Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 5
AIDML 2024

Advancements in Traffic Sign Recognition and Detection: Harnessing the Power of Intelligent Systems

Jiaxi Huang [1, *], Yaokun Dong [2] and Yichen Zhang [3]

[1] College of Physics and Electronic Information Engineering, Zhejiang Normal University, Jinhua, China

[2] Wuhan Experimental School, Wuhan, China

[3] Sino Canada School, Suzhou, China

* Corresponding Author Email: huangjiaxi031016@zjnu.edu.cn

Abstract. This paper delves into the current research landscape of traffic sign recognition and detection, aiming for a comprehensive understanding of the field's developmental trajectory, significant achievements, and persistent enigmas. This is achieved through an extensive review of pertinent literature from both domestic and international sources. Within the research backdrop, it explores the diverse methodologies and challenges encountered in traffic sign detection and recognition, with a specific emphasis on algorithms powered by deep learning. The paper provides an in-depth analysis of the notable advancements these deep learning-based algorithms have achieved in recent years, marking a significant stride in the field. Furthermore, it scrutinizes the implications of these advancements for the future of automated and intelligent transportation systems. Building on this analysis, the paper forecasts future research directions, intending to offer a referential and guiding framework for subsequent, more profound investigations. This comprehensive approach not only showcases the current state of the art but also lights the path for future explorations that could revolutionize the landscape of traffic sign recognition and its application in autonomous driving technologies.

Keywords: Traffic sign recognition; Deep learning; Convolutional neural networks.

1. Introduction

In recent years, the comprehensive optimization and advancement of intelligent transportation systems have increasingly highlighted the role of artificial intelligence in vehicular safety. Traffic signs, with their conspicuous, standardized, and clear text, graphics, and symbols within a certain range, have consistently been the primary focus of image recognition technology in the field of traffic. Since the inception of image recognition technology, traffic sign recognition algorithms and models have garnered widespread attention [1]. This technology has matured significantly, finding applications in autonomous driving. Despite the existence of numerous recognition methods, the challenge of detecting and recognizing traffic signs with higher accuracy and speed remains a primary concern in current research [2]. The evolution of machine vision and strides in image processing technology have enhanced the significance of traffic sign recognition in transportation, offering substantial practical value for the advancement of intelligent transportation.

Traffic sign detection and recognition typically involve segmentation positioning and recognition classification. Generally, traffic signs are extracted from natural scenes by utilizing their shape and color features. The recognition classification aims to identify the content of the detected traffic signs. Presently, deep learning algorithms like convolutional neural networks demonstrate high accuracy and robust anti-interference capabilities in traffic sign recognition applications. They represent a key breakthrough in overcoming the reliability issues of automatic traffic sign recognition [3]. Recent developments in machine learning have significantly transformed image recognition. By setting simple judgment conditions and enriching the feature library with learning samples in various forms and scenes, the system progressively enhances its cognition and recognition probability of traffic signs. Machine learning enables recognition to transcend reliance on specific fixed parameters, allowing the system to identify targets through a series of conditional judgments, thereby elevating the accuracy and flexibility of recognition [4]. This burgeoning field has become a focal point of research, significantly refining image recognition accuracy. Against this backdrop, this paper reviews traffic sign recognition technology based on deep learning, outlining several key steps in the recognition process. It highlights the advantages of deep learning-based traffic sign recognition over other methods, and analyzes the substantial progress made in recent years. This review aims to offer fresh perspectives for the development of traffic sign recognition technology and contribute practical insights for the construction of intelligent transportation systems.

2. Relevant Theories

According to the categories, traffic signs can be divided into warning signs on a yellow background with black edges, prohibition signs on a white background with red circles, and indication signs on a blue background with white words. The shapes are mainly triangles, circles and rectangles. The clear shape color distinction and the limited number of signs provide a relatively stable application environment for image recognition [5].

As shown in Figure 1. Traffic sign detection and recognition using image recognition technology is generally divided into the following working steps:

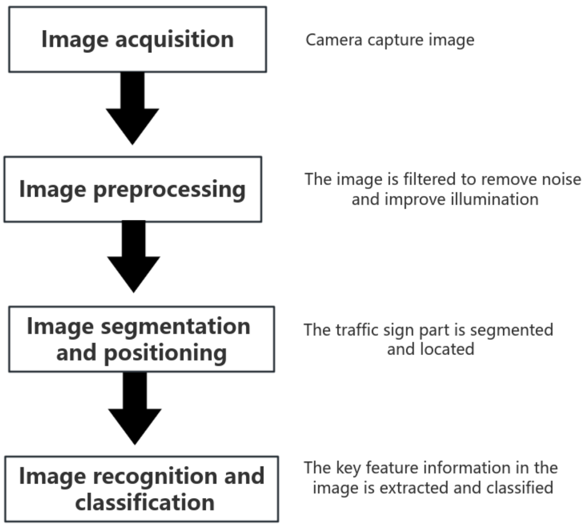

Fig. 1 Traffic sign detection process (Photo/Picture credit: Original).

2.1. Pre-processing procedures

Preprocessing methods include histogram equalization, white balance, brightness adjustment and so on. In the actual traffic scene, due to the influence of various factors such as motion jitter, natural light, weather conditions and so on, it is inevitable that a certain degree of interference and noise will be introduced into the collected images [6]. Therefore, it is necessary to preprocess the collected images to eliminate these adverse factors. Through image equalization, image enhancement and image denoising algorithms, the light of the image is equalized, so that the image becomes clear and easy to identify, so as to highlight the key information. The pre-processing is good for feature extraction and classification, which greatly improves the timeliness and accuracy of detection, and reduces the complexity of the following problem. Image preprocessing often determines the final effect to a certain extent.

2.2. Traffic Signs Segmentation and Positioning

The preprocessed image still contains a lot of information, and traffic signs only account for a very small part of it. In order to reduce the amount of data processed and speed up the processing speed, the area of the traffic sign is generally detected and located first, and then the specific meaning of the traffic sign in this area is judged [7]. In order to play a warning role, traffic signs have certain particularity and easy to distinguish in color and shape, and can be classified to a certain extent according to the following figure 2. Therefore, two methods are usually used to detect and locate traffic signs: image binarization processing based on color segmentation and ROI extraction based on shape detection.

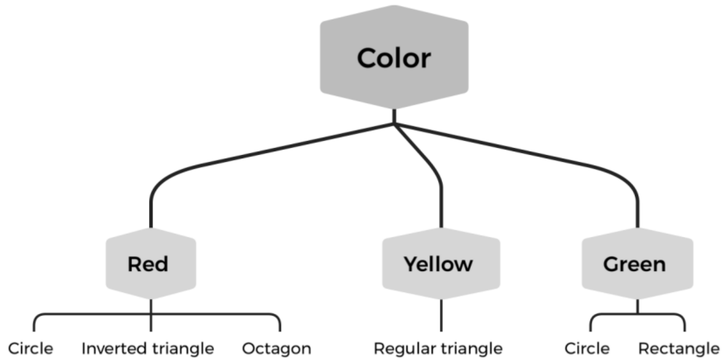

Fig. 2 Traffic signs classified by color and shape (Photo/Picture credit: Original).

2.3. Traffic sign Classification and Recognition (Feature extraction)

After the image segmentation and positioning is completed, only the key information of traffic signs is left in the image. At this time, it is necessary to further extract and compare these image features to identify the specific information [8]. The key features of the image are the key factors to identify the specific information, and the quality of the features directly determines the accuracy of recognition. Generally speaking, these key features need to have several elements such as distinguibility, parsimony and anti-interference to ensure the detection speed and efficiency, while ensuring as little as possible to be affected by noise, natural light and image distortion.

3. Traffic Sign Classification and Recognition Based on Deep Learning

Deep learning has significantly advanced image recognition, speech recognition, and natural language processing domains. Notably, it was first applied to image recognition in 1898 by LeCun and colleagues, who successfully utilized a deep neural network with a convolutional structure known as a Convolutional Neural Network (CNN) [9].

The LeNet-5 network stands as a quintessential model of CNNs. This network initially employs local connections and shared weights to extract features from input grayscale images, significantly reducing the model's training complexity and markedly enhancing its recognition and classification accuracy. Subsequently, the AlexNet and VGGNet networks were introduced, characterized by additional convolutional and connection layers, thereby refining the model's classification and recognition capabilities. Currently, CNN-based detection models are predominantly categorized into single-stage and two-stage models, which encounter challenges with slow detection speeds due to their process complexity. In response, GirshickR et al. introduced the R-CNN algorithm based on candidate region extraction, followed by RenS et al. who developed the enhanced FasterR-CNN algorithm, improving the candidate box generation approach. The introduction of an end-to-end convolutional neural network led to the proposal of the YOLO series network [10]. Subsequently, more streamlined object detection network models such as SSDLite and YOLO-LITE emerged, along with an improved SSD network-based traffic sign detection method. Yin et al. employed the YOLOv5 network, an evolution of YOLOv3, which demonstrates superior performance in traffic sign recognition. The continuous optimization of related algorithms significantly bolsters algorithmic performance. Nonetheless, comprehensive enhancement of recognition performance remains a challenge, necessitating further advancements.

As shown in Figure 3. Deep learning-based detection algorithms generally fall into two categories: one-stage and two-stage. The one-stage method derives output directly from input, while the two-stage approach first selects a subset of features, i.e., pre-selected boxes, during the input-output process. A comparative analysis of these one-stage and two-stage methods is as follows:

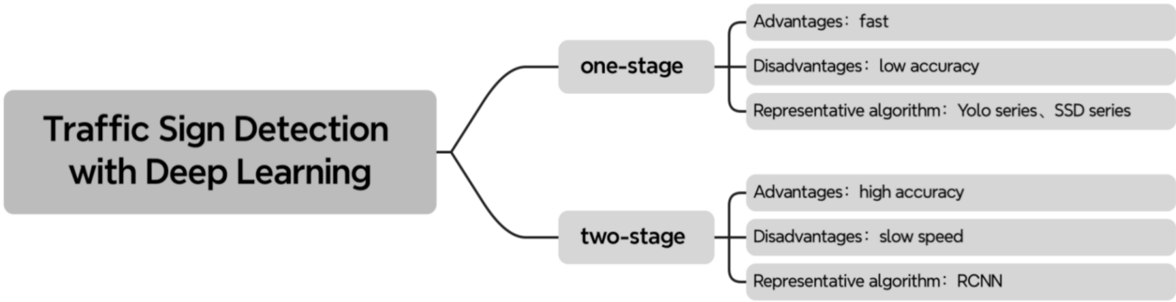

Fig. 3 Comparison of the two methods (Photo/Picture credit: Original).

The biggest difference between convolutional neural network deep learning algorithm and traditional machine recognition algorithm is that convolutional neural network can extract features while training, which can effectively improve the efficiency of detection and training. This gives convolutional neural networks a huge advantage in terms of detection performance and accuracy, especially for traffic sign recognition under complex conditions such as enhanced illumination, weather conditions, object occlusion, and shooting angles.

In recent years, with the unremitting efforts and innovation of researchers in the field of image recognition, more and more deep learning models and their variants applied to image recognition have been published to the public, and deep learning has been widely used in the field of intelligent transportation.

3.1. Based on the Improved LeNet-5

The LeNet model was proposed by Yann Lecun, the father of convolutional neural network. It is a 7-layer neural network. The network model is simple. The algorithm takes a long time and the recognition rate is low in complex traffic sign recognition. Zhang Meng et al improved and optimized the original LeNet-5 network by constructing sub-convolutional networks. By changing the size of convolution kernels, increasing the number of convolution kernels, adding nonlinear activation function ReLU and Dropout layer to improve feature extraction, reduce overfitting, and to make the network more stable. The accuracy of the improved network for traffic sign recognition reaches 93.558%. This recognition method of multi-convolution kernel and multi-scale features can achieve high accuracy in a short time, and has good robustness and generalization. Li Da et aleffectively preprocessed and augmented the data set, added convolutional layer, pooling layer and fully connected layer to the model to increase the depth of the model, and used ReLU activation function instead of Sigmoid activation function to reduce the computational complexity of the algorithm. The average accuracy of the improved LeNet-5 network in the validation data set reaches 99.42%, which is 4.93% higher than that of the traditional model. Wang Guiping et al introduced the Inception convolution module group on the basis of the traditional network to extract the rich features of the target. At the same time, he not only increased the depth of the network and introduced the BN layer to normalize the input batch samples, but also changed to Relu activation function with better performance.

3.2. Based on the improved residual network

ResNet was proposed by Dr. He from Microsoft Research. Compared with the VGG-Net network, the residual network can not only obtain deeper network layers without affecting the gradient stability, but also better control the number of model parameters. Building a target classification model based on CNN convolutional neural network will cause a series of problems such as model gradient explosion and disappearance with the deepening of the network. The emergence of ResNet can solve this problem well. How to ensure the accuracy of traffic sign recognition while still well controlling the iteration time and calculation amount of model training is a technical difficulty for researchers to overcome. Huang Shangan constructed an "eight-block ResNet model" according to the characteristics of traffic signs itself, and combined the log-likelihood cost function. It can solve the problem of gradient explosion and disappearance, not only control the model training iteration speed, but also improve the recognition accuracy. The improved model is tested on the public data set GTSRB of traffic sign recognition, and the test accuracy reaches 99.74%.In view of the continuous loss of high-level information of the image in the downsampling process of the traditional residual network model, which leads to insufficient feature extraction, Liang Zhengyou et al improved the ResNet network structure and proposed a traffic sign recognition method based on multi-scale features and attention mechanism, which fused the multi-scale features of each level of the model. Through the attention mechanism to strengthen the characteristics of different channels, the accuracy of the improved model on the GTSRB and BelgiumTS traffic sign datasets reaches 99.31% and 98.96% respectively, which is better than the state-of-the-art traffic sign recognition algorithms. The recognition accuracy of the improved ResNet model is significantly improved, which can better meet the high accuracy requirements of traffic sign recognition tasks. Although the proposed method effectively improves the accuracy of automatic traffic sign recognition, it increases the complexity and calculation of the model. Therefore, how to reduce the complexity of the model, reduce the time and computing resources required by the algorithm, and better meet the real-time applications and requirements under the premise of ensuring the recognition accuracy will be their next research direction. As shown in Figure 4.

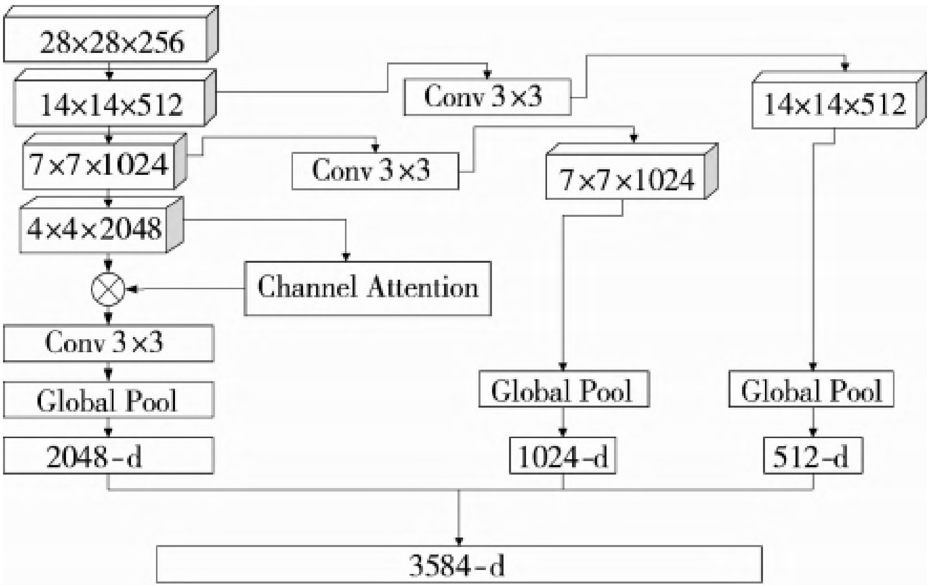

Fig. 4 Model structure (Photo/Picture credit: Original).

3.3. Based on YOLO-v5

YOLO algorithm has attracted much attention in the field of traffic sign recognition because of its good recognition performance. Among them, YOLO-V5 has the best recognition speed and recognition accuracy in the YOLO series. It is more suitable for practical engineering applications. Aiming at the problems of low accuracy and large volume in the traditional model, Zhang Yu et al proposed a lightweight algorithm for traffic sign detection based on the YOLO-v5s basic model. They used the attention mechanism SENet to improve the C3 module, integrated the CBAM module into the model, and cut half of the convolution kernel to realize the lightweight model. TensorRT technology is used to deploy the model on Jetson nano. The accuracy of the improved model is improved to 92.9%, and the FPS reaches 24.1f/s, which meets the requirements of real-time traffic sign detection. Zheng Hongbing et al proposed to introduce the GAM into the YOLO-v5 model. And replaced the GIoU loss function used in the YOLOv5 algorithm with the CIoU loss function with more regression characteristics to optimize the model. The improved algorithm is trained on the Tsinghua-Tencent 100K dataset. The mean average accuracy of traffic sign recognition reaches 93.00%, which is 5.72% higher than that of the original algorithm, and has better recognition performance.

4. Challenges

Many traffic accidents are attributed to obstacles in traffic sign recognition, leading to incorrect assessments of road conditions. The primary challenges faced by traffic sign detection can be broadly categorized into three aspects:

Traffic signs are constantly exposed to the elements throughout the year. They must withstand various complex weather conditions, which can result in color fading, and are also subject to obstructions and tilting.

Traffic signs exhibit a diverse range of colors, and a single image may capture multiple traffic signs, adding to the complexity of recognition.

Due to high speeds and other factors, the captured traffic signs can appear deformed, blurred, or distorted, complicating real-time detection and recognition.

Given the significance of traffic sign detection in reducing traffic accidents, it is a vital concern for both the country and society. Despite the formidable challenges, developing a traffic sign detection system with high real-time performance and accuracy remains a key focus for many researchers.

5. Conclusion

Traffic sign detection and recognition technology is an essential component of intelligent transportation systems and presents extensive prospects for development. An increasing number of scholars are focusing their research on this area. Currently, many traffic sign detection algorithms, particularly those based on deep learning, have shown impressive results in various datasets. As per the ongoing developmental trends, future traffic sign detection is expected to integrate multiple methodologies, leveraging the strengths of each to enhance overall performance. Advancements in traffic sign detection and recognition are pivotal in driving the progress of autonomous vehicle technology, thereby advancing human scientific and technological achievements for the greater good.

Authors Contribution

All the authors contributed equally.

References

[1] Zhang, K., Hou, J., Liu, M., & Liu, J. (2021). Traffic sign recognition based on improved convolutional networks. International Journal of Wireless and Mobile Computing, 21(3), 274-284.

[2] Zhang, R., Zheng, K., Shi, P., Mei, Y., Li, H., & Qiu, T. (2023). Traffic sign detection based on the improved YOLOv5. Applied Sciences, 13(17), 9748.

[3] Zhang Yu, Luo Suyun. Research on Traffic sign Detection based on Lightweight YOLOv5s [J]. Agricultural Equipment and Vehicle Engineering,2023,61(10):130-135. Huang Shangan. Traffic sign recognition based on improved ResNet residual network model [A]. Science and Technology Innovation,2021.

[4] Zheng-you, L. I. A. N. G., G. E. N. G., & Yu, S. U. N. (2022). Traffic sign recognition algorithm based on improved residual network. Computer and Modernization, (04), 52.

[5] Wang, G. P., Sheng, G. F., Huang, H., Wang, H. F., & Wang, P. (2018). Traffic sign recognition method based on improved LeNet-5 network. Science Technology and Engineering, 18(34), 83-89.

[6] Zhu, X., Huang, Y., Wang, X., & Wang, R. (2023). Emotion recognition based on brain-like multimodal hierarchical perception. Multimedia Tools and Applications, 1-19.

[7] Lim, X. R., Lee, C. P., Lim, K. M., Ong, T. S., Alqahtani, A., & Ali, M. (2023). Recent advances in traffic sign recognition: approaches and datasets. Sensors, 23(10), 4674.

[8] Chourasia, K., & Chourasia, J. N. (2011). A review and comparative analysis of recent advancements in traffic sign detection and recognition techniques. SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology, 2(01), 24-34.

[9] Hussain, S., Abualkibash, M., & Tout, S. (2018, May). A survey of traffic sign recognition systems based on convolutional neural networks. In 2018 IEEE International Conference on Electro/Information Technology (EIT) (pp. 0570-0573). IEEE.

[10] Gunasekara, S., Gunarathna, D., Dissanayake, M. B., Aramith, S., & Muhammad, W. (2022). Deep Learning based autonomous real-time traffic sign recognition system for advanced driver assistance. International Journal of Image, Graphics and Signal Processing, 14(6), 70-83.