

# Studies advanced in traffic sign detection in autonomous driving scenarios

Hao Li<sup>1,†</sup>, Yifei Ma<sup>2,†</sup>, Jinghao Yu<sup>3,5,†</sup>, Ziyu Zhang<sup>4,†</sup>

<sup>1</sup>Chengdu College of University of Electronic Science and Technology of China, Sichuan province, 610000, China

<sup>2</sup>Xi'dian University, Xi'an, Shanxi province, 710126, China

<sup>3</sup>Hebei University of Science & Technology, Shijiazhuang, Hebei province, 050091, China

<sup>4</sup>Zhengzhou Foreign Language School, Zhengzhou, Henan province, 450007, China

<sup>5</sup>Corresponding author: jinghaoyu@students.federation.edu.au

<sup>†</sup>These authors contributed equally.

**Abstract.** Traffic sign detection and recognition is a hot research topic in the environmental awareness module of autonomous driving, which aims to build a model of road traffic information and provide a decision basis for driving scheme design. Early traffic sign recognition methods were mostly based on color features, shape features, or multi feature fusion. Thanks to the rapid development of convolutional neural networks, traffic sign recognition methods based on deep learning have continuously made breakthroughs in both accuracy and speed. In this paper, focusing on the four major types of mainstream frameworks mentioned above, we will introduce the latest research progress in traffic sign detection and recognition technology, including the principles, steps, advantages and disadvantages of representative algorithms. In addition, we quantitatively compared the performance of different recognition methods on common data sets. Finally, we discussed the existing problems in traffic sign recognition and its future development direction.

**Keyword:** traffic sign detection, autonomous driving, deep learning.

## 1. Introduction

Thanks to the rapid development of various hardware sensors and pattern recognition methods, autonomous driving technology has gradually matured and started commercial applications. Environmental awareness, especially road information awareness, is the key module of autonomous driving technology, which mainly includes subtasks such as road recognition and detection, vehicle detection, pedestrian detection, traffic sign detection. As a kind of symbolic information specifically set up for road driving, traffic signs can remind drivers of road information such as speed limits, road conditions, and road guidance. In this case, assisted driving systems or autonomous driving systems must detect and recognize the information of traffic sign efficiently, attracting a large amount of research attention from academia and industry.

The key step in the accurate and rapid recognition of traffic signs is traffic sign detection, which is to discover and locate traffic signs from complex scenes. Due to the variability of the actual application environment, the accuracy of traffic sign detection will directly affect the recognition accuracy. Specifically, lighting conditions are difficult to control, which means there are significant differences in exposure, reflection, and dimming of images because of varying brightness in different weathers, seasons, and backgrounds. On the one hand, some signs may have faded, blurred, damaged surfaces, and other phenomena. On the other hand, the captured images may appear blurred, artifacts, and ghosting due to the vibration and fluctuation of the vehicle during movement, especially when the road surface is uneven. The above issues make traffic sign detection extremely challenging.

Early traffic sign detection methods mainly relied on image color features, which can be well distinguished from the surrounding environment because traffic signs are usually colored in bright and visible colors. The color based traffic sign detection method is to divide the captured image into subsets of similar color attributes, and then segment it by processing color thresholds to extract traffic signs. However, color feature based detection methods are easily affected by weather and light. Further, some scholars consider the specific shapes of traffic signs such as simply geometric figure and extract shape features to achieve detection of traffic sign. The shape based traffic sign detection method first finds its contour, and then makes decisions based on the number of contours. Hough transform method that isolates specific shape features in a given image is a common method of detection. In addition, other shape based detection methods include distance transform matching and edge detection. Shape based detection methods are more sensitive to occlusion, and when the traffic signs to be detected have partial occlusion, the accuracy decreases sharply. In order to make up for the shortcomings of the above two methods, some scholars pay more attention on the color and shape features, and the accuracy of the traffic sign detection method based on multiple feature fusion has been further improved.

With the development of deep learning, traffic sign detection and recognition methods based on convolutional neural networks has been developed. The deep learning model can automatically extract features without manual design, which can avoid the limitations of manual design features. Focusing on the four mainstream frameworks mentioned above, this article introduces the research progress of traffic sign detection based on color, shape, multi feature fusion, and deep learning, including the steps, advantages, disadvantages, and performance of representative algorithms. What's more, the problems that need to be solved in traffic sign detection and recognition based on deep learning are analyzed, and its future development trend is prospected.

## 2. Method

### 2.1. Color-based traffic sign detection

As we all know, traffic signs are usually set with special colors easy to find color, which can bring drivers or some unmanned monitoring equipment relevant road information in real time to assist them to drive safely. The common traffic signs such as no traffic, no parking and left turns are shown in Figure 1.



**Figure 1.** Common traffic signs.

Therefore, the intelligent detection system is very sensitive to color, and the most important thing for color-based traffic sign detection is to divide the color of the color [1]. Color segmentation is under a specific color space model, according to the segmentation scale is divided into small areas, that is, objects, and then the small objects are merged into large objects through regional object characteristics and relative position information. The most basic three colors space models are RGB, HIS, HSV model, and the latter two models are improved and strengthened based on the first model. The following is a brief introduction to the three models:

(1) RGB model. RGB model[2] is a commonly used color information expression, which uses the brightness of the three primary colors of red, green, and blue to quantitatively represent the color. This model also known as the additive color mixing model, is a method of superimposing RGB three-color light on each other to achieve color mixing.

Using the RGB model for color segmentation, it is characterized by its suitability for hardware devices, fast speed, and small amount of calculation. The relative disadvantage is that it does not simulate human visual perception of color well. Since the image will be affected by different lighting when actually shooting, and the definition of RGB color needs to be considered, the image must be preprocessed for color segmentation using the RGB model in order to select a unique threshold for any image.

(2) HIS model. The HSI model [3] describes its color characteristics with three parameters, H, S, I, where H defines the frequency of the color. S represents the depth of the color and I means intensity or brightness. The HIS model is based on the human visual system to depict some colors with color, color saturation and brightness.

(3) HSV model. The HSV model [4] is like an enhanced version of the HIS model, which is closer to human perception of color than the HIS model. H represents hue, S represents saturation, and V represents luminance value, the coordinate system of the HSV model can be a cylindrical coordinate system, but it is generally represented by a hexagonal pyramid, so it is also called a hexagonal cone model, which is very similar to the HIS model.

## 2.2. Shape-based traffic sign detection

In addition to the color, another distinguishing feature of traffic signs is their shape, because they are all designed with special geometric figures, such as circles, triangles, squares, diamonds, etc. Therefore, many researchers use shape features as a means of detecting traffic signs. Shape-based traffic sign detection is less affected by lighting factors, and it is generally judged whether it is a traffic sign by its shape characteristics and the symmetry of the area of interest [5]. The most common method of shape detection is to use some form of Hof transform [6] and a Histogram of Oriented Gradients (HOG). The Hough transform is a method of finding special shape features, which can detect simple geometries such as circles or straight lines, and is mainly used in the detection of speed limit signs. However, its calculation is too cumbersome and does not occupy a dominant position in the detection of traffic signs that require high real-time.

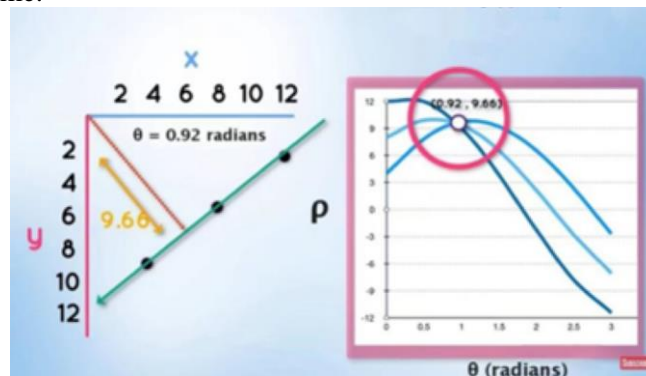


Figure 2. Hough transform.

### 2.3. Traffic sign detection based on multi-feature fusion

Whether it is color-based traffic sign detection or shape, they will be affected by external factors and produce large errors, so relying on a single detection method may lead to failure. Therefore, the fusion of multiple features such as color and shape is more conducive to the detection of traffic signs [7-8]. In the stage of traffic sign recognition, corresponding to the traffic sign to be classified and identified, we know its color, shape and other characteristics, and we mainly have to distinguish the small differences inside the pattern.

### 2.4. Traffic sign detection based on deep learning

Whether it is traffic sign detection based on color or shape, it is ultimately a traditional machine learning detection method controlled by humans. Although traditional machine learning has the advantages of high detection accuracy and strong real-time performance, it also relies on manual design, so it requires researchers to have rich experience, so this method also has certain limitations. The reason why deep learning is better than traditional machine learning in some aspects is that it can spontaneously extract features without manual control, so it can better avoid the limitations of manual control.

The detection based on deep learning is mainly divided into two types: two-stage object detection algorithm and first-stage object detection algorithm, and the two-stage object detection algorithm mainly represents the network R-CNN series, SPP-net, etc.; Typical one-stage detection algorithms include SSD and its improved algorithms, yolo series, RetinaNet, RefineDet and other networks [9].

## 3. Experiment

Road traffic signs are road facilities used to convey road conditions and traffic rules to drivers, using combinations of different colors, shapes, and patterns to express different information, and have significant color, shape, and scale features. In order to evaluate the performance of different traffic sign detection methods, this paper quantitatively compared the results of representative algorithms on the same dataset, as shown in Table 1. The algorithm used in this experiment conducted by Sichuan University [10] has been experimentally verified to be feasible and effective, with good performance in both detection and recognition. The purpose of the experiment is to demonstrate the superiority of the multi-feature fusion recognition method in the accuracy of identifying road signs by recording the accuracy of different methods for identifying road signs. It can also be seen that the advantages of multi-feature fusion recognition are quite obvious. In this experiment, in addition to the above two features, feature vectors were also fused for training the classification detection model, and single feature recognition experiments were performed for comparison. Further improvements in recognition robustness can be achieved by adding image restoration methods, such as dealing with the problem of faded or slightly deformed road signs mentioned in the single feature recognition experiment above. However, due to the fusion of multiple features, the complexity and time consumption of this method will be much higher than that of single feature recognition, so the selection of this method needs to be further judged based on actual conditions.

**Table 1.** Experiment results of different traffic sign detection algorithms.

	group 1	group 2	group 3	group 4	group 5
Sample space (test samples)	500 (80)	500 (80)	500 (80)	500 (80)	500 (80)
Test samples number (segmented area images)	80 (272)	80 (272)	80 (272)	80 (272)	80 (272)

**Table 1.** (continued).

Test method	Color feature and SVM	Texture feature and SVM	Shape feature and SVM	Multi-vector space accumulation and SVM	Multi- feature fusion and SVM
Number of correct areas of the test sample	248	251	242	256	264
Accuracy of detection	91.2%	92.28%	88.97%	94.11%	97.01%

#### 4. Discussion

In the current research, there already are some algorithms with higher accuracy, but they are all built on the standard database, when applied to the actual traffic environment, the collection of traffic signs is easily affected by natural scenes, and the factors affecting the accuracy mainly include the following aspects:

- (1) Natural environment: weather conditions and light intensity will change at different times, such as sunny days, rain and snow weather temperature and light intensity will change greatly, affecting the visibility of traffic signs;
- (2) Changes in traffic signs: After prolonged exposure to sunlight, traffic signs are prone to noticeable fading. If the red fades and turns into light white, it will lead to inaccurate identification;
- (3) Occlusion: In the actual environment, there may be traffic signs that are obscured by pedestrians or obstacles, making it impossible to detect or recognize them;
- (4) Blurred signs: The shaking that occurs during high-speed driving of the vehicle can easily cause blurred images collected.

Among the various technologies that currently exist for traffic sign detecting, the influence in the actual environment cannot be completely overcome. For color-based traffic sign detection, when the RGB color model method is used in the real traffic environment, the traffic sign may be mixed with background noise in turns, which cannot achieve good results. The HIS algorithm requires better hardware processing to improve real-time performance. For shape-based traffic sign detection, the direction gradient histogram (HOG) method is more common, which has the advantages of rotation and scale invariance, but the calculation is too large and the process is cumbersome. For traffic sign detection based on multi-feature fusion, the color and shape multi-feature fusion can improve the detection accuracy, but there are still defects, such as Tang Kai and his team proposed a method based on color, shape, scale and other multi-feature, this method is for the extracted closed contour curvature histogram code for scale normalization processing, but the small-scale curvature histogram is easily affected by edge noise, which will make it quite difficult to detect traffic signs with smaller scales. For the traffic sign detection based on deep learning, the accuracy and speed are greatly improved compared with traditional methods, but there are still shortcomings. For RCNN, some algorithm optimization can improve accuracy and stability of the model in practical application scenarios, but it still needs to be further improved for the actual and changeable traffic environment. For the YOLO algorithm, the model has high requirements for the computing power of the experimental platform, and its real-time performance cannot be guaranteed under the condition of improving the accuracy. For the SSD algorithm, the current algorithm optimization can solve the recognition of small targets with low light and long distance, but the detection speed still needs to be further improved.

Compared with traditional methods, deep learning has more and more obvious advantages, and how to achieve the compression of neural network scale without affecting accuracy is the focus of research, and with the deepening of convolutional neural network level, gradient propagation will become more difficult, how to design algorithms to help gradient propagation is an important research content. It can be seen that the detection of traffic signs based on deep learning still has a lot of room for improvement.

With the rapid development of computer technology, algorithms and computing power will continue to improve, so as to feedback the continuous development of deep learning methods.

## 5. Conclusion

Starting from the detection technology of traffic signs, this paper introduces the traditional target recognition method and the target recognition method based on deep learning, and sorts out the advantages and disadvantages of the method. At present, the deep learning algorithm with great advantages is bound to become the mainstream of traffic sign detection methods, and scholars are paying more attention to creating lightweight network structures to achieve a more advanced detection, such as replacing the single Resenet53 residual network in the YOLO v3 network structure with the dense series mechanism of Dense Net to ensure the detection rate while increasing the speed. From the perspective of technical category, the first-stage algorithm needs to find a more reliable way to filter out redundant information in the network; The two-stage algorithm needs to pay more attention to finding excellent backbone networks to extract features with more learning ability, optimize the region, propose that the network has completed accurate screening, and optimize the network structure at the regional level to improve its speed. In short, the current stage of traffic sign detection research always needs to find a balance between real-time and accuracy.

## References

- [1] LU Shengrong, LIU Lifeng, LI Cuihua. 2011. J. Traffic sign detection based on color segmentation and multi-feature fusion. Journal of Xiamen University: Natural Science Edition, 2011, 50(4): 685-689.
- [2] ZHANG Rui, ZHANG Peng. 2010. J. Traffic sign detection and classification based on color and shape. Microcomputer information, 2010(35): 226-228.
- [3] WANG Bingfei, SONG Haiyu, LI Houjie, WANG Peichang, MA Biao. 2015. J. Research on traffic sign extraction method based on color model and regional characteristics. Journal of Dalian University for Nationalities, 2015, 17(3): 274-277.
- [4] ZHANG Zhijia, LI Wenqiang, ZHANG Dan, ZHONG Ling. 2015. J. Traffic sign detection method based on color and shape features. Computer technology and development, 2015, 25(7): 174-178.
- [5] XU Guoyan, WANG Chuanrong, GAO Feng, WANG Jiangfeng. 2010. J. Vehicle video detection region of interest determination algorithm. Journal of Beihang University, 2010(7): 781-784.
- [6] YE Zhouhai, CHEN Fumin. 2006. J. Improvement of a generalized Hough transform. Microcomputer application, 2006, 22(6): 23-251
- [7] WU Xiaohui, TIAN Qichuan. 2020. J. Review of traffic sign recognition methods. Computer Engineering and Application, 2020, 56(10): 20-26
- [8] TANG Kai, LI Shiyong, LIU Juan, LI Renfa. 2015. J. Traffic sign detection based on multi-feature collaboration. Computer Engineering, 2015, 41(03): 211-217
- [9] LI Kejian, CHEN Shaobo, LI Wanqi. 2020. J. Simulation of traffic sign detection system based on deep learning. Software Guide, 2020, 19(9): 31-34
- [10] ZHU Guokang, WANG Yunfeng. 2011. J. Traffic sign detection based on multi-feature fusion. Signal Processing, 2011, 27(10): 1617-1620