Journal of Network Intelligence ©2019 ISSN 2414-8105 (Online)
Taiwan Ubiquitous Information Volume **4**, Number **2**, May **2019**

Research and Implementation of Road Traffic Sign Identification System

Chien-Chung Wu, Shiue-Ling Wu

Department of Computer Science and
Information Engineering Southern Taiwan
University of Science and Technology,
No.1, Nantai St., Yongkang Dist., Tainan City 710, Taiwan (R.O.C.)
wucc@stust.edu.tw, ma4g0105@stust.edu.tw

Received December 2018; revised January 2019

Abstract. *It is helpful to enhance traffic safety by driving information which is provided by the roadside traffic signs. However, as the traffic sign information cannot be seen immediately while driving, drivers always ignore a lot of important driving information. To solve this problem, an automatic traffic sign identification system which took road as an example was proposed. The system has been developed with image recognition technology and can provide the identification result of related driving information to drivers. The system proposed in this study was divided into two parts. Firstly, through the brightness of the image, the possible traffic sign image and its position could be quickly detected with the classifier trained by Haar Cascade in Haar-like with AdaBoost method. Secondly, the detected images could be classified by the convolutional neural network (CNN) classifier designed with GoogLeNet Network Architecture in TensorFlow library and Python program language. To speed up the recognition process and the correctness of recognition result, a method was proposed in this study to avoid analyzing those areas without traffic signs. In the method, the collected road videos were divided into 12 regions to find out the non-traffic sign areas. After carrying out this method, the recognition speed has been improved 11.24 frame/sec from to 12.04 frame/sec. Besides, the precision of the system can be improved from 80.97*
Keywords: Haar-like, CNN, Road Traffic Sign Identification, TensorFlow

1. **Introduction.** Road traffic sign detection and recognition are the important and popular modules in the driving assistant systems, and also can be used as the navigation system for the blind. Drivers and the blind can predict the road situation ahead by analyzing the detected and recognized results and avoid dangers. There are many methods to conduct the image recognition, no matter through hardware to software.

In hardware, Bekircan Yavas et al. [1] had proposed that through OpenCV training to do the red circular outline detection and text recognition for speed limit traffic signs. The training result was placed in RFID, and the portable RFID devices were integrated into road traffic signs. On the label, the vehicle was able to identify road traffic signs as long as it had an RFID reader installed. 1 As for software, Hengliang Luo et al. [2] had proposed a road traffic sign identification system. The system can be divided into three phases. It was to extract the ROI (Regions of Interest) from the image, and then to filter out the non-traffic ROIs. Finally, the remaining ROIs were classified. The method of using multi-task convolutional neural networks was proposed in the classification of modified ROIs. The classification was mainly divided into two decision layers, namely, the binary classification layers, which were used to distinguish between background and road traffic signs very quickly, while the other was a multi-category traffic classification layer, which would use the ROI of the binary classification layer to perform more in-depth feature comparisons to obtain the road traffic sign classification results.

There are many ways to identify images. For example, Haar-like Cascade [3], [4], [5] is superior in fast image feature computing and has high stability against anti-noises and brightness changes, but the disadvantage is that the error rate for detected images is relatively high. With the aid of this algorithm, the target objects can be obtained quickly from the video frames. In the past decade, the convolution neural networks (CNNs)[6] became a popular method for the image recognition, and the image features could be predicted and classified, but the disadvantage is the size of input image was fixed. To increase the correctness of the road traffic sign detection, a method composed of Haar-like cascade and CNNs on road traffic sign detection has been proposed in this work. The road traffic sign were detected by the Haar-like cascade only in the first method. In the experiment, if video frame detected by this method, some incorrect results might be detected. The second method is to implement the CNNs classification after the Haar-like cascade. The incorrect detection will be fixed in the second method.

2. **Motivation.** In recent years, the complexity of driving routes has caused serious injuries and deaths from traffic accidents. While driving, drivers must pay attention to road conditions and also have to watch out for road traffic signs in the vicinity at the same time. A traffic accident may occur because the message provided by the traffic signal would usually be ignored by drivers. For example, nowadays, mobile phones are very common in our daily life. People use them so often, no matter walking, parking, or even driving. Therefore, pedestrians or drivers might easily overlook the message from traffic signals in the surrounding environment, then the unexpected accident would happen. The identification and application of road traffic signs have played an important role in the driver assistance system. With a great improvement of the high-resolution image equipment and the related technology of image processing, the development of image vision in the application of the driver assistance systems has been promoted. In the image processing, deep learning has flourished in recent years and combined with the computer's computing capabilities, so the image recognition by deep learning has become easier, with good results and high accuracy. If the identification of traffic signals by deep learning and combine it with the driver assistance system, the function of the driver assistance system can be more complete. Through the analysis and identification of images, drivers can receive correct traffic signal alerts by the system when the road traffic signal is approaching. Drivers can be notified of the traffic situation beforehand. Therefore, the system has proposed to provide drivers the information of road traffic signs ahead through the analyzed road traffic sign images with image processing and deep learning.

3. **Road Traffic Sign Detection and Identification Methods.** There are many methods to detect objects. The main purpose is to find the regions of interest (ROI) in the image. ROI is a block that represents a target in the image. For instance, if the detection object is a road traffic sign, and the area in the image which includes road traffic sign is called as ROI. There are many methods to detect the object position in the image, for instance, selective search by hierarchical grouping, bounding box with non-maximum suppression, Haar-like features with AdaBoost, etc.

In recent years, deep learning has become the mainstream of image identification. Deep learning is one of the machine learning. Its purpose is to learn data from multi-layers network and automatically obtain features that can represent data. The precision of classifiers and time-consuming on identifying objects are often affected by its network architecture and network depth. CNN is one of the common deep learning network architectures. Through the convolutional layer and the pooling layer in the network architecture, the relationship between the identification data is strengthened, and the image and voice recognition would be well-achieved. The popular deep learning network architectures include Alexnet[7], GoogLeNet[8][9], VGG16[10], YOLO[11], etc.

In order to reduce the search time for the entire image through CNN, a special design is of using strong classifiers composed of multiple weak classifiers to find the target of road traffic signals quickly. Then the road traffic sign identification module which based on GoogLeNet[8] neural network architecture is used to classify the detected road traffic sign, while eliminating some misjudgment. Each module will be described next.

3.1. **Road Traffic Sign Detection.** The Haar-like feature combined with Ada Boost method is used for road traffic sign detection in this system and the training process is shown as Figure 1. The classifier is trained through the Cascade Classifier Training tool provided in the OpenCV library. Firstly, the positive and negative samples collected by the sample retrieval tool will be normalized into specific size. Secondary, the positive sample vector file is created through opencv_createsample.exe in OpenCV library. And the negative sample information is recorded into a file which contains the image file names and file paths. Then the classifier is created through the training tool opencv_traincascade.exe in OpenCV library

3.2. **Road Traffic Sign Identification.** The road traffic sign identification is based on the GoogLeNet[8] network architecture and trained on TensorFlow library with required sample images. The training process is as Figure 2. Firstly, collect the positive and negative sample images through the image capture software, and the necessary positive images collected in classifier are pedestrians ahead, slow, steep descent, etc.; the negative images collected in classifier are road, trees, building, etc. After the sample images were collected, the collected positive samples would be divided into warning signs and prohibitive signs based on the nature and the outline. The classified method is shown as Figure 2. Different types of images are placed in different folders and named according to their categories, such as pedestrians ahead, left reverse bend, and steep descent. After completion, the classifier training can be started. When the training is completed, there will generate trained_graph.pb and trained labels.txt two files. All of the parameters related to the weights and features in the training are recorded in the trained_ graph.pb, while all of the categories recorded during the training are in the trained labels.txt. In particular, it is important to note that at least 100 samples of each category of training should be sufficient to validate the training classifier.

4. **System Design and Architecture.** Since the system is designed to identify road traffic signs in Taiwan, and there is no road traffic sign data base available publicly. Collecting Taiwan road traffic sign training samples are necessary. The sample images were collected by an image capture software and the sample sources were from the video captured by the dashboard camera. The videos were downloaded from YouTube and self-captured. To speed up the system process, a method to limit the detection area is proposed in the system. From 12 testing videos, the road traffic sign positions were recorded and the road traffic signs were not located in some area on the testing videos. Those areas would be ignored when detecting the targets. The system is combined with

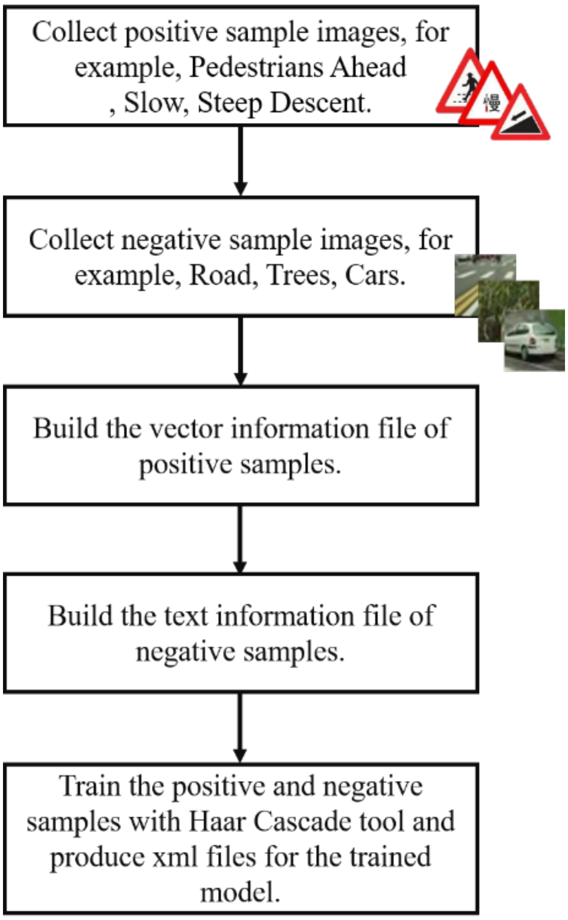


Figure 1. Training steps of road traffic sign detection classifier

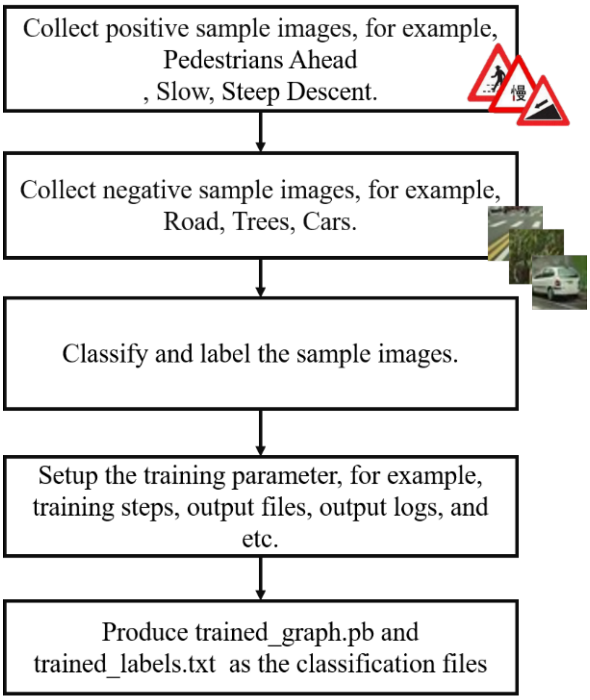


Figure 2. Training steps of road traffic sign detection identifier with GoogLeNet architecture

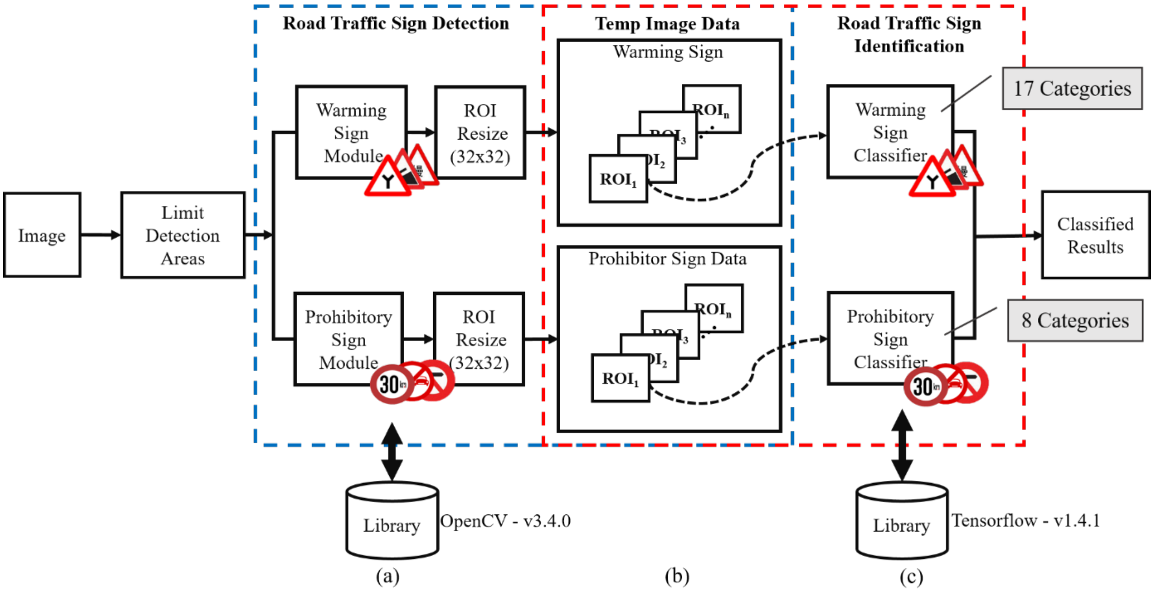


FIGURE 3. System process: (a) normalize the ROIs which detected by Fast Detection Classifier, (b) save ROIs in temp storage, (c)identify the ROIs and output results

road traffic sign detection and identification. As shown in Figure 3, firstly, with the feature of Haar-like with AdaBoost, through the brightness of the images, the target in input image or video can be detected quickly. After targets detected, the detected images would be transferred to the road traffic sign identification classifier which was built with TensorFlow library

4.1. **Experiment and Result.** The experimental operating system is under Ubuntu 16.04 64-bit. In terms of hardware, the processor is Intel i7-3770@3.40GHZ x 8, and the memory is DDR3.24GB. The development tools include Python v3.5.2, OpenCV v3.4.0, NumPy v1.13.3 and TensorFlow v1.4.1 suites. To evaluate and measure the performance in the road traffic sign identification system, the Precision and Recall definitions in the field of information retrieval were used to evaluate the performance of detection and identification results. As the formula (1), precision is defined to represent the proportion of all the correct information retrieved.

As the formula (2), recall is defined to represent the percentage of all the correct information. True positive (TP) refers to the positive targets which were correctly detected or identified by the classifier. False positive (FP) refers to the negative targets and were incorrectly detected or identified as positive by the classifier. True negative (TN) refers to the negative targets which were correctly detected or identified by the classifier.

$$Precision = \frac{TP}{TP + FP} \times 100 \tag{1}$$

$$\mathrm{Re}\,call = \frac{TP}{TP + TN} \times 100 \tag{2}$$

The positive and negative sample images are collected from 34 different road videos which were downloaded in YouTube and self-captured. There are a total of 25 types of road traffic signs that can be classified by the road traffic sign identification system proposed in this paper. There are 17 types of warning signs, and there are 8 kinds of the prohibitive signs, and the road traffic signs to be recognized are as Figure 4. To

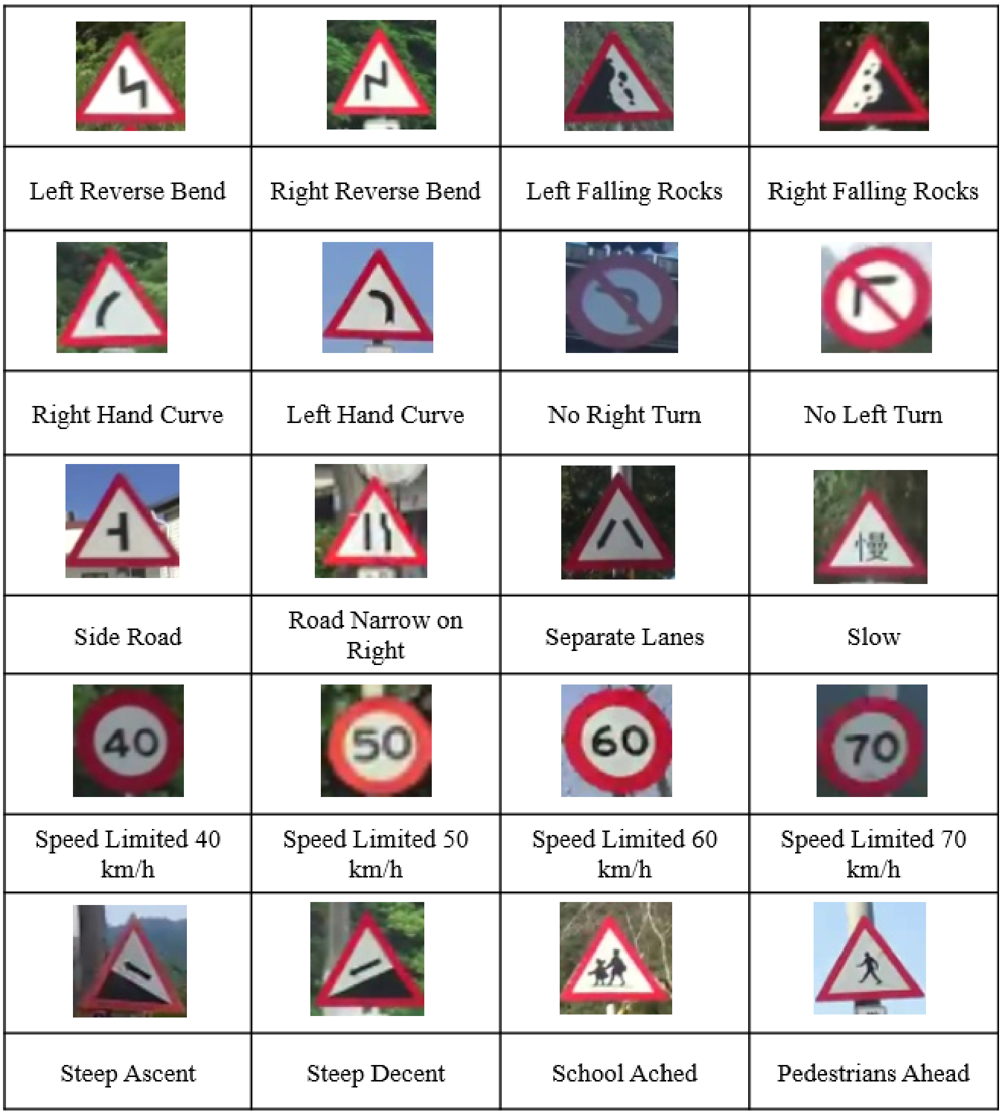


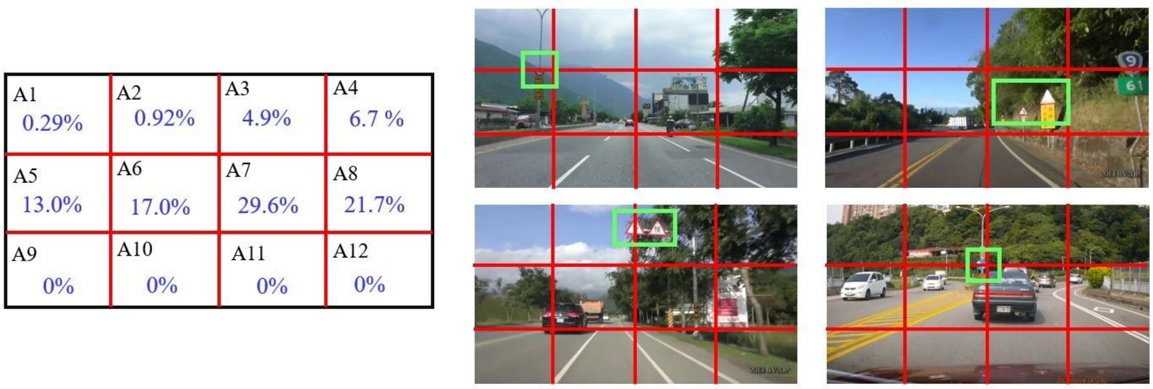
FIGURE 4. Recognition road traffic sign types

analyze and define the areas which road traffic signs usually appeared, those 10 testing videos were captured in different roads, for example, Hualien area, Taitung, etc. and the analyzed result shown as Table 1. The calculated result of 10 testing videos is as Figure 5(a) and 12 areas shown as Figure 5(b). After evaluating the limited detection area by comparing the time consumed only through road traffic sign detection classifier with through both road traffic sign detection and identification classifier, as shown in Table 2, the average detection time per frame in the detection of the full image with only road traffic sign detection classifier took 15.63 milliseconds, and after limited the detection area, the average detection time required for each frame was 11.00 milliseconds, and each frame could be reduced by 4.63 milliseconds. Under the condition of through both road traffic sign detection and identification classifier, the detection time of each frame in the full image was estimated to be 88.96 milliseconds, and after the detection area was limited, the average detection time per frame consumed 83.07 milliseconds, and each frame could be reduced 5.89 milliseconds

After evaluating the average precision and recall of road traffic sign detection with full detection area and limited detection area, the result shows as Table 3. The average precision of full detection area is 80.97 and with limited detection area is 86.36. The

TABLE 1. Position distribution table of road traffic sign in Taiwan

Video Inform (29 FPS)	Video Total Frames	Frames with Traffic Sign	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
Video 1 (00:02:05)	3,762	170		21	94	42		53	103	19				
Video 2 (00:02:12)	3,969	252			26	44		30	185	62				
Video 3 (00:01:40)	3,003	130			2	23			103	41				
Video 4 (00:02:02)	3,681	97		38	37	8		68	39	16				
Video 5 (00:01:00)	1,796	38							29	14				
Video 6 (00:01:55)	3,468	150		5	67	51		12	80	23				
Video 7 (00:02:07)	3,807	277		9	115	16		57	158	10				
Video 8 (00:02:02)	3,681	1,437	16		16	214	623	776	920	931				
Video 9 (00:01:51)	3,331	74				2			50	31				
Video 10 (00:02:03)	3,704	276			14	17	105	13	99	88				
Total Frames	34,202	2,901	16	73	371	417	728	1,009	1,766	1,235	0	0	0	0
Frames Including Traffic Sign			5,615											
The Ratio of Area Including Traffic Sign(%)			0.285	0.926	4.933	6.679	12.965	17.026	29.617	21.656	0	0	0	0



(a) (b)

FIGURE 5. Road traffic sign position: (a) analyzed road traffic sign location ratio, (b) 12 areas in the frame

average precision improves 5.39 after limited the detection area. The average recall of full detection area is 62.20, and with limited detection area is 62.47. The average recall reduces 0.27 after limited the detection area. The average precision and recall of both road traffic sign detection and identification are shown as Table4. The average precision of full detection area is 97.47 and with limited detection area is 98.01. The average precision improves 0.54 after limited the detection area. The average recall of full detection area is 53.66, and with limited detection area is 53.85. The average recall improves 0.19 after limited the detection area.

Comparing Table 3 and Table 4, the average precision improves after road traffic sign identification, due to the incorrect road traffic sign images being filtered out. The filtered

Table 2. The time impact of limited detection area

Video (720x1080)	Only Traffic Sign Detection		Traffic Sign Detection and Identification	
	Detect with Full Area (ms)	Limited Detection Area(ms)	Detect with Full Area (ms)	Limited Detection Area(ms)
Test Video1	58910.34	42886.16	393628.04	335603.45
Test Video2	66629.92	48932.89	485138.07	451708.19
Test Video3	42116.40	30479.98	160949.06	148535.65
Test Video4	59595.27	43313.22	271322.37	244542.10
Test Video5	30415.87	19496.74	74755.96	60704.16
Test Video6	54453.84	38468.08	292447.03	274324.78
Test Video7	49589.40	33547.52	561443.67	539584.45
Test Video8	60932.55	43482.03	241300.22	236538.28
Test Video9	52857.28	35825.32	248318.23	237904.60
Test Video10	58989.85	39885.84	313302.92	311590.67
Time for per frame(ms)	**15.63**	**11.00**	**88.96**	**83.07**

(a) (b)

Figure 6. The incorrect road traffic sign image filtered result: (a) through road traffic sign detection, (b) after road traffic sign identification

result is shown as Figure 6. The frame was detected by road traffic sign detection, and there were two detected images as Figure 6(a). After road traffic sign identification, the incorrect image in the middle of the frame was filtered out and the road traffic sign identification result also displayed on frame as Figure 6(b) and the identification results are shown as Figure 7.

To integrate the road traffic sign detection with the identification functions, transferring images between OpenCV library and TensorFlow becomes very important. There are two methods to transfer detected road traffic sign images to identification function. One is to save the images detected by Haar Cascade method into temp folder then transfer to the identification classifier trained with TensorFlow, and the other is to transfer the images detected by Haar Cascade with array method to identification classifier trained with TensorFlow. As shown in Table 5, the average FPS is 7.56 if transferring with

TABLE 3. Average precision and recall of comparing with full area and limited detection area for only road traffic sign detection

Test Video (720x1080)	Frames	Detect with Full Area for Traffic Sign Detection			Limited Detection Area for Traffic Sign Detection		
		TP_{Haar}	FP_{Haar}	FN_{Haar}	TP_{CNN}	FP_{CNN}	FN_{CNN}
Video 1 (00:02:05)	**3762**	286	74	163	287	23	162
Video 2 (00:02:12)	**3969**	425	43	163	427	38	160
Video 3 (00:01:40)	**3003**	120	6	70	122	6	68
Video 4 (00:02:02)	**3681**	205	13	153	207	8	151
Video 5 (00:01:00)	**1796**	29	8	7	29	6	7
Video 6 (00:01:55)	**3468**	239	21	153	243	15	149
Video 7 (00:02:07)	**3807**	292	249	121	294	231	119
Video 8 (00:02:02)	**3681**	1305	211	973	1285	115	993
Video 9 (00:01:51)	**3331**	50	56	24	50	48	24
Video 10 (00:02:03)	**3704**	223	65	69	222	10	69
Average Precision (%)		**80.97**			**86.36**		
Average Recall (%)		**62.60**			**62.47**		

FIGURE 7. Recognition result: the left image shows the result of "speed limited 60" and the right image shows the result of "right falling rocks"

saved detected image, and if transferring the detected image with array, the FPS is 11.24. Compare the FPS with above two methods, and the average FPS can be improved 3.68 after replacing the saved image with array.

5. **Conclusions.** In this work, after limited detection area, the ROIs of possible road traffic signs would be quickly detected from the input images, and those ROIs would be transmitted using the image array method to the classifier which was trained by CNNs for identifying the road traffic sign types. The classifier which was trained by Haar-like feature and AdaBoost could quickly identify the features of the road traffic sign based on the difference of lightness and darkness. The classifier which was trained with

TABLE 4. Average precision and recall of comparing with full area and limited detection area for road traffic sign detection and identification

Test Video (720x1080)	Frames	Detect with Full Area for Traffic Sign Detection and Identification			Limited Detection Area for Traffic Sign Detection and Identification		
		TP_{Haar}	FP_{Haar}	FN_{Haar}	TP_{CNN}	FP_{CNN}	FN_{CNN}
Video 1 (00:02:05)	**3762**	260	7	185	261	5	187
Video 2 (00:02:12)	**3969**	324	6	214	265	4	183
Video 3 (00:01:40)	**3003**	107	0	83	111	0	79
Video 4 (00:02:02)	**3681**	185	1	183	187	1	171
Video 5 (00:01:00)	**1796**	23	4	13	24	4	12
Video 6 (00:01:55)	**3468**	219	5	173	222	3	170
Video 7 (00:02:07)	**3807**	270	10	143	272	8	141
Video 8 (00:02:02)	**3681**	1050	30	1228	1054	25	1224
Video 9 (00:01:51)	**3331**	46	5	28	47	3	27
Video 10 (00:02:03)	**3704**	215	2	81	210	1	80
Average Precision (%)		**97.47**			**98.01**		
Average Recall (%)		**53.66**			**53.85**		

GoogLeNet network architecture by TensorFlow library could accurately identify the road traffic sign type. Besides, after the experiments, the average precision can be improved 0.54%. In addition to the methods proposed above, this paper also proposes to define the classifier into two categories according to the property of the road traffic signs. The two categories are the warning sign and the prohibition sign. The signs in two categories can be detected and identified parallelly in the system. At the same time, it is hoped that the sound module can be added into the system and alert the driver of the traffic information around the environment by sound in the future.

Acknowledgment. Ministry of Science and Technology Grant Number: MOST 105-2221-E-218 -024 -

REFERENCES

[1] Bekircan Yavas, Metin Mete Ozbilen and Yasemin Yilmaz, Traffic Speed Sign Recognition with RFID Support, *2014 IEEE 8th International Conference on Application of Information and Communication Technologies* (AICT), pp. 1-4, 2014.

[2] Hengliang Luo, Yi Yang, Bei Tong, Fuchao Wu and Bin Fan, Traffic Sign Recognition Using a Multi-Task Convolutional Neural Network, *IEEE Transactions on Intelligent Transportation Systems, vol*, pp. 1-12, 2017.

[3] C. Papageorgiou, M. Oren, and T. Poggio, A General Framework for Object Detection, *In International Conference on Computer Vision*, pp. 555–562, 1998.

[4] R. Lienhart, A. Kuranov and V. Pisarevsky, Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection, Technical report, MRL, Intel Labs, 2002.

Table 5. Compare the FPS of the saving temp image data methods

Test Videos (720x1080)	Video Length (29 fps)	Frames	Transfer with Saved Detected Images	Transfer with Image Array
Test Video1	00:02:05	3762	540.4237	393.6280
Test Video2	00:02:12	3969	655.5504	485.1381
Test Video3	00:01:40	3003	309.8349	160.9491
Test Video4	00:02:02	3681	426.9080	271.3224
Test Video5	00:01:00	1796	134.7602	74.7560
Test Video6	00:01:55	3468	409.7258	292.4470
Test Video7	00:02:07	3807	730.7865	561.4437
Test Video8	00:02:02	3681	418.6144	241.3002
Test Video9	00:01:51	3331	369.8210	248.3182
Test Video10	00:02:03	3704	526.4013	313.3029
Average FPS			**7.56 FPS**	**11.24 FPS**

[5] P. Viola, and M. Jones, Rapid Object Detection Using a Boosted Cascade of Simple Features, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 511–518, 2001.

[6] Khan S., Rahmani H., Shah S. A. A. and Bennamoun M., A guide to convolutional neural networks for computer vision, *Synthesis Lectures on Computer Vision,* vol. 8(1), pp. 1-207, 2018.

[7] Krizhevsky A., Sutskever I., and Hinton G. E., Imagenet Classification with Deep Convolutional Neural Networks, I*n Advances in neural information processing systems*, pp. 1097-1105 ,2012.

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, Imagenet Classification with Deep Convolutional Neural Networks, Advances in neural information processing systems, pp. 1097.

[9] C. Szegedy, W. Liu, Y. G. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich, Going Deeper with Convolutions, 2015 *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, 2015.

[10] K. Simonyan and A. Zisserman., Very Deep Convolutional Networks for Large-Scale Image Recognition, arXiv preprint arXiv: 1409.1556, 2014.

[11] Joseph Redmon, Santosh Divvala, Ross Girshick and Ali Farhadi, You Only Look Once: Unified, Real-Time Object Detection, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779-788, 2016.