# Traffic Sign Detection Based on Deep Learning Methods

Xiqing Huang[1](✉), Fan Wang[2], Shibin Yang[3], and Hongfei Zhang[4]

[1] Science in Information Technology, University of Technology Sydney, Sydney 2007, Australia
xiqing.huang1@student.uts.edu.au

[2] College of Information and Computer Engineering, Northeast Forestry University, Harbin 150040, Heilongjiang, China

[3] Faculty of Engineering, Architecture, and Information Technology, University of Queensland, St Lucia, QLD 4072, Australia
shibin.yang@uqconnect.edu.au

[4] School of Information Science and Engineering, Jinling College, Nanjing University, Nanjing, Jiangsu, China

**Abstract.** This article gives a brief overview of a few current research on traffic signs detection, which briefly reviews the concept and structure of traffic signs detection in the last decade. The methodology varies in different ways which is generally separated into two exact dimensions. The first one is the traditional method using the theory of computer vision with machine learning to detect the traffic signs, while the other one uses deep learning to train the model to detect the objects. In recent years, the methods based on deep learning have gradually replaced the traditional methods since they can extract features from traffic signs better and do predictions. Therefore, this paper mainly focuses on the deep learning methods for traffic signs detection and reviews previous work and their corresponding datasets and performance. The results based on different methods are compared. Finally, we made a conclusion based on this review.

**Keywords:** component · Deep learning · Traffic sign detection · CNN

## 1 Introduction

In recent years, with the rapid popularity of driverless driving, a large number of cars have been equipped with driverless technology, so many new technologies related to driverless technology are constantly improving. In general, most driverless technology will rely on the detection of traffic signs to ensure that vehicles operate safely and smoothly on the road. Besides, traffic target detection is also an important part of advanced driver-assistance systems, which helps drivers to make decisions and improve driving safety and comfort. To sum up, detection for traffic signs is a necessary and promising direction, which deserves much attention in the cases mentioned above.

X. Huang, F. Wang, S. Yang and H. Zhang—Contributed equally.



Z. Zhan et al. (Eds.): ICBDIE 2022, AHCS 5, pp. 700–708, 2023.
https://doi.org/10.2991/978-94-6463-034-3_72

Many methods have been proposed for detecting and classifying traffic signs in the last decade. Previously, traffic signs were detected using a classical method based on the distinctive shape or color of the traffic signs. It is a widely used method to distinguish regular traffic signs because they have bright colors and neat shapes such as triangles or circles that are easily recognized. Such methods have been widely applied to the detection of traffic signs before, however, the shapes of traffic signs are easily distorted due to being affected by complex occlusion of the road in the real application. To solve this limitation, there has been a variety of works, such as [1], proposed some related solutions to detect multiclass partial occluded traffic signs. However, the detection of traffic signs still faces other challenges that traditional methods are hard to overcome. For example, the performance of the detection will be easily affected by extreme weather such as rain, snow, or different lighting patterns. Therefore, limitations of traditional methods reduced the interest in traffic sign detection based on shape and color has waned over the years. Instead, new methods such as deep learning have become particularly important in recent years.

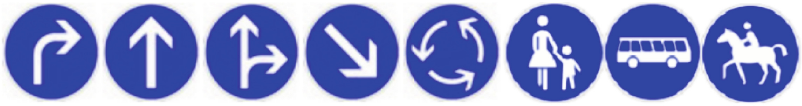
Machine learning-based methods to detect traffic signs have also been a focus in recent years. The "Viola-Jones" detector is a cascade classifier trained with AdaBoost, combining integral images, which improves object detection performance. However, machine learning methods have limitations when faced with realistic and complex traffic scenarios. Deep learning is an algorithm in machine learning that is based on learning representations of data that can better extract image features. It has also been used frequently in recent years to process far more data than ever before. The development of Convolutional Neural Networks (CNN) has greatly improved the accuracy of traffic sign detection. For example, YOLO (You Only Look Once) is a real-time object detection method that was proposed to detect traffic signs quickly and provide drivers with more reaction time. In addition, Baojun Zhang et al. designed a four-scale detection structure based on the three-scale detection structure of the YOLOv3 [1] network, aiming to enhance the recognition rate of long-distance small targets.

In this paper, the progress of the traffic sign detection will be reviewed, and it mainly focuses on some important aspects including comparing the strengths and weaknesses with a different model proposed in the last decade.

The structure of this paper is organized as follows. First, different common datasets of traffic detection and methods will be integrated and introduced in Sects. 2 and 3. Then, Sect. 4 and 5 introduce the typical measurements for this field and compares the detection results obtained by different experimental methods. Finally, in Sect. 6, this paper makes a summary of the whole paper.

## 2 Dataset

The dataset for traffic sign detection should consider as many situations as possible so that the model can be more robust and work well in a variety of situations. Although some teams [2] used their local datasets, most of them [3–5] used datasets called GTSDB German data set, tt100k data set, and CCTSDB dataset for the training the proposed algorithms or models. In the following subsections, these datasets that are usually used in this case are introduced in detail to illustrate their characteristics and advantages. Figure 1 and Fig. 2 show the sample images for these datasets.



(a) Mandatory

(b) Danger


(c) Prohibitory

**Fig. 1.** Demo of GTSDB dataset.

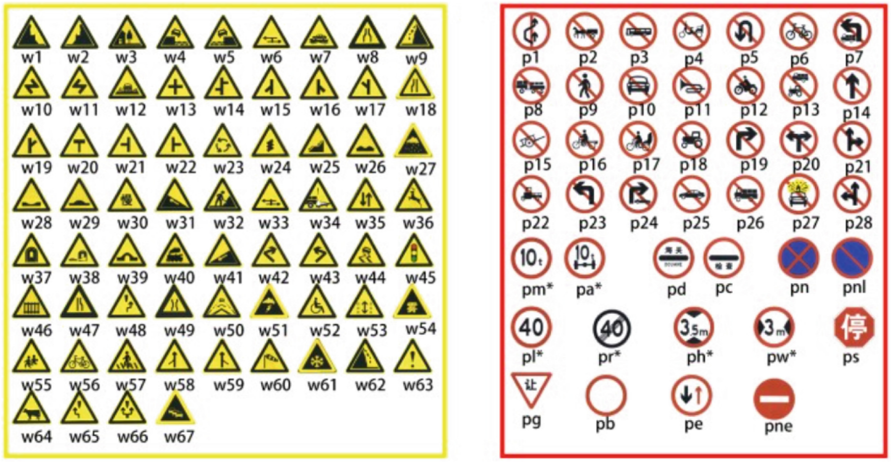



**Fig. 2.** Demo of TT100k dataset.

### 2.1 GTSDB

German Traffic Sign Detection Benchmark is a traffic sign dataset in Germany that contains up to 900 images that are in 136 $\times$ 800 pixels, which are divided into three categories [6].

### 2.2 TT100k

This dataset was created by researchers from Tsinghua University from 100,000 Tencent street views, providing 100,000 pictures, including 30,000 traffic sign examples, covering the huge changes in light and weather conditions, and many of the traffic signs are only $20 \times 20$ pixels in size and occupy a small space [7].

### 2.3 CCTSDB

Chinese Traffic Sign Detection Benchmark is produced by Zhang et al. [8]. It contains 15,734 pictures. The marked data is divided into three categories: indication signs, prohibition signs, and warning signs.

## 3 Method

Essentially speaking, traffic sign detection is also one kind of target detection. Before machine learning was introduced into target detection, most researchers used the traditional image processing algorithm for detection. It mainly includes three parts: first, extract the model from the candidate area, and then extract the features of the candidate area, such as (HOG), (SIFT), and (ORB) features, etc. Finally, the image is classified, and the position of the candidate box is corrected. A famous algorithm called Support vector machine (SVM) is often used in this stage [5].

However, the traditional detection algorithms have great limitations, their detection accuracy is not high, and the detection speed is usually slow. It can only achieve good results in some specific scenes. In face of some interference, especially when the traffic signs are small, the illumination is insufficient, or the signs are blocked, false detection would be very easy to appear.

Generally, there are several methods based on deep learning methods for traffic sign detection in various situations.

### 3.1 Multi-scale Problem

For the multi-scale problem of traffic signs detection, the method of Data Augmentation is usually adopted. This method is used to solve the problem of insufficient training data. Data Augmentation increases the number of training graphics available by using slight changes in translation, rotation, etc. [4].

### 3.2 Occlusion Problem

For the occlusion problem causes by a complex environment, many relevant solutions have been proposed. For example, the Haze Removal technique using the Retinex theory has been adopted and well used in traffic sign detections. Other Haze Removal techniques or Data Augmentation are also used to deal with this problem [9].

### 3.3 Brightness Problem

For the detection problem causes by illumination changes, the methods using shape analysis are better than the methods using color segmentation, since the change of light usually makes the color differentiation not obvious, and there are only several shapes of traffic signs such as triangle or square. Furthermore, Data Augmentation is still a good method. It is used to increase the data of different lighting degrees in real situations when training the neural network [10, 11].

## 4 Measurements

In the detection of traffic signs, most researchers use a few indicators: Precision, Recall, ACC, FPS. The specific definition of these indicators can be found below.

### 4.1 FPS

Frame Per Second is used to evaluate the speed of target detection. It is used to predict the proportion of correct samples.

### 4.2 ACC

It is used to predict the proportion of correct samples, which can be found in Eq. (1).

$$ACC = \frac{TP + TN}{N} \tag{1}$$

### 4.3 Precision

It is used to show among all the predicted objectives, the proportion of correctly predicted, which can be found in Eq. (2).

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

### 4.4 Recall

It is used to predict the ratio of the correct target box to the ground truth box, which can be found in Eq. (3).

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

### 4.5 F1Score

In addition to these criteria, there are other criteria, such as the F1 score, used to locate the harmonic average of the accuracy and recall, which can be found in Eq. (4).

$$F1Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} = \frac{2TP}{2TP + FN + FP} \tag{4}$$

### 4.6 mAP

Besides, mAP (mean average precision) as a typical indicator is also employed to measure the average quality of all categories. In this Eq. (5), the AP value is used to calculate the area under the P-R curve, it is the most commonly used evaluation standard for target detection.

$$mAP = \frac{\sum_{q=1}^{Q} AP(q)}{Q} \tag{5}$$

## 5 Result and Discussion

Table 1 compares various methods using deep learning to detect the traffic signs and their results. The Yolo model exceeds Faster R-CNN in its third generation which is YoloV3 in both FPS and accuracy. However, as vividly shown in the table, with using structures (e.g. FPN, SPP), the improved YoloV3 method provides a more precise and faster performance. It is worth mentioning that, by using Integrated CapsNet based on the bagging, Qu et al. reach the accuracy of 97.86% which tops the rank list in terms of accuracy.

To further compare the performance of different methods based on the same baseline. Table 2 illustrates the variation of results of both traditional methods as well as the deep learning methods when using the same dataset as the training sample. It is obvious that the

**Table 1.** Performance based on different methods for traffic sign detection.

| Method | Dataset | FPS | Accuracy | mAP |
|---|---|---|---|---|
| YoloV3 with improved FPN and SPP [12] | TT100K | 31.3 | - | 75.2% |
| CNN + MBSP + RPN [13] | TT100K | - | - | 87.14% |
| Faster R-CNN [14] | TT100 K | 3.6 | 50.3% | - |
| YOLOv1[15] | TT100 K | 41 | 16.9% | - |
| YOLOv1-tiny | TT100 K | 54.9 | 11.5% | - |
| YOLOv2[16] | TT100 K | 31.2 | 19.8% | - |
| YOLOv3[1] | TT100 K | 25.6 | 52.9% | - |
| YOLOv3 + fourth detection layer [3] | TT100 K | 13.3 | 54.2% | - |
| Improved YOLOv3 (YOLOv3 + fourth detection layer + soft-NMS) [3] | TT100 K | 13.3 | 54.3% | - |
| SSD + MRFeature + VSSA [17] | OPTTSR | 21 | - | 53.45% |
| Integrated CapsNet based on the Bagging [18] | LISA + Belgium TS | - | 97.86% | - |

**Table 2.** The performance on the same dataset using different methods.

| Method | Dataset | FPS | Accuracy | mAP | Dataset | FPS |
|---|---|---|---|---|---|---|
| Cascaded + R-CNN [19] | GTSRB | - | 88.60% | - | 96.80% | 92.52 |
| CascadedR-CNN + Multiscale-Attention [4] | GTSRB | - | 90.50% | - | 98.70% | 94.42 |
| Integrated CapsNet based on the Begging [18] | GTSRB | - | - | 99.07% | - | - |
| Multiscale Attention + Cascaded R-CNN [4] | GTSDB | - | 83.62% | - | - | 90.82 |
| color probability model + color HOG + CNN [5] | GTSDB | 6 (overall) | - | 98.24% | - | - |
| color probability model + color HOG + CNN [5] | CTSD | 6 (overall) | - | 98.77% | - | - |

method using deep learning overperforms a lot in comparison to the traditional methods (e.g. the method using AdaBoost, Haar wavelet feature, and Bayesian classification).

What is more, from the Table 2, we can safely conclude that the addition of multiscale attention does benefit improving the accuracy of the model. For example, with multiscale attention added, the cascaded R-CNN reaches an accuracy of 90.5% which is about 2% than the initial one.

## 6 Conclusion

In this paper, several commonly used methods were introduced for traffic signs detection problems, especially approaches focusing on deep learning. Some recent articles that address the problems associated with the complex task of traffic sign detection such as brightness, occlusion due to complex weather conditions, etc. were also reviewed. This paper summarizes the most popular methods in recent years and compares the strengths and weaknesses of each model with the same indicators. In the future, applying these methods to other detection tasks will be considered in further study.

## References

1. J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," 2018.
2. A. Aggar, A. A. Rahem, and M. Zaiter, "Iraqi Traffic Signs Detection Based On Yolov5," in 2021 International Conference on Advanced Computer Applications (ACA), pp. 5–9, 2021.
3. B. Zhang, G. Wang, H. Wang, C. Xu, Y. Li, and L. Xu, Detecting Small Chinese Traffic Signs via Improved YOLOv3 Method, Mathematical problems in engineering, pp. 1–10, 2021.
4. J. Zhang, Z. Xie, J. Sun, X. Zou, and J. Wang, "A Cascaded R-CNN With Multiscale Attention and Imbalanced Samples for Traffic Sign Detection," IEEE access, vol. 8, pp. 29742–29754, 2020.

5. Y. Yang, H. Luo, H. Xu, and F. Wu, "Towards Real-Time Traffic Sign Detection and Classification,"IEEE transactions on intelligent transportation systems, vol. 17, no. 7, pp. 2022–2031, 2016.
6. S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing and C. Igel, "Detection of traffic signs in real-world images: The German traffic sign detection benchmark," The 2013 International Joint Conference on Neural Networks (IJCNN), pp. 1–8, 2013.
7. J. Zhang, M. Huang, X. Jin, and X. Li, "A real-time Chinese traffic sign detection algorithm based on modified YOLOv2," Algorithms, vol. 10, no. 4, p. 127, 2017.
8. J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition," Neural networks, vol. 32, pp. 323-332, 2012.
9. C. Liu, F. Chang, and C. Liu, "Occlusion-robust traffic sign detection via cascaded colour cubic feature," IET intelligent transport systems, vol. 10, no. 5, pp. 354–360, 2016.
10. M. Garcia-Garrido, M. Sotelo, and E. Martin-Gorostiza, "Fast traffic sign detection and recognition under changing lighting conditions," in 2006 IEEE Intelligent Transportation Systems Conference, pp. 811–816, 2006
11. W. Shadeed, D. Abu-Al-Nadi, and M. Mismar, "Road traffic sign detection in color images," in 10th IEEE International Conference on Electronics, Circuits and Systems, 2003. vol. 2, pp. 890–893, 2003.
12. Z. Liu, L. Yuan, M. Zhu, S. Ma, L. Chen, "YOLOv3 Traffic sign Detection based on SPP and Improved FPN," Computer Engineering Applications, pp.164–170, 2021.
13. Z. Liu, C. Shen, X. Fan, et al. "Scale-aware limited deformable convolutional neural networks for traffic sign detection and classification." IET Intelligent Transport Systems, vol.14, pp.1712-1722, 2020.
14. S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems", vol.28, 2015.
15. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779–788, 2016.
16. J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," in Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 6517–6525, Honolulu, HI, USA, July 2017.
17. Y. Yuan, Y. Xiong, and Wang, Q, "VSSA-NET: Vertical spatial sequence attention network for traffic sign detection." IEEE transactions on image processing, vol. 28, pp.3423–3434, 2019.
18. Z. Qu, L. Wang, Y. Shao,et al, "A CapsNet Integration Algorithm for Multi-View Traffic Sign Recognition", Journal of Chongqing University of Technology (Natural Science), vol. 34, pp. 160-168, 2020.
19. Z. Cai and N. Vasconcelos, "Cascade R-CNN: Delving into high quality object detection," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Salt Lake City, UT, USA, pp. 6154-6162, 2018.