Traffic Sign Detection and Recognition Methods, Review, Analysis and Perspectives

Kanan Miniyar

July 6, 2022

Traffic Sign Detection and Recognition Methods, Review, Analysis and Perspectives

1. ABSTRACT

Road Sign Recognition is a branch of applied computer vision research that deals with the automatic detection and classification of traffic signs in images of traffic scenes. The purpose of this review paper is to look after the various classification techniques that can be used to build a system that recognizes road signs in images. The primary goal is to study literature available on the road traffic signs detection and recognition that can identify various types of road signs from static digital images in a reasonable amount of time. In this review paper, we will look at various learning systems for classification that are based on prior knowledge.

A road sign recognition system must solve a classic pattern recognition problem: distinguishing between different road signs. Furthermore, the location of the road sign in the image is unknown. Once these hurdles are cleared, such a system could be integrated into a Smart Driver system.

2. INTRODUCTION

Traffic Sign Detection and Recognition (TSD&R) is used in traffic situations to control traffic signs, alert drivers, and mandate or prohibit particular activities. Fast Real-Time and robust traffic sign detection and recognition can assist and relieve the driver, increasing driving safety and comfort dramatically. An automatic road sign recognition system could be useful in warning drivers to dangerous road conditions, making driving safer. In general, traffic signs convey a variety of information to drivers in order to ensure safe and effective travel. As a result, automatic traffic sign recognition is critical for autonomous intelligent driving vehicles or driver assistance systems [1]. However, identifying traffic signs against a variety of natural background viewing situations remains a difficult task. An autonomous road sign recognition system should be able to detect and identify a group of road signs in images. A system like this should be able to assess the camera's acquired road scene image, extract the road sign region, and make intelligent decisions. It must also properly inform the vehicle to the approaching road sign. Automatically recognizing road signs is a difficult challenge. There are a number of important issues that need to be taken into consideration. These include illumination conditions, direction of sign's face, status of paint on signs, placement of multiple signs near each other, torn and tilted signs, variation in sign's scale, obstacles such as tree, image sensor's properties, car vibrations etc. The systems usually have been developed into two specific phases:

First Phase: It is normally related to the detection of traffic signs in an image by using image processing.

Second Phase: It is related to the recognition of those detected signs, which deals with the interest of performance in an artificial neural network.

The detection system's efficiency and speed are crucial. Various methods for automatic traffic sign detection have been developed with promising results in order to recognize traffic signs. Neural networks accurately portray a traffic sign recognition technique.

2.1 Traffic Signs

Traffic signs are placed along the roads with the function of informing drivers about the front road conditions, directions, restrictions or text information. Though traffic signs have different structures and appearances in different countries, the most essential types of traffic signs are prohibitory, danger, mandatory and text-based signs. The prohibitory, danger or mandatory signs often have standard shapes, such as circle, triangle and rectangle, and often have standard colors such as red, blue and yellow. The text-based signs usually do not have fixed shapes and contain informative text.

(a)

(b)

(c)

Figure 2.1 Different types of traffic signs from Germany, China and America.

(a) German signs, (b) Chinese signs, (c) American signs.

Signs from Germany and China are classified into prohibitory signs, danger signs, mandatory signs and other types of signs. American signs are classified into regulatory signs, warning signs, guide signs and other signs according to Wikipedia. More signs from these three countries can be found in German GTSDB dataset [4], Chinese TT100K dataset [5], and American LISA dataset [6].

In Figure 2.1, some types of German signs, Chinese signs, and American signs. Signs from Germany and China are classified into prohibitory signs, danger signs, mandatory signs and other types of signs have been listed. American signs are classified into regulatory signs, warning signs, guide signs and other signs. More signs from these three countries can be found in German GTSDB dataset [4], Chinese TT100K dataset [5], and American LISA dataset [6].

2.2 Traffic signs for human driving safety

Though traffic signs play an important role in traffic safety and regulating drivers' behavior, they are often unattended. In the study of [7], Costa et al. show that different types of signs have different ability to capture the attention of drivers. During gazing, the drivers may not remember the content of a sign or may miss some other important signs. During driving, traffic signs with different distances and different presentation times have different influences on the accuracy of sign identification for human drivers [7]; the study in [7] shows the drivers have 75% accuracy with less than 35ms presentation time and 100% accuracy with 130ms presentation time; this study also shows the drivers need enough time to correctly recognize the signs in front. According to [8], the sign context and drivers' age have effect on traffic sign comprehension; their experiments show that younger drivers perform better than older drivers on both accuracy and response time, and that the sign context increase the comprehension time.

2.3 Challenges involved in the traffic sign detection and recognition

The challenges involved in the detection and recognition of traffic signs are,

- Illumination variations with the time of the day and due to the weather conditions
- The paint present on the signs fades with the time
- Geometric distortions due to storms or wind
- Variation in scale as the vehicle approaches the traffic sign
- Other similar shapes in the image
- Occlusion by the trees, street lamps, buildings, pedestrians etc.
- Characteristics of the image acquisition system
- Detection may suffer due to vibration and motion blur

3. Traffic Sign Detection

The Traffic Sign Detection stage entails locating the region in the image or video frame that contains the traffic signs. Because traffic signs are made of specific colours and shapes, they can be identified using colour and shape information. This section provides a brief overview of existing detection methods such as color-based detection, shape-based detection, color-and-shape-based detection, and other approaches.

A. Colour Based Detection

Colour-based detection techniques use colour as the primary feature to identify the region of the image or video frame containing the traffic sign. Colour segmentation is a popular method for colour-based detection. The colour segmentation technique eliminates unnecessary objects, reducing the search area of the image or video frame.

The authors in [9] proposed red colour traffic sign detection in video using colour as the main feature. The original video frame is converted into gray scale and only the red channel is extracted from the original frame. The gray scale frame is subtracted from the red component of the original frame and then the subtracted frame is converted to binary by applying the threshold. Morphological operations are performed to remove the background noise.

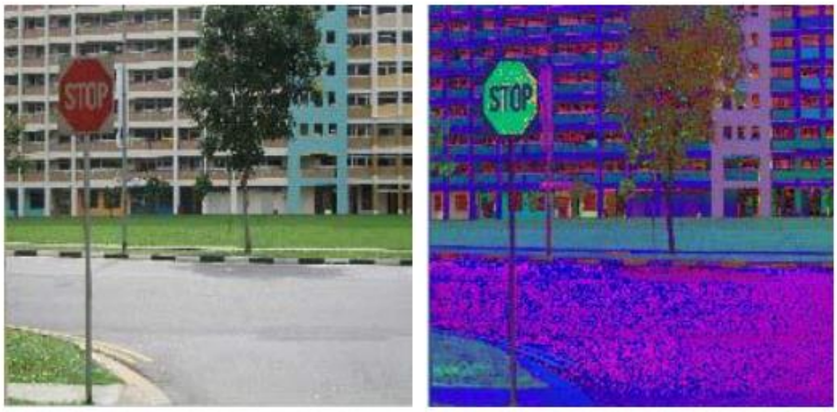
The authors in [10] discussed the colour distance method for the traffic sign detection from image sequences. The colour distance is similar to the Euclidean distance between two points and it is calculated by taking the difference between the two colours as in Equation 1, which indicates the similarity between two pixels i and j.

$$CD(i,j) = \sqrt{(R_i - R_j)^2 + (G_i - G_j)^2 + (B_i - B_j)^2} \quad (1)$$

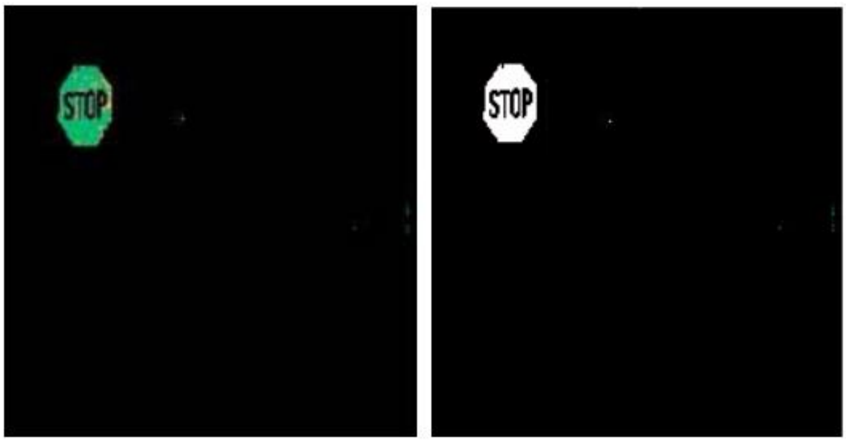
As the colour distance decreases the similarity increases. A colour standardization method is proposed in [11] for detecting the traffic signs. Here, each pixel in the RGB space is compared with the threshold values defined for each component. Only the component value larger than the threshold is converted to 255 otherwise it is set to 0. After colour standardization the image contains only black, red, green, blue, cyan, magenta and white. Only the interesting colour is retained in the image and others are discarded to get a binary image.

In the above methods, the RGB colour space is used for the segmentation. Even though the segmentation takes less computation time, it is sensitive to illumination changes.

In [12], the detection of the traffic sign is performed in the Hue–Saturation–Intensity (HSI) colour space. Since it separates achromatic and chromatic components, the detection can be performed even in the challenging lighting conditions. The image in RGB space is converted to HSI colour space and the segmentation is done based on the Hue and Saturation values. The image is transformed to binary by setting the pixels of interest to 255 and others to 0. The result for the STOP sign detection from [12] is given in the Figure 1.



(a) RGB representation b) HSI representation



(c) Pixels of interest in HSI b) Binary image

Figure 1. Result of colour segmentation in HSI colour space, Courtesy: [12]

To detect traffic signs, the authors in [13] used the Hue–Saturation–Value (HSV) colour space. The image in RGB colour space is transformed to HSV colour space, and the Hue and Saturation values are used to segment the image. Morphological operations such as erosion, dilation, opening, and closing are used to remove small background noise in the segmented image. The segmented image is then subjected to connected component analysis to identify possible traffic sign regions that meet the height, breadth, and area constraints.

In [14], YCbCr color space is chosen for the traffic sign detection. The red component is obtained from the YCbCr colour space by applying the dynamic thresholding on the Cr component value. The objects with the smaller area than some predefined value are eliminated as they are less likely to be the traffic signs.

Although the computation time for colour space conversion is long, segmentation in other colour spaces produces better results under different lighting conditions.

B. Shape Based Detection

The main issue with color-based detection is the variation in ambient illumination, which is why shape-based detection methods were developed. Traffic signs are typically triangle, rectangle, octagon, or circular in shape. This section discusses some popular algorithms for shape analysis.

In [15], traffic signs are detected by using Hough transform. Hough transform for circumference is used to detect the circular signs and triangular signs are detected by using the Hough transform for straight lines. Canny edge detection method is used to detect the edges in an image. Analysing all the contours in the image will lead to high computational cost and hence to reduce the computations some contours are rejected based on the area and perimeter of the contours. Then the Hough transform is applied to the contours. By using Hough transform, detection rates of 97.2% and 94.3% are achieved for the speed limit and warning signs respectively.

The radial symmetry transform is used in [16] to detect speed limit traffic signs. The original image is grayscaled, and the gradient of each pixel is calculated. Two vote images are constructed in such a way that they are both the same size. The two images, which contain gradient orientation and magnitude information, are referred to as orientation and magnitude images. Then, by combining two vote images, a radial symmetry image is created. Applying the threshold to the radial symmetry image detects the circles. Because the radial symmetry algorithm can only detect regular polygons, it is ineffective for detecting traffic signs with geometric distortion.

In [17], the traffic signs are detected using a template-based method. In this case, the binary images involved in the detection process are feature image (I) and feature template (T).

The matching between T and I is measured by the Distance Transform of the feature image. In the DT image, each pixel value represents the distance between that pixel and its nearest edge. The template is compared to the DT image to identify the shape of interest. As a result, the computational cost is extremely high.

C. Both Colour and Shape Based Detection

Using chromatic and shape information separately increases the number of interferences when the image contains objects of the same colour and shape. As a result, to reduce interferences, both chromatic and shape information are combined for detection. This consists of two stages: the first is colour segmentation in any colour space, and the second is traffic sign detection using shape analysis.

To detect traffic signs, the authors of [18] used both colour and shape information. RGB ratios are used to segment the RGB image. The Douglass-Peucker (DP) algorithm is then used to perform shape analysis. The DP algorithm is a contour approximation technique that detects objects based on the number of boundaries.

Figure 2. Results of traffic sign detection using both colour and shape information [18]

As a result, even if the traffic signs have some geometric distortions, they are detected. The results of [18] for traffic sign detection are shown in Figure 2. To detect traffic signs with circular shapes, both colour and shape are used as features. In this case, the red colour is segmented in the HSI colour space, and circular signs are detected using the Hough transform for circumference.

In [19], first the colour segmentation is done in the HSV color space and then boundary boxes are inserted for all the regions detected through the color segmentation. The traffic sign is detected by using the features of the bounding boxes such as mean color, size and number of pixels enclosed in the boundary box.

D. Other Approaches

The Maximally Stable Extremal Regions are used to detect traffic signs for the first time in [20]. (MSERs). The original RGB frame is converted to grayscale, and each frame is binarized at several threshold levels, with connected components found at each level. MSERs are components that retain their shape after multiple thresholding. It is not affected by changes in lighting.

The authors in [21] have used the colour probability model and the colour Histogram of Oriented Gradient (HOG) features to detect the traffic signs in the images. The colour probability model transfers the image into probability maps and enhances the colours associated with the traffic signs by suppressing the other colours. Then the MSER region detector is used to detect the regions that are likely to contain the traffic signs. After detecting the regions color HOG features are computed on the probability maps and are employed to train the Support Vector Machines (SVM).

The authors in [22] used the HSI colour space to improve the segmentation performance. The Distance to Border (DtB) features are computed for the blobs obtained from the segmentation process and are used to classify the blobs using SVM. An algorithm for the classification of pixels using Support Vector Machines (SVM) is presented in [23]. Samples of the target colour to be detected and some other colours from the training images were labelled and used to train the SVM. For simplicity RGB colour space is used. The results of the SVM outperformed the conventional colour thresholding methods. However, the main disadvantage of SVM is its speed.

4. Traffic Sign Recognition

After detecting traffic signs in images or video frames, the next step is to identify which traffic sign it is. The recognition can be accomplished through the use of either feature matching algorithms or machine learning algorithms. This section provides a brief description of the algorithms used for traffic sign recognition.

A. Feature Matching or Template Matching Algorithms

The authors of [18] used scale and rotation invariant Binary Robust Invariant Scalable Key point (BRISK) features for traffic sign recognition. BRISK detection and descriptor computation are faster than the accelerated robust features (SURF). The detection of keypoints and the extraction of binary bit-strings are the two main stages in the BRISK. The first stage involves identifying points of interest, and the second stage involves comparing brightness pixel by pixel to form the descriptor vector. The target class is the template image with the greatest number of keypoint matches.

For traffic sign recognition in [19], an interest point detector based on the SURF algorithm is used. SURF can function as both a detector and a descriptor. The former detects interest points in the image such as blobs, corners, and so on, and the latter creates a feature vector that represents the detected interest points. To determine the similarity match, the feature vectors of the target traffic sign are compared to those in the database.

In [24], a template matching method is introduced to classify the detected traffic signs into a specific class. In this step, the detected traffic signs are compared to a template stored in the database. To find a match between the detected traffic signs and the template stored in the database, the normalized cross-correlation (NCC) method is used.

Authors from [25] conducts a comparative analysis of three feature matching techniques (SIFT, SURF, and BRISK). The recognition process is carried out in two scenarios. In the first scenario, the road signs are manually segmented and compared to the signs in the database; in the second, the output of the detection system (i.e. after colour segmentation in HSV and detection using the Hough transform) is compared to the signs in the database. Figure 3 depicts the matching results using the SIFT, SURF, and BRISK descriptors from [25].

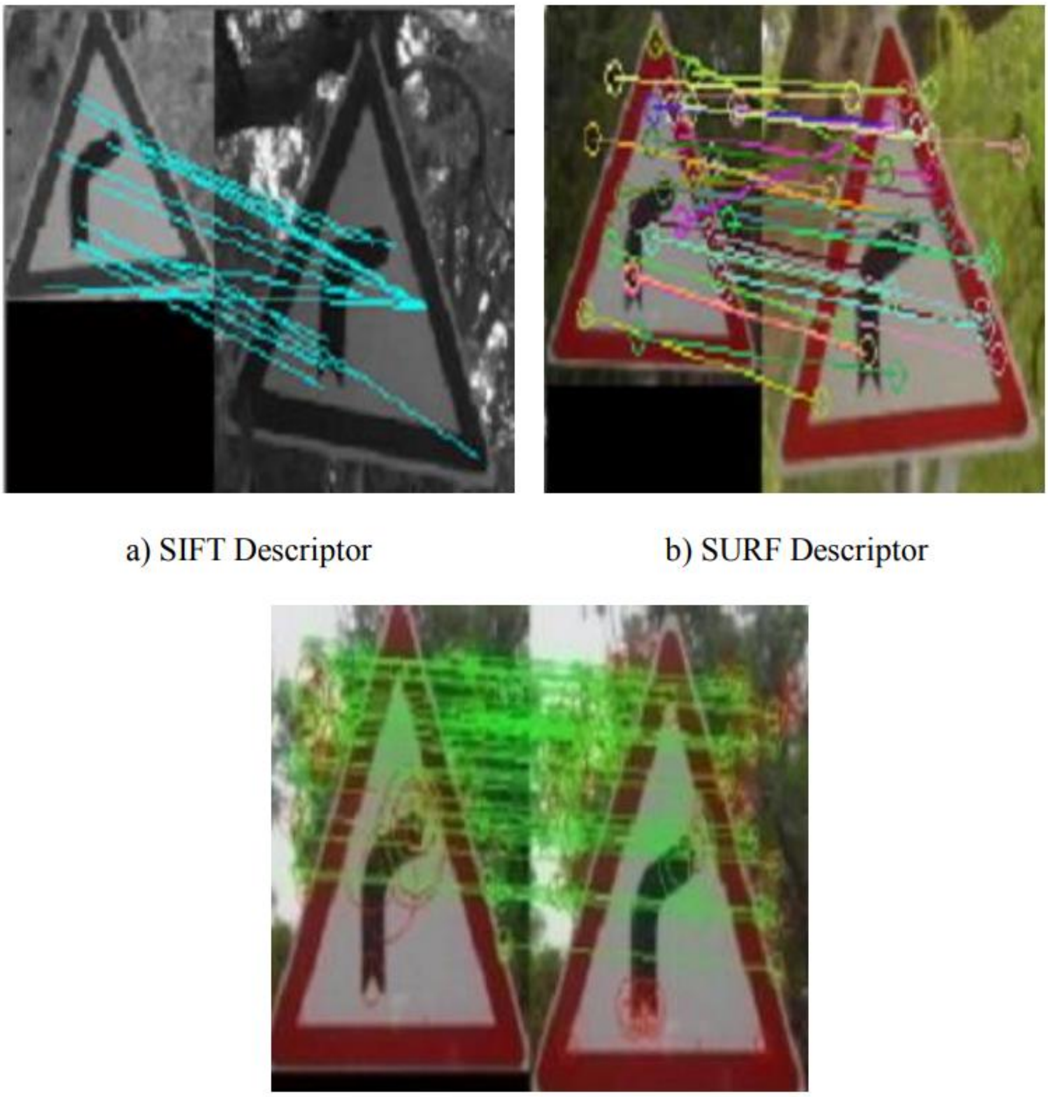


Figure 3 : Feature matching using 3 descriptor, Courtesy: [25]

B. Machine Learning Approaches

In the field of traffic sign recognition, learning-based methods are very popular. They are trained to determine an optimal separation between the classes. Neural networks and SVM are the most often utilised approaches for traffic sign recognition. Artificial Neural Networks (ANNs) are modelled after neurons. The biological neural network is thought to be responsible for human intelligence. This type of system is created artificially with the help of Artificial Neural Networks (ANN).

The authors of [9] employed an auto-associative neural network to recognise traffic signals in video frames and achieved 100% and 94.7 percent recognition rates in both daylight and shadow environments. The authors in [12] employed one-to-one multilayer perceptron neural networks trained with the robust back propogation method to recognise the detected traffic signs. A traffic sign recognition system based on the deep convolutional neural network is proposed in [26]. It achieves a recognition accuracy of 98.83%. Even though, it can provide good recognition rate, it requires a high computation complexity.

SVM is a method of supervised learning. SVM entails constructing a hyperplane to determine the class of the given data. The support vectors obtained from the data points define the margin of hyper-maximum plane. Zernike moments are combined with SVM in [11] for traffic sign recognition. Image scale, rotation and translation have no effect on Zernike moment features. The authors of [13] used SVM to classify traffic signs. The HOG features are used to train the SVM. HOG was initially used for pedestrian detection. HOG is invariant to scale variations.

The authors of [27] used a Genetic Algorithm to improve recognition accuracy under different environmental conditions (GA). The original template image is modified here by using the genetic algorithm's gene coding. The gene coding is represented by 24 bits, the first 6 bits representing the scaled version of the template, the second 6 bits representing the rotated version of the template, and the third and fourth representing the intensity modified and blurred version of the template image. The image's detected traffic sign is compared to the modified images. The training phase is not required for gene coding recognition. The system's overall accuracy is 94.7%.

Hu moment invariants and neural networks are used to detect traffic signs in [28]. Hu moments are rotation, scale, and translation invariant, resulting in great recognition accuracy. In addition, the computational complexity is modest. The authors of [29] employed the LBP feature extraction technique in conjunction with SVM to classify Chinese traffic signs. To account for scale discrepancies, the detected ROIs are resized to 64u64 pixels before being recognised. Gray scale and rotation invariance are not a problem for the LBP technique, and it is computationally simple. The uniform patterns provide rotation invariance while also reducing the complexity of the feature vector from 256 to 59.

5. CONCLUSION

The issues of traffic sign detection, as well as a brief review of known approaches for traffic sign detection and recognition, are described in this paper. The report details the extensive study that has been done in this area. In addition, the paper highlights a number of areas that require further investigation. These problems include recognising images that have been damaged by motion blur, images captured in occlusion, and signs captured in snow and rain.

REFERENCES

[1] Shawkat Ali, M. Ameer Ali, G.M. Shafiullah, Colin Cole, “Smart Driving: A New Approach to Meeting Driver Needs”, International Conference on Industrial Engineering and Operations Management”, 2010.

[2] H.s.g, Supreeth., & Patil, C. M. (2016). An approach towards efficient detection and recognition of traffic signs in videos using neural networks. 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), 456-459.

[3] Qin, F., Fang, B., & Zhao, H. (2010). Traffic sign segmentation and recognition in scene images. 2010 Chinese Conference on Pattern Recognition (CCPR), 1-5.

[4] S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing, andC.Igel, ‘‘Detection of traffic signs in real-world images: The German traffic sign detection benchmark,’’ in Proc. Int. Joint Conf. Neural Netw., Dallas, TX, USA, Aug. 2013, pp. 1–8.

[5] Z. Zhu, D. Liang, S. Zhang, X. Huang, B. Li, and S. Hu, ‘‘Traffic-sign detection and classification in the wild,’’ in Proc. CVPR, Las Vegas, NV, USA, Jun. 2016, pp. 2110–2118.

[6] A. Mogelmose, D. Liu, and M. M. Trivedi, ‘‘Detection of U.S. traffic signs,’’ IEEE Trans. Intell. Transp. Syst., vol. 16, no. 6, pp. 3116–3125, Sep. 2015.

[7] M. Costa, A. Simone, V. Vignali, C. Lantieri, and N. Palena, ‘‘Fixation distance and fixation duration to vertical road signs,’’ Appl. Ergonomics, vol. 69, pp. 48–57, May 2018.

[8] T. Ben-Bassat and D. Shinar, ‘‘The effect of context and drivers’ age on highway traffic signs comprehension,’’ Transp. Res. F, Traffic Psychol. Behav., vol. 33, pp. 117–127, Aug. 2015.

[9] H.s.g, Supreeth., & Patil, C. M. (2016). An approach towards efficient detection and recognition of traffic signs in videos using neural networks. 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), 456-459

[10] Qin, F., Fang, B., & Zhao, H. (2010). Traffic sign segmentation and recognition in scene images. 2010 Chinese Conference on Pattern Recognition (CCPR), 1-5.

[11] Xing, M., Chunyang, M., Yan, W., Xiaolong, W., & Xuetao, C. (2016). Traffic sign detection and recognition using color standardization and Zernike moments. 2016 Chinese Control and Decision Conference (CCDC), 5195- 5198.

[12] Nguwi, Y., & Kouzani, A. (2006). Automatic road sign recognition using neural networks. The 2006 IEEE International Joint Conference on Neural Network Proceedings, 3955-3962.

[13] Romdhane, N. B., Mliki, H., & Hammami, M. (2016). An improved traffic signs recognition and tracking method for driver assistance system. 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), 1-6.

[14] Fatmehsari, Y. R., Ghahari, A., & Zoroofi, R. A. (2010). Gabor wavelet for road sign detection and recognition using hybrid classifier. 2010 International Conference on Multimedia Computing and Information Technology (MCIT), 25-28

[15] Gracia-Garrido, M., Sotelo, M., & Martin-Gorostiza, E. (2006). Fast traffic sign detection and recognition under challenging lighting conditions. 2006 IEEE Intelligent Transportation Systems Conference, 811-816.

[16] Barnes, N., & Zelinsky, A. (2004). Real-time radial symmetry for speed sign detection. IEEE Intelligent Vehicles Symposium, 2004, 566-571.

[17] Garvila, D. M. (1999). Traffic sign recognition revisited. Informatik aktuell Mustererkennung 1999, 86-93.

[18] Zheng, Z., Zhang, H., Wang, B., & Gao, Z. (2012). Robust traffic sign recognition and tracking for Advanced Driver Assistance Systems. 2012 15th International IEEE Conference on Intelligent Transportation Systems, 704-709.

[19] Oruklu, E., Pesty, D., Neveux, J., & Guebey, J. (2012). Real-time traffic sign detection and recognition for in-car driver assistance systems. 2012 IEEE 55th International Midwest Symposium on Circuits and Systema (MWSCAS), 976-979.

[20] Greenhalgh, J., & Mirmehdi, M. (2012). Real-time detection and recognition of road traffic signs. IEEE Transactions on Intelligent Transportation Systems, 13(4), 1498-1506.

[21] Yang, Y., Luo, H., Xu, H., & Wu, F., (2014). Towards realtime traffic sign detection and classification. 17th International IEEE Conference on Intelligent Transportation Systems, 17(7), 2022-2031..

[22] C.g., K., Prabhu, L.V., A. R., & K., R. (2009). Traffic sign detection and pattern recognition using support vector machine. 2009 Seventh International Conference on Advances in Pattern Recognition, 87-90.

[23]] Gomez-Moreno, H., Maldonado-Bascon, S., Gil-Jimenez, P., & Lafuente-Arroyo, S. (2010). Goal evaluation of segmentation algorithms for traffic sign recognition. IEEE Transactions on Intelligent Transportation Systems, 11(4), 917-930.

[24] Farhat, W., Faiedh, H., Souani, C., & Besbes, K. (2016). Real-time recognition of road traffic signs in video scenes. 2016 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), 125-130

[25] Malik, Z., & Siddiqi,I. (2004). Detection and recognition of traffic signs from road scene images. 2014 12th International Conference on Frontiers of Information Technology, 330-335.

[26] Qian, R., Zhang, B., Yue, Y., Wang, Z., & Coenen, F. (2015). Robust Chinese traffic sign detection and recognition with deep convolutional neural network. 2015 11th International Conference on Natural Computation (ICNC), 791-796.

[27] Kobayashi, M., Baba, M., Ohtani, K., & Li, L. (2015). A method for traffic sign detection and recognition based on genetic algorithm. 2015 IEEE/SICE International Symposium on System Integration (SII), 455-460.

[28] Hossain, M. S., Hasan, M, M., Ali, M. A., Kabir, M. H., & Ali, A. B. (2010). Automatic detection and recognition of traffic signs. 2010 IEEE Conference on Robotics, Automation and Mechatronics, 286-291

[29] Zhang, H., Wang, B., Zheng, Z., & Dai, Y. (2013). A novel detection and recognition system for Chinese traffic signs. 2013 32nd Chinese Control Conference (CCC), 8102-8107.