Proceedings of the Eastern Asia Society for Transportation Studies, Vol.12,2019

Traffic Sign Detection Using Convolutional Neural Networks

Ngoc Dung BUI [a*], Hieu Vu TRAN [b], Binh T. H. NGUYEN [c]

[a,b] *University of Transport and Communications, Hanoi, Vietnam*
[a] *E-mail: dnbui@utc.edu.vn*
[b] *E-mail: hieutv@utc.edu.vn*
[c] *University of Transport and Communications, HCM City's Campus, Vietnam*
[c] *E-mail: binhnth@utc2.edu.vn*

Abstract: Safety in driving can be greatly enhanced by automated and specialized devices. However, many tasks are still heavily depend on the operation of the drivers. To improve the safety of the daily life transportation, - many methods have proposed to help drivers keep attention in any cases that may cause accidents. This paper presents an approach to automatically recognize the traffic signs using cameras on the vehicles. The proposed approach extracts the features, which is local information from image sequences and recognize the signs using a neural network. The proposed algorithm was tested in several roads in Vietnam. The experimental results demonstrate a great accuracy.

Keywords: Driving Safety, Traffic Sign Recognition, Convolutional Neural Networks

1. INTRODUCTION

Intelligent vehicles, including vehicles with driving assistance systems and autonomous vehicles, have been gaining attentions from both academic and industry. In order to ensure the efficiency and safety of intelligent vehicles, robust automatic traffic sign recognition algorithms are required. The traffic sign recognition is a challenging task due to several reasons, such as partial occlusion, viewpoints, illuminations, and weather conditions. The existing traffic sign recognition algorithms mostly reply on two main techniques, which are color segmentation (Tsai *et al*. 2008; Gao *et al*. 2006; De la Escalera *et al*. 2003; Broggi *et al*. 2007) and template matching (Betke and Makris, 1995; Torresen *et al*. 2004; Miura *et al*. 2002). The eigen color model was used to detect road signs in (Tsai *et al*. 2008), the authors used statistical analysis to collect data of R, B and G channels and then applied Karhunen-Loeve transform to segment the color. In (Gao *et al*. 2006), color ranges of traffic signs were found under different conditions such as rainy, sunny and cloudy in the training phase. In the test phase, traffic sign images were converted from RGB color model to LCH (Lightness, Chroma, Hue) color model using CIECAM97 model. After that a quad-tree histogram was used to segment the traffic sign based on color information. HIS (Hue, Saturation, Intensity) color model was used in (De la Escalera *et al*. 2003), this algorithm analyzed only hue and saturation components of the image, parts of the image which satisfy some given conditions will be selected and then will be compared with all possible collected signs. Segmentation using color information also was used in (Broggi *et al*. 2007). The authors used color information together with linear transformation to detect road signs in real-time in Italy. In (Betke and Makris, 1995), the authors matched all objects in images with online templates using location, shape, size, and orientation of objects. Templates were

[*] Corresponding author.

suggested in (Torresen *et al*. 2004) to detect speed limit signs and then recognize displayed number of speed limit signs in Norway. Another similar approach is used in (Miura *et al*. 2002), the authors built the templates of speed signs and character of guidance signs in a real-time traffic sign system.

These above-mentioned algorithms encounter problems when weather changes rapidly in real world condition, this rapid change can greatly affect color and displayed character of traffic signs. To overcome this difficulty, researchers rely on features which are weather-invariant, such as edges (Franke *et al*. 1998), Haar wavelets (Bahlmann *et al*. 2005). Edges of each sign is utilized to detect a sign in (Franke *et al*. 1998), each pixel was checked at four direct neighbors to see if each neighbor's brightness is brighter, darker or is same as the central pixel. Haar wavelets features were trained from Ada-Boost in (Bahlmann *et al*. 2005) to recognize traffic signs. In (Nassu et al. 2010), Nassu and Ukai suggested an approach which utilizes SIFT (Scale-Invariant Feature Transform). This method can avoid problems of color segmentation and template matching techniques.

A new era of object classification and detection has been opening since the birth of deep neural networks (DNNs). Inspirited by the results of DNNs in many applications, several papers have been proposed CNN-based traffic sign recognition and detection (Wu et al. 2013; Boujemaa et al. 2017; A. Shustanov and P. Yakimov, 2017). The proposed algorithms are invariant to scale and viewing angle. However, they are slow and sensitive to weather change. This paper proposes an algorithm to recognition traffic signs by combining color-based traffic sign extraction and convolutional deep neural networks to solve the weakness of the existing algorithms. The paper is organized as follows: Section 2 present the proposed method to detect and recognize traffic signs. The results is also presented in this section to illustrate the performance of the proposed algorithm. Finally, the conclusions are stated in Section 3.

2. PROPOSED TRAFFIC SIGN RECOGNITION ALGORITHM

The proposed method consists of two main steps. The first step extracts the all regions of interest (ROI) from an image using the color information. The second step utilizes a convolutional neural network (CNN) to recognize the traffic signs in the extracted ROIs. The overall flowchart of the proposed algorithm is shown in figure 1.

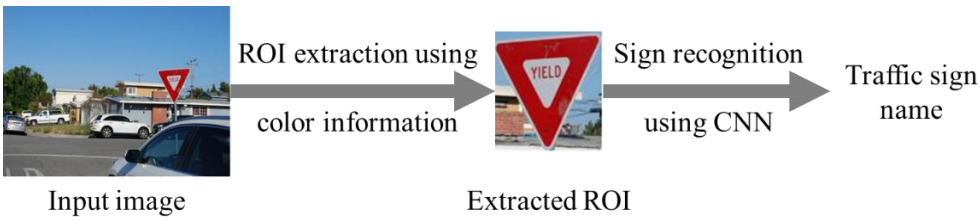


Figure 1. Flowchart of the proposed sign recognition algorithm

2.1. ROI Extraction

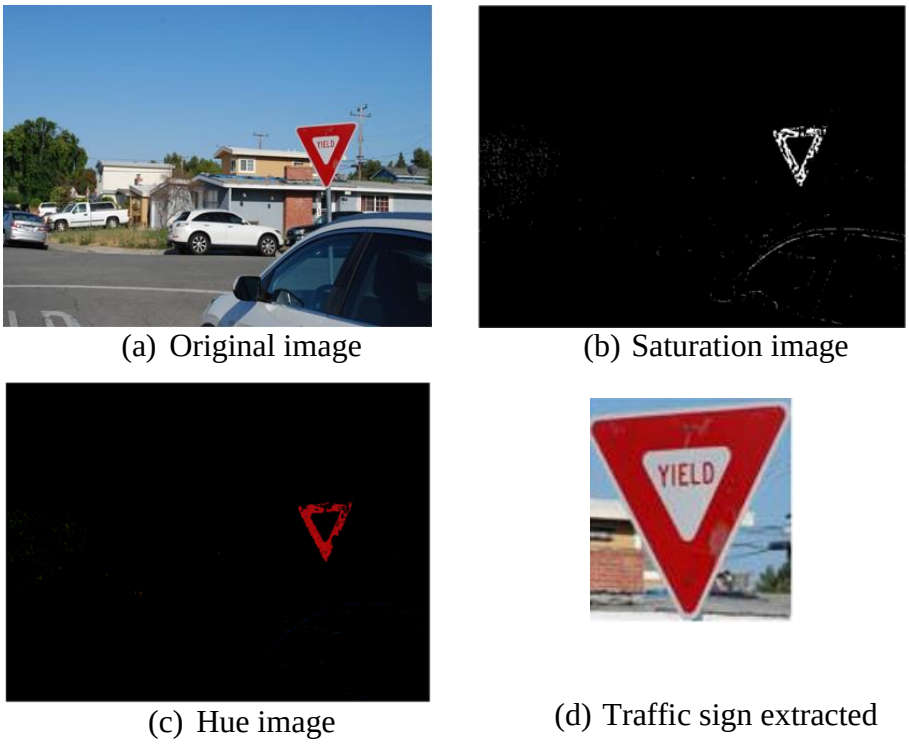
In order to reduce the search space for the sign detection task, the regions of interest are extracted from images using color information. The proposed ROI extraction algorithm utilizes color since this feature is invariant to translation, rotation, and scaling. Also, it is well-known that traffic signs are made up of highly saturated Red, Blue and Yellow colors (Gaurav *et al*, 2014). To reduce the effect of illumination, instead of using RGB color model, HSV model is used to extract ROIs. The process of ROI extraction starts with converting the

input image from RGB color model to HSV color model. The hue and saturation images are then binarized using predefined thresholds. The thresholds are 0.9 for hue and 0.9 for saturation. Outputs of the binarization are then combined using (1) to create one image which contains traffic sign candidates.

$$I_{candidate} = H_B \otimes S_B \tag{1}$$

where H_B and S_B are the output of binarizing the hue and saturation images, $\otimes$ is the pixel-wise multiplication. All pixels which may belong to a traffic sign will have value 1 in the image $I_{candidate}$. Examples of this step is shown in Figure 2b and 2c.

Since $I_{candidate}$ may contains noise, the characteristic that most pixels in traffic signs are in red, blue, and yellow color is used to remove noise. For each connected component in $I_{candidate}$, the number of red, blue, and yellow pixels is counted. If this number is smaller than a threshold (in the experiment, the threshold is set to 0.9), the connected component will be removed from the traffic sign candidate list. Traffic sign ROIs are then extracted using the input image and the $I_{candidate}$ image. Figure 2d displays an example of traffic sign ROI.



(a) Original image (b) Saturation image

(c) Hue image (d) Traffic sign extracted

Figure 2. The traffic sign extracted

2.2. Traffic Sign Recognition using Convolutional Neural Networks

The convolutional neural network (CNN) used to recognize traffic signs consists of three convolutional layers, three max pooling layers, and two fully connected layers. The architecture of the CNN is display in figure 3. The inputs of the CNN are images of 32×32 pixels. Hence, ROIs extracted from the previous step must be resized to 32×32 before feeding to the CNN. All convolutional layers of the proposed CNN share the same layout, which consists of convolutional filters, batch normalization and ReLU activation function. However, the numbers of filters and the filter sizes are different. The numbers of filters are 50, 100, and 200 in the first, second, and third convolutional layers. All conv layers use the stride of one,

but the sizes of the filters are 7×7, 5×5, and 3×3. The max pooling layers are conducted using 2×2 filters with a stride of two. The final activation function of the CNN is Softmax, which is suitable for the classification problem. In this paper, the proposed algorithm is developed to recognize ten different traffic signs. Hence, outputs of the CNN are probabilities of being ten traffic sign classes. Input images are assigned to the classes that maximize the probabilities.

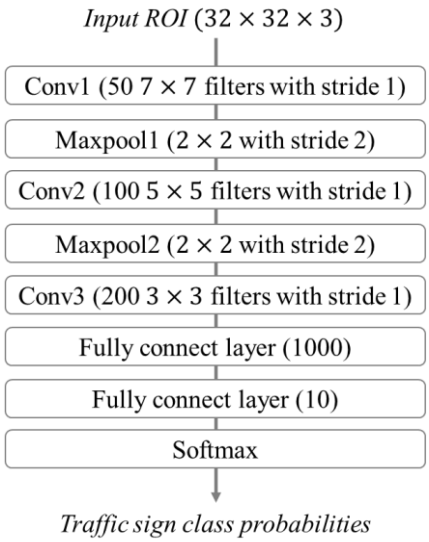


Figure 3. The proposed CNN architecture

2.3. Results and Discussion

The proposed algorithm is trained on a datasets of ten different traffic signs. The total number of images in the dataset is 223. Since the number of training images is small, the data augmentation method, including random rotation, scaling, and translation, is used. Total number of training images after augmentation is 55,920. Some of the signs are shown in figure 4. The CNN is trained with the following parameters: the learning rate is 0.001, the number of epochs is 200, and the batch size is 128.

Figure 4. A part of training images

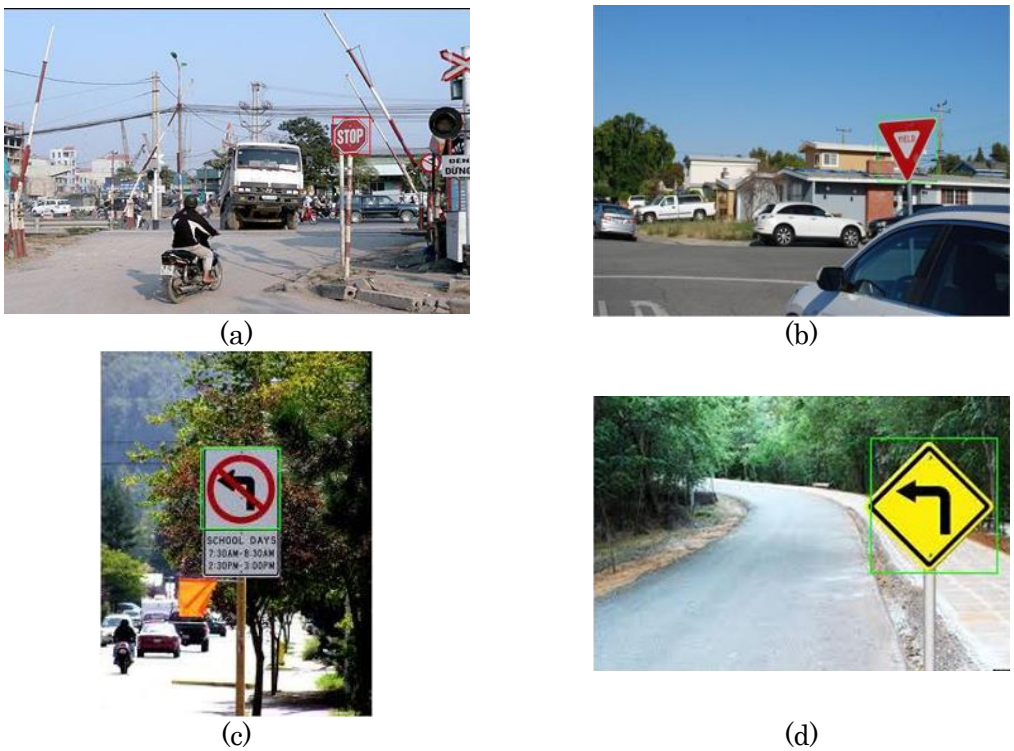
For the testing, the images sequence can be captured by camera on vehicles. The specification of the camera is described following. The camera model is anytekT6 the image resolution is 1Mp, and the frame rate (FPS) is 25f/s. The speeds of the vehicles range from 5km/h to 20km/h. The test images were capture under different weather conditions, including sunny, cloudy, light and heavy rain. The acquisited images have low quality due to non-static camera and other conditions, such as weather and occlusions. The ROIs are extracted from test images using the process presented in Section 2.1. The proposed CNN is then used to determine the traffic sign in each ROI. ROIs are resized to 32×32 before feeding to the CNN.

Figure 5 presents the results of the method. Table 1 and figure 6 show the accuracy the proposed method. The detection rate on the B5 traffic sign, which is the limited under 20km/h, is low because of the occlusion problem. In the future, we will include images with various occlusion conditions to the training set to solve that problem. To prove the advantages of the proposed algorithm, SIFT-based traffic sign recognition was implemented. The result shows that CNN improves the recognition accuracy 5% % compared to the SIFT-based method.

Table 1. Detection results of the proposed algorithm for different traffic signs

	B1	**B2**	**B3**	**B4**	**B5**	**B6**	**B7**	**B8**	**B9**	**B10**
B1	23	0	0	0	0	0	0	0	0	0
B2	1	26	0	0	0	0	0	0	0	0
B3	0	0	16	0	0	0	0	0	0	0
B4	0	0	0	24	0	0	0	0	0	0
B5	0	0	0	0	8	3	0	0	0	0
B6	0	1	0	4	4	27	0	4	2	0
B7	0	0	0	0	0	0	15	0	0	0
B8	0	0	0	0	0	0	0	25	0	0
B9	1	0	0	0	0	0	0	0	29	0
B10	1	0	0	0	0	0	0	0	0	29

The content of the B1-B10 traffic signs are Turn left, Turn right, Stop, Compulsory turn left, Yield, No turn left, No turn right, No U turn, Speed limit, No stop, respectively.



(a) (b) (c) (d)

Figure 5. The signs detection. (a): stop sign detected, (b): give way to other driver sign detected, (c): no turn left sign detected, (d): turn left sign detected

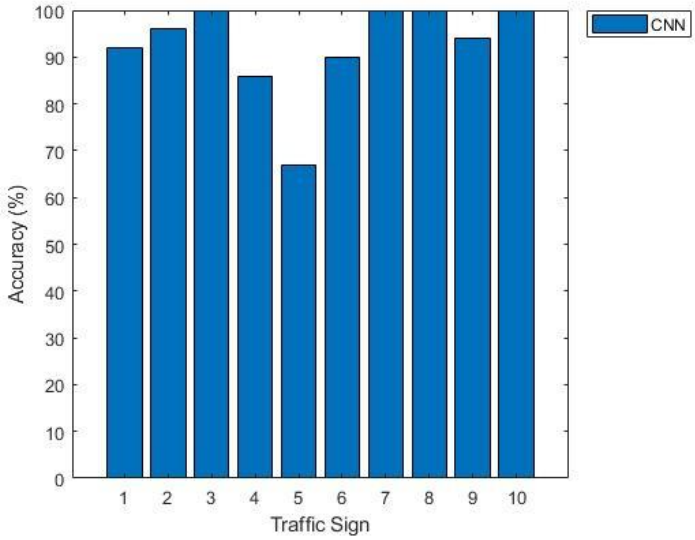


Figure 6. Accuracy of the proposed algorithm for different traffic sign

In the case of more than one signs appear in the same image, the method is missing in recognition because the distribution of colors in separate areas. Also, the fail caused by the presence of fully saturated Red, Blue or Yellow colors along with traffic signs. To improve the recognition, in future we may use more information about the color space and shape.

3. CONCLUSIONS

This paper present the method for traffic signs recognition. The proposed method was built on the local information which is extracted from the signs and recognized by the convolution neural network. The result proved the accuracy and robustness of the algorithm. This approach which is based on computer vision can greatly support drivers and as a result can enhance the safety of intelligent vehicle. In the near future, we will include more traffic signs in the systems and solve the problem of illumination changing.

REFERENCES

Alexander Shustanov and Pavel Yakimov (2017), CNN Design for Real-Time Traffic Sign Recognition, *Procedia Engineering*, 201, pp. 718-725.

Bahlmann, C., Zhu, Y., Ramesh, V., Pellkofer, M. & Koehler, T., (2005) A system for traffic sign detection, tracking, and recognition using color, shape, and motion information. *Proceedings 2005 IEEE Intelligent Vehicles Symposium, IEEE*, pp. 255-260.

Betke, M. & Makris, N.C., (1995) Fast object recognition in noisy images using simulated annealing. *Proceedings of Fifth International Conference on Computer Vision 1995, IEEE*, pp. 523-530.

Boujemaa, K. S., Berrada, I., Bouhoute A. and Boubouh K., (2017) Traffic sign recognition using convolutional neural networks. *2017 International Conference on Wireless Networks and Mobile Communications (WINCOM), Rabat*, pp. 1-6.

Broggi, A., Cerri, P., Medici, P., Porta, P.P. & Ghisio, G., (2007) Real time road signs recognition. *Intelligent Vehicles Symposium, 2007 IEEE*, pp. 981-986.

De la Escalera, A., Armingol, J.M. & Mata, M., (2003) Traffic sign recognition and

analysis for intelligent vehicles. *Image and Vision Computing*, 21, no. 3, pp. 247-258.

Franke, U., Gavrila, D., Gorzig, S., Lindner, F., Puetzold, F. & Wohler, C., (1998) Autonomous driving goes downtown. *IEEE Intelligent Systems and Their Applications*, 13(6), pp. 40-48.

Gao, X.W., Podladchikova, L., Shaposhnikov, D., Hong, K. & Shevtsova, N., (2006) Recognition of traffic signs based on their colour and shape features extracted using human vision models. *Journal of Visual Communication and Image Representation*, 17 (4), pp. 675-685.

Gaurav, M., Aabhas. Majumdar., Sudarsh. Rathi., (2014) Traffic Sign Identification System, *Technical Report*.

Ian Nabney, (2004) Netlab: Algorithms for Pattern Recognition. *Advances in Pattern Recognition, Springer*.

Miura, J., Kanda, T., Nakatani, S. & Shirai, Y., (2002) An active vision system for on-line traffic sign recognition. *IEICE Transactions on Information and Systems*, 85(11), pp. 1784-1792.

Nassu, B.T. & Ukai, M., (2010) Automatic recognition of railway signs using SIFT features. *Intelligent Vehicles Symposium (IV) 2010, IEEE*, pp. 348-354.

Ripley, B.D., (1996) Pattern Recognition and Neural Networks. *Cambridge University Press, UK*.

Torresen, J., Bakke, J.W. & Sekanina, L., (2004) Efficient recognition of speed limit signs. *Proceedings The 7th International IEEE Conference on Intelligent Transportation Systems 2004*, pp. 652-656.

Tsai, L., Hsieh, J., Chuang, C., Tseng, Y., Fan, K. & Lee, C., (2008) Road sign detection using eigen colour, *IET computer vision*, 2(3), pp. 164-177.

Wu Y, Liu Y., Li J., Liu H. & Hu X., (2013) Traffic sign detection based on convolutional neural networks. *The 2013 International Joint Conference on Neural Networks (IJCNN), Dallas, TX*, pp. 1-7.