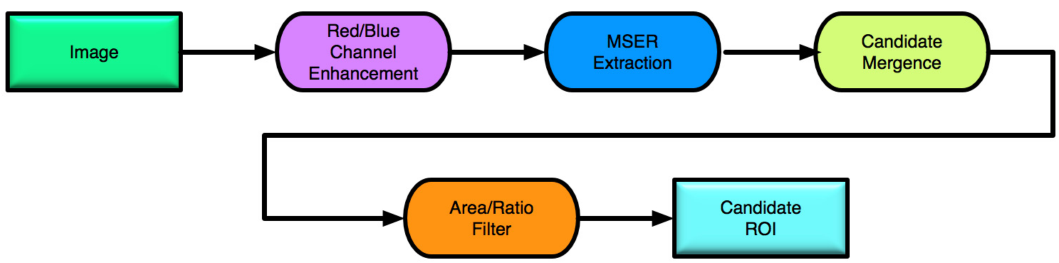


Figure 2. Detail operations of the proposed TSD method.

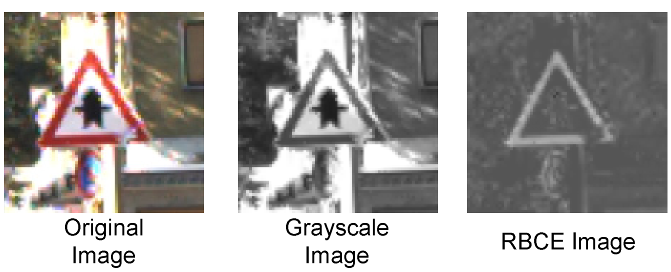


Figure 3. Comparison of the original image, the grayscale image and the RBCE image.

## II. The Proposed Traffic Sign Detection and Recognition System

### A. System Overview

As shown in Figure 1, the proposed TSDR system is divided into two processes: an offline process for training a LSVM classifier model [13] and an online process for performing TSDR task. We focus on the design of the proposed TSDR method in this section.

### B. Detection of Traffic Signs

Detection is the first priority of the TSDR system. Figure 2 shows the whole operations of the proposed TSD method. In general, the input RGB image is firstly converted to the grayscale image. Then, the grayscale image is binarized at several threshold levels to extract the MSER candidates [14] with bounding boxes. However, we empirically found that a lot of MSER candidates are obtained when applying the MSER extraction process on the grayscale image. This is mainly caused by that the grayscale image contains a lot of background textures. To overcome this issue, we employed the Red/Blue Channel Enhancement (RBCE) method proposed in [15] to suppress background textures while emphasizing red/blue colors. Firstly, we compute the maximum value between red value and blue value in each pixel. Secondly, the maximum value is divided by the sum of red, green and blue value in the pixel. Finally, the value is multiplied by 255 to produce the RBCE image. In summary, the following is the formula of RBCE method:

$$\sigma_{RB} = \frac{\max(R,B)}{R+G+B} \times 255, \qquad (1)$$

where the value of   is the enhancement of red and blue colors with suppression of other colors. Figure 3 shows a comparison of the original image, the grayscale image and the RBCE image.

Next, the MSER extraction process is applied on the RBCE image to find out the stable regions which may contain traffic signs. As a result, the candidate bounding boxes are not so

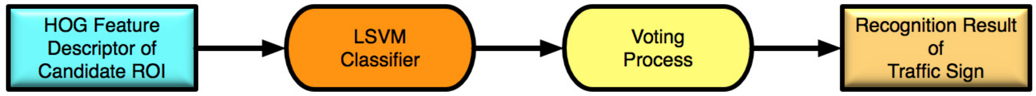


Figure 4. Detail operations of the proposed TSR method.

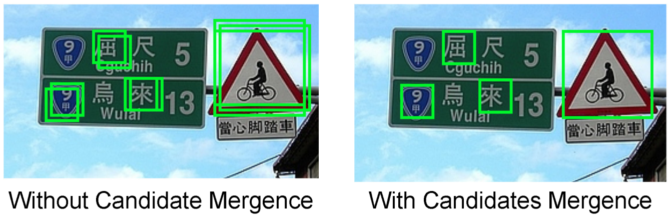


Figure 5. Comparison of candidate extraction results obtained with and without the candidate mergence processing.

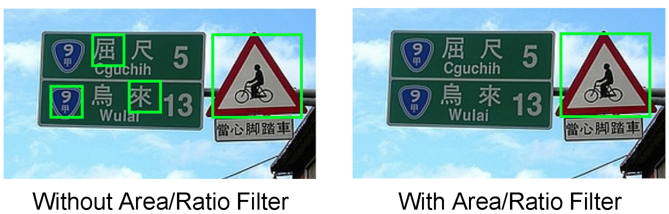


Figure 6. Comparison of candidate extraction results obtained with and without the area/ratio filter.

sensitive to the background textures. Although the input color image has been processed to enhance red and blue colors, there are still some candidate regions overlapped to each other. In order to decrease the number of candidates, a candidate mergence method is proposed, which finds the two overlapped candidates in the image and groups them together to form a new candidate. Figure 5 shows the comparison of candidate extraction results obtained before and after the candidate mergence processing.

Although the number of candidates has been drastically reduced by the RBCE and candidate mergence methods, there are only a few candidates that we are not interested. To further exclude these non-candidates, some geometric features of traffic signs, such as height, width and shape, can be used. By observing these features, we deign a filter called "Area/Ratio Filter", which aims to filter out the candidates that do not satisfy the conditions of area size and aspect ratio of the bounding box. Figure 6 shows the comparison of candidate extraction results obtained before and after the area/ratio filter processing. By filtering a specific aspect ratio and area size of candidates, we then get the final regions of interest (ROI) in the image.

### C. *Recognition of Traffic Signs*

The TSR module aims to classify each detected traffic sign correctly. Figure 4 shows the whole operations of the proposed TSR method. For the recognition of each ROI detected in the image, the HOG feature description method [16] is then applied to each detected ROI in the image. The HOG feature descriptor is computed by dividing an image patch into several cells, and the HOG descriptor is then obtained by stitching all histogram of gradient orientations in every single cell together.

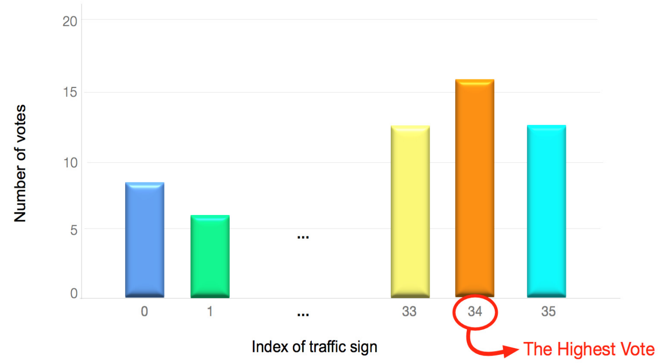


Figure 7. Voting histogram of the recognition results.

After extracting the HOG feature descriptor from each detected ROI, a LSVM classifier is employed to recognize the traffic sign in the ROI. For training the LSVM classifier model, an online training dataset of German Traffic Sign Recognition Benchmark (GTSRB) [17] was used in this work. The GTSRB dataset is suitable for training the LSVM classifier model because it contains a variety of traffic signs and sorts them in a clear way. Although the GTSRB contains many kinds of traffic signs, the number of training data in every single category is still insufficient for training LSVM classifier model. To increase the number of training samples, a data augmentation algorithm is implemented to generate synthetic training data, which is inspired from the deep learning and machine learning approaches.

However, when the trained LSVM classifier was applied on the road image captured in the real-world environment, we found that the recognition result usually is not very stable in every single frame having the same traffic signs. This issue deteriorates the recognition performance of the proposed TSR method in practice. In order to overcome this problem, a recognition voting approach is also proposed to stabilize the recognition result in a sequence of road images. In other words, we take the recognition process as an election task, in which the number of classes is the number of candidates and the vote is the result obtained from a single process of the LSVM classification. By doing so, we will get a voting histogram that presents the traffic sign voting statistics in the road image sequence. According to the statistics, the final recognition result can be determined as the traffic sign that gets the highest number of votes in the histogram. Figure 7 shows the concept of recognition voting approach implemented on the proposed TSR method.

Figure 8. Comparison of the real (left) and the augmentation (right) data.

### *D. Synthetic Training Data Generation*

To prevent misclassification of unknown signs, it is very important to train the LSVM classifier model on all possible traffic signs. However, it is also difficult and inefficient to gather a sufficient amount of real data for training the LSVM classifier model. Although there are a lot of training data in GTSRB [17], it is insufficient for training a robust and stable LSVM classifier model used in real world. The proposed solution to this problem is to use a data augmentation algorithm which is inspired from the deep learning and machine learning approaches.

There are a variety of data augmentation methods, including average blur, Gaussian blur, Gaussian noise, image shift, image rotation and image contrast etc. Similar to these methods, the proposed algorithm synthetically generates variations and distortions of traffic sign images to create various forms of training samples to improve the robustness of the LSVM classifier model. To achieve this, we took the GTSRB database as a base, and then applied the data augmentation algorithm to the GTSRB database. Each image in the GRSRB is able to generate a large number of synthetic images for training the LSVM classifier model. Figure 8 shows the comparison between the real and the augmentation data. With the help of the data augmentation algorithm, we can avoid tiring manual hand labelling for collecting training data and get more reliable recognition results after TSR.

## III. EXPERIMENTAL RESULTS

The proposed TSDR method was implemented on the Apple iPhone 6s smartphone. Table I shows the specification of the iPhone 6s smartphone, which is with screen resolution at 640 x 480 pixels. When implemented, the proposed TSDR method can run about 30 frames per second on the iPhone 6s smartphone. Figure 9 shows the experimental results of the proposed TSDR method tested in the real world environment. The left image was taken in an underground parking lot lacking of light, and the right image was taken in a street at a daytime with sufficient light. The experimental results show that the proposed method detects and recognizes the traffic sign successfully.

Although we apply the area/ratio filter method in TSD, there still are some backgrounds whose features are similar to traffic signs. Figure 10 shows the experimental results of recognizing non-traffic-sign candidate ROI. The experimental result shows that the proposed method successfully detects and recognizes both the traffic and non-traffic signs in the road image.

Because there are a variety of traffic signs in the real world environment, it is very difficult to cover all types of the traffic signs. Figure 11 shows a failure case of the proposed method to recognize the traffic sign which is not included in the training data. In Figure 11, it is clear that the ROI candidates of both traffic signs are detected successfully by the proposed TSD method; however one of the detected traffic sign is not recognize correctly. This is because the incorrect recognized traffic sign is not included in our training database. Moreover, traffic signs in Western are different from those in Eastern countries. For example, Figure 12 shows two cases of traffic

TABLE I. IPHONE 6S SPECIFICATION

| | | |
|---|---|---|
| Platform | OS | iOS 9.3.1 |
| | Chipset | Apple A9 |
| | CPU | Dual-core 1.84 GHz Twister |
| | GPU | PowerVR GT7600 |
| Memory | Internal | 64GB ROM with 2GB RAM |
| Camera | Primary | 12 MP, f/2.2, 29mm |


Figure 9. Experimental results of the proposed TSDR method implemented on the iPhone 6s smartphone.

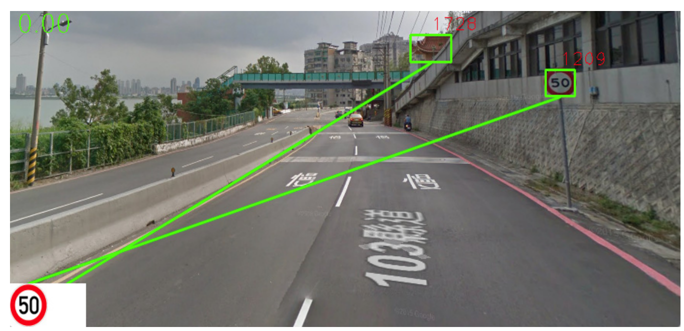


Figure 10. Experimental result of recognizing non-traffic-sign candidate ROI.

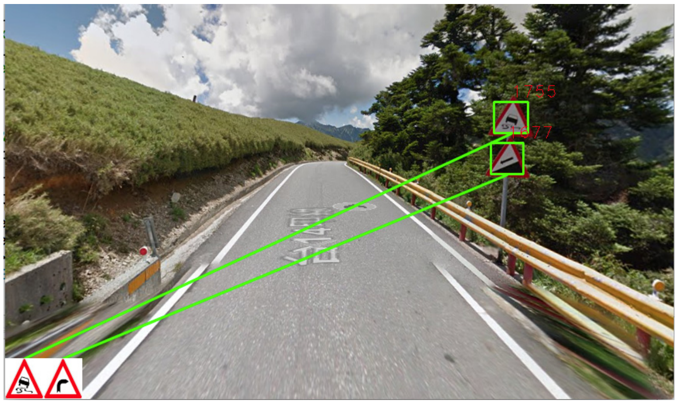

Figure 11. The failure case of recognizing the traffic sign which is not included in the training data.


Figure 12. Two traffic signs different in Western and Eastern countries.

signs which are different in Western and Eastern countries.

In the current design, the proposed TSDR method can recognize 35 different types of traffic signs and takes about 28.73 ms in the TSD process and about 0.25 ms in the TSR process. We also test the proposed TSDR method using the online dataset German Traffic Sign Detection Benchmark (GTSDB) [18] and GTSRB to evaluate the recognition performance of the proposed method. We obtain 94.48% in recall and 96.12% in precision, respectively.

## IV. Conclusions and Future Work

In this paper, a real-time TSDR algorithm is designed and developed. The proposed method can achieve real time performance when implemented on a common iPhone 6s smartphone. The proposed method uses MSER extraction to find appropriate candidates in the RBCE image. To get more appropriate ROI, the candidate mergence and area/ratio filter methods are proposed. We apply HOG feature description and LSVM classification to recognize traffic signs. Finally, we stabilize the recognition result by a voting histogram. Experimental results validate the proposed TSDR method in the real world environment.

In the future, we will increase the types of traffic signs in our training dataset to recognize more types of traffic signs. In addition, cloud services [19] are becoming more and more popular in recent years. It is also an important research issue to extend the proposed TSDR method to cooperate with a cloud server. By doing so, the position information and the recognized traffic signs will be sent to the cloud server, which makes the location-aware traffic sign recognition possibly to be realized in practice.

## References


[1] P. Paclîk, J. Novovicõvá, P. Pudil, and P. Somol, "Road sign classification using laplace kernel classifier," Pattern Recognition Letters, vol. 21, no. 13-14, pp. 1165-1173, 2000.

[2] W.-J. Kuo and C.-C. Lin, "Two-stage road sign detection and recognition," Multimedia and Expo, 2007 IEEE International Conference on, pp. 1427-1430, July 2007.

[3] L. Estevez and N. Kehtarnavaz, "A real-time histographic approach to road sign recognition," Image Analysis and Interpretation, 1996., Proceedings of the IEEE Southwest Symposium on, pp. 95-100, Apr 1996.

[4] G. Loy, "Fast shape-based road sign detection for a driver assistance system," in IEEE/RSJ International Conference on Intelligent Robots and Systems, 2004, pp. 70-75.

[5] C.Paulo and P.Correia, "Automaticdetectionandclassificationoftraffic signs," in Image Analysis for Multimedia Interactive Services, 2007. WIAMIS '07. Eighth International Workshop on, June 2007.

[6] D. Gavrila, "Traffic sign recognition revisited," in DAGM-Symposium, 1999, pp. 86-93.

[7] P. Sermanet and Y. LeCun, "Traffic sign recognition with multi-scale convolutional networks," in Proceeding IEEE International Joint Conference Neural Network, San Jose, CA, USA, 2011, pp. 2809-2813.

[8] G. Wang, G. Ren, Z. Wu, Y. Zhao, and L. Jiang, "A hierarchical method for traffic sign classification with support vector machines," in Proceeding IEEE International Joint Conference Neural Network, Dallas, TX, USA, 2013, pp. 1-6.

[9] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "Man vs computer: Benchmarking machine learning algorithms for traffic sign recognition," in Proceeding IEEE International Joint Conference Neural Network, San Jose, CA, USA, 2011, pp. 323-332.

[10] F. Zaklouta, B. Stanciulescu, and O. Hamdoun, "Traffic sign classification using K-d trees and random forests," in Proc. IEEE International Joint Conference Neural Network, San Diego, CA, USA, 2011, pp. 2151-2155.

[11] C.-Y. Tsai, H.-C. Liao and Y.-C. Feng, "A Novel Translation, Rotation, and Scale-Invariant Shape Description Method for Real-Time Speed-Limit Sign Recognition," *IEEE International Conference on Advanced Materials for Science and Engineering*, Tainan, Taiwan, November 12-13, 2016, pp. 486-488.

[12] C.-Y. Tsai, H.-C. Liao and K.-J. Hsu, "Real-time embedded implementation of robust speed-limit sign recognition using a novel centroid-to-contour description method," *IET Computer Vision*, vol. 11, no. 6, pp. 407-414, 2017.

[13] C. G. Kiran, L. V. Prabhu, R. V. Abdu, and K. Rajeev, "Traffic sign detection and pattern recognition using support vector machine," in International Conference on Advances in Pattern Recognition, 2009, pp. 87-90.

[14] J. Matas, "Robust wide-baseline stereo from maximally stable extremal regions," International Vision Conference, vol. 22, no. 10, pp. 761-767, 2004.

[15] J. Greenhalgh and M. Mirmehdi, "Traffic sign recognition using MSER and random forest," in European Signal Processing Conference, 2012, pp. 1935-1939.

[16] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Computer Vision and Pattern Recognition, 2005, pp. 886-893.

[17] German Traffic Sign Recognition Benchmark, Traffic Signs Image Database. Available : http://benchmark.ini.rub.de/?section=gtsrb&subsection=dataset.

[18] German Traffic Sign Detection Benchmark, Traffic Signs Image Database. Available : http://benchmark.ini.rub.de/?section=gtsdb&subsection=dataset.

[19] Cloud Services. Available: https://www.comprompt.co.in/services/cloud-services/