# Traffic Sign Detection and Recognition for Intelligent Transportation Systems: A Survey

Shuanghu Luo[1,2], Ling Yu[3], Zhongqin Bi[3], Yongbin Li[3]

[1] School of Computer Engineering and Science, Shanghai University, China
[2] Shanghai Municipal Educational Examinations Authority, China
[3] College of Computer Science and Technology, Shanghai University of Electric Power, China
luosh@shmeea.edu.cn, yuri705@126.com, zqbi@shiep.edu.cn, lybin40000@163.com

## Abstract

With the development of intelligent transportation systems (ITSs), traffic sign detection and recognition methods play a vital role in unmanned vehicle driving (UVD) and traffic flow forecasting. However, due to the real-time and reliability features in traffic systems, each traffic sign should be handled in a specified interval to ensure the accuracy of the detection and recognition results. To this point, images captured by the vehicle's image sensors recognize traffic signs on the road to help drivers correctly operate their vehicles and even drive themselves. In this research area, the first step is to locate the area of traffic signs as the detection range, and the second step is to recognize the types of traffic signs detected. This paper presents a survey of the research results of scholars over the years, analyses the research results and presents the findings in summarized categories. Then, the prospects of future research regarding intelligent system implementations for traffic sign detection and recognition are elaborated. We hope this summary will lead to a boom in more original research on our smart city.

**Keywords:** Traffic sign, Object detection, Image recognition, Intelligent transportation system

## 1 Introduction

Vehicles play an important role in our society, providing a convenience for our trips, improving our living standard and changing our perspective. However, with the development of the automobile industry, the increasing number of vehicles has led to the frequent occurrence of traffic accidents, causing heavy casualties and tremendous property losses. To overcome this problem, intelligent transportation systems (ITSs) have emerged as the times have required. ITSs can effectively use the latest information and communication technology to integrate cars, people and roads into a unified system. In addition to supporting safe driving, preventing accidents and effectively alleviating congestion, ITSs are also helpful to realize automatic driving, improve material transport efficiency, reduce exhaust emissions and improve the environment along the road.

One of the main functions of ITSs is to work as an auxiliary pilot or replace the driver; this function is referred to as UVD. The UVD system can sense the road environment through the vehicle-mounted sensing system, automatically plan the driving route and control the vehicle to reach the predetermined destination. It has such functions as path planning, obstacle avoidance, navigation and traffic signal monitoring. A traffic sign detection and recognition module is a part of a UVD system [1-6]. As public infrastructure, traffic signs are designed to provide critical road information to drivers. If the driver or driving system does not detect or recognize traffic signs in time, then accidents may occur. Thus, the challenging task of traffic sign detection and recognition systems should be correct and reliable within a limited time.

In this area, there are many difficult problems that still need to be solved and improved. To the best of our knowledge, the following catalogues are listed as key points.

- Changes in lighting conditions. The intensity of the light is related not only to the season and the time of day but also to the weather. The illumination conditions of day and night, sunny days and rainy days may affect the colour appearance of traffic signs in the image. In addition, the pattern of traffic signs in the image may also be affected by the shadows of the surrounding objects.
- Traffic signs may experience deformation. The vibration of the moving vehicle may cause a distortion of the image of the traffic sign taken by the on-board camera. The concrete manifestations include deformation, fuzziness, inaccurate focus, etc.
- Traffic signs may fade in colour due to prolonged exposure to sunlight and rain. Over time, the paint on the sign may fade or even peel off. Additionally,

*Corresponding Author: Ling Yu; E-mail: yuri705@126.com
DOI: 10.3966/160792642020112106018

the camera also affects the image quality.

- The shape of a traffic sign may cause problems. Strong wind or heavy rain may cause traffic signs to tilt. For example, traffic signs will become brighter and lighter than normal in foggy weather, and in dusty weather, traffic signs will become yellower than normal.
- The scale of traffic signs. Along with mobile vehicles, the size of the traffic signs will change from small to large. The proportion of traffic signs in the original map is related to the difficulty of traffic sign recognition and affects the accuracy of traffic sign recognition.
- Obstacles blocking traffic signs. For example, trees, buildings, vehicles and pedestrians may block part of the area of signs, affecting the detection and identification of traffic signs.

The remainder of this paper is organized as follows. Studies regarding the traffic sign detection algorithm are reviewed in Section 2, and the traffic sign recognition algorithm is examined in Section 3. In Section 4, technical difficulties and future prospects are introduced. Our paper is concluded in Section 5.

## 2 Traffic Sign Detection Algorithm

The traffic sign detection stage includes recognizing the region in the image or video that contains traffic signs. Because the basic meaning of a traffic sign is defined by three attributes, namely, the colour, shape and pattern, we can use the colour and shape information to detect a traffic sign. As shown in Table 1, we list several exciting traffic sign detection algorithms.

**Table 1.** Comparison of traffic sign detection approaches

| Name of Technique | Cat | Theory | Advantage | Possible Issue(s) | Ref | Dev. Year | Performance | Dataset |
|---|---|---|---|---|---|---|---|---|
| Colour Threshold Segmentation | C* | Compare the colour properties to a set of value ranges and decide which category a pixel belongs to. | Simple. | Prohibition signs are particularly difficult to detect due to segmentation problems with achromatic colours. | [8]<br>[10] | 1997<br>2011 | No detailed detection test results given in the reference. | Self-created dataset. |
| Dynamic Pixel Aggregation | C* | Colour segmentation is performed by introducing a dynamic threshold in the pixel aggregation process in the HSV colour space. | Reduces hue instability. | Performance decreases when images are characterized by sign pixels with coordinates that fall in unstable HSV areas. | [11] | 2001 | 86.3%~95.7% segmentation hit rate. | Self-created dataset tested using 620 outdoor images. |
| Colour Indexing | C* | By comparing the colour histograms [7] of two images, the colour object in an image is assessed. | Fast and straight-forward. | The complexity of the traffic scene will greatly influence the computation. | [12]<br>[13]<br>[14]<br>[15] | 1990<br>1991<br>1995<br>1994 | | |
| Transformation Based on the CIECAM97 Model | C* | Converts the RGB space of the images into the standard XYA space of the CIE and then uses the CIECAM97 model to obtain the LCH to segment objects. | Fast and efficient. | Separates objects that are of similar colour as traffic signs. | [16]<br>[17] | 2002<br>2002 | [16] has a detection rate of 90% on sunny days. [17] has a detection rate of 90%. | British traffic signs scanned from the book of the Highway code; Russian traffic signs obtained from the web site; Self-created dataset. |
| HSI Transformation | C* | The RGB space of the image can be converted into the HSI colour space. | Not susceptible to light. | Non-colour traffic signs may be undetectable. | [22] | 2008 | ~96% detection rate. | Self-created dataset. |
| Region Growing Segmentation | C* | Starts with a seed pixel and then expands to a group pixels of similar colour affinity together. This can be done in the HSV space. | Simple and straightforward. | Region growing often leads to undergrowth or overgrowth due to the setting of non-optimal parameters. | [23] | 1993 | >84% detection rate. | Self-created dataset. |

**Table 1.** Comparison of traffic sign detection approaches (continue)

| Name of Technique | Cat | Theory | Advantage | Possible Issue(s) | Ref | Dev. Year | Performance | Dataset |
|---|---|---|---|---|---|---|---|---|
| Hierarchical Spatial Feature Matching | S* | Search for geometric objects in an image that resemble traffic signs. Suitable for traffic scene images with grey-level inputs. | Independent of traffic sign colour. | Results are inferior to those of colour-based algorithms. | [25] | 2002 | No detailed detection test result shown in the given reference. | Self-created dataset; Dataset of 558 images was acquired with digital cameras under general illumination conditions. |
| Hough Transform | S* | Separate the features of a particular shape from the image. | Can tolerate gaps in the feature boundary description and is not affected by image noise. | Due to the complexity of the calculation and the need for memory, there is a lack of detection speed. | [26] | 2006 | 92.5% average detection rate. | The dataset used in experiments was recorded with a VGA camera under diverse lighting conditions. |
| Distance Transform Matching | S* | This method uses a template hierarchy to capture objects of various shapes. | Can detect objects of any shape. | Matching still depends on a reasonable contour-segmentation approach. The detection rate drops significantly under severe weather conditions. | [27] | 1999 | ~95% detection rate. | A self-created dataset of 1000 traffic sign images taken during both day- (sunny, rainy) and night-time conditions. |
| Colour-shape Combination | B* | Image colour space colour segmentation and shape analysis to detect traffic signs. | Comprehensive. | The detection speed may be inadequate. | [28]<br>[29]<br>[30]<br>[31] | 2012<br>2012<br>2011<br>2008 | [28] has a detection rate of 93.77%. No detailed detection test results were given in [29]. [30] reported a detection rate of 97.36%. [31] reported a detection rate of ~93-98%. | [28-29] and [30] used self-created datasets, and [31] used the BelgiumTSC dataset. |
| Deep Learning | D* | Uses a CNN for feature extraction. | High speed and high performance. | High cost. | [38]<br>[39] | 2020<br>2019 | [38] had a detection rate of 98%. [39] had a detection rate of 97.8%. | [38] used the GTSDB, CCTSDB and Lisa datasets. [39] used the GTSDB dataset. |

*Note.* C*: Colour-based; S*: Shape-based; B*: Colour-shape combination; D*: Deep learning.

## 2.1 Traditional Traffic Sign Detection Algorithm

### 2.1.1 Detection Algorithm Based on Colour

The most important and obvious visual feature of traffic signs is colour. Most of the traditional traffic sign detection algorithms use a colour segmentation algorithm to extract traffic signs from the background for recognition. Here, are some common colour detection algorithms:

Colour Thresholding Segmentation: This is one of the first techniques used to segment digital images [7]. The threshold can be done by classifying the image pixels into traffic signs or backgrounds. When matching a pixel, if the pixel reaches or exceeds the threshold, then it is considered a traffic sign pixel. If the threshold is not reached, then it is considered a background pixel. Refs. [8] and [9] use this algorithm, and the colour threshold value during the test stage are described and analysed. However, this method cannot deal with non-colour traffic signs. To solve this problem, H. Gomez-Moreno et al. [10] proposed an algorithm for the detection of non-colour traffic signs.

Dynamic pixel aggregation: This is an algorithm based on colour segmentation. Vitabile et al. [11] implemented this algorithm with a segmentation rate of 86.3% ~ 94.6%. The colour segmentation process of this method introduces a dynamic threshold in the pixel aggregation process of the hue-saturation-value (HSV) colour space and adjusts the threshold according to the change in the image brightness to reduce the tonal instability of real scenes.

Colour Indexing: By comparing the colour histogram of two images, the colour object in the picture is compared. Colour histograms are used to index images and store them in a database. This algorithm is simple, fast and effective. However, the problem is that the complexity of the traffic scene will greatly increase the computation. This algorithm is described in detail in [12-15].

Transformation based on the CIE 1997 Colour Appearance Model (CIECAM97) Model: Refs. [16] and [17] used this algorithm and achieved a detection rate greater than 90%. First, the algorithm converts the red-green-blue (RGB) space of the image into the standard XYZ space of the Commission Internationale DE l 'eclairage (CIE) and then uses the CIECAM97 model to obtain the lightness, chrome, and hue (LCH) to segment the traffic signs. The disadvantage of this algorithm is that it will separate objects that are similar in colour from the traffic signs.

HSI Transformation: The RGB space of the image can be converted into the HSI colour space. The HSI colour model encodes the colour information by separating an overall brightness value from the two values encoding hue and saturation so that it is not affected by the changes in light. Therefore, this model is favoured by many researchers. Henry et al. [18] implemented this algorithm and achieved a detection rate of 95%. In addition, the algorithm is also used in [19-22].

Region Growing Segmentation: Region growing is the process of aggregating pixels or subareas into larger areas according to predefined criteria. This process starts with a pixel, merges adjacent pixels that satisfy the smoothing constraint, and outputs them in the form of a cluster of points. Each cluster point set is regarded as belonging to the same plane. The disadvantage of this algorithm is that region growing often leads to undergrowth or overgrowth due to the setting of non-optimal parameters. Its actual implementation can be referred to in Refs. [23] and [24].

One of the main problems of traffic sign detection algorithms based on colour is that external light may affect the image colour obtained by vehicle cameras or imaging sensors. Most colour-based technologies encounter problems because the external light affects both the image intensity and colour.

### 2.1.2 Detection Algorithm Based on Shape

Another apparent representation of a traffic sign is shape; therefore, the shape of traffic sign may also affect the detection algorithm. Traffic signs generally have triangular, rectangular, octagonal and circular shapes. This subsection discusses some common shape detection algorithms.

Hierarchy Spatial Feature Matching (HSFM): This detection algorithm is suitable for traffic scene images with grey-level input. A list of regions is generated by searching for geometric images in the image that resemble traffic signs. The contents of the list are then passed to the next step for recognition. Ref. [25] implements this algorithm.

Hough Transform: The Hough transform algorithm is generally used to separate the image characteristics of a specific shape. Its advantage is that this algorithm can tolerate a gap in the description of feature boundaries and is not affected by image noise. In [26], the Hough Circle Transform is used to detect circular symbols, and the Hough Line transform is used to detect triangular symbols. Using the Hough transform, the detection rates of the speed limit and alarm signals are 97.2% and 94.3%, respectively. The disadvantage of this algorithm is that the computation is complex, and the memory requirement is large, which makes it lack detection speed, so it is not an ideal choice for real-time detection algorithms.

Similarity Detection: This method is achieved by calculating the similarity coefficient between a segmented region and a set of binary image samples representing the shape of each traffic sign. The algorithm assumes that the dimensionality of the sampled image is the same as that of the segmented image. In [11], this algorithm is used and achieved a maximum detection rate of 95.7%. However, the detection rate of the triangle symbol is the lowest, which is 86.3%.

Distance Transform Matching: This algorithm uses the template hierarchy to capture objects of various shapes, and the stochastic optimization technique is introduced to generate the hierarchy of a given shape distribution online [27]. This algorithm can detect objects of any shape, which is an advantage over other techniques in dealing with nonrigid objects. As a function of template transformation parameters, the similarity measure obtained by this algorithm is smoother. However, its limitation is that matching still relies on a reasonable contour segmentation, and the detection rate drops seriously from 95% to 80% under severe weather conditions.

Shape-based detection algorithms do not require colour information. However, these algorithms may have more problems than colour-based detection algorithms. Traffic signs in a messy scene, with an imperfect shape, object occlusion, differences in the scale, size, or angle of orientation, etc., all make this algorithm very challenging.

#### 2.1.3 Both Colour-Based and Shape-Based Detection Algorithms

When an image contains traffic signs with the same colour or shape, a detection algorithm that is based only on colour or shape information may cause noise. Therefore, many researchers have been working on detection algorithms that combine colour information and shape information. This detection algorithm is generally composed of two stages: image colour space colour segmentation and shape analysis.

In [28], the RGB ratio is used to segment the RGB input image and then the Douglas-Peucker (DP) algorithm is used for shape analysis to detect traffic signs. The DP algorithm is a contour approximation technique based on the number of object boundaries. Therefore, traffic signs can be detected even if they have some geometric distortions.

According to [29], colour segmentation is first carried out in HSV colour space, and then boundary boxes are inserted into all regions detected by colour segmentation. Finally, the traffic signs are detected by the average colour, size and number of pixels in the boundary box.

In [30], a real-time driver assistance system is presented that integrates single view detection with region-based 3D tracking of traffic signs. This algorithm has high detection accuracy and excellent detection efficiency.

In [31], an algorithm is proposed to extract traffic signs from dynamic or static complex scenes. The algorithm combines invariant features and a support vector machine. The support vector machine is divided into two stages, which include the determination of the shape of the sign rim and a pictogram of the sign. Moreover, the best performance achieved is 98% for sign rims and 93% for speed limit signs.

Compared with detection based on a single feature, this algorithm makes full use of the advantages and bypasses the disadvantages to improve the detection performance.

### 2.2 Traffic Sign Detection Algorithm Based on Deep Learning

In the traditional traffic sign detection algorithm, most of the algorithms consider the inherent colour and shape characteristics of traffic signs and only use the design to detect a specific category of the traffic sign detection algorithm. However, both colour information and shape information are easily affected by environmental conditions. Thus, in severe weather conditions and complex environments, the algorithm tends to fail.

In 2014, the region-based convolutional neural network (R-CNN) [32] successfully applied deep learning to target detection and obtained the optimal effect at that time. This success made researchers in the field of computer vision pay more attention to deep learning algorithms and promote the development of CNNs. At present, most advanced traffic sign detection algorithms are based on deep learning.

In [33], an algorithm is proposed that combines the support vector machine (SVM) and CNN methods. First, SVM was used to transform an RGB image into a grey-level image, and then the image was input into the next CNN layers for detection and recognition. Then, bootstrapping was used to improve the accuracy and avoid the problem of overfitting. The algorithm obtained an average detection rate of 98.68% on the GTSDB dataset. However, the algorithm is time-consuming with respect to the processing time and may not meet the requirements for real-time detection.

In [34], a traffic sign detection algorithm is proposed based on a deep CNN and used a regional proposal network (RPN) to detect traffic signs. The dataset used in the experiment of this algorithm consists of traffic signs in China. The detection rate of continuous image sequences reaches 99%, and the average detection time of each image is approximately 51.5 ms. This algorithm seems to be suitable for real-time detection but requires high hardware quality and may not be a good choice for unmanned driving systems.

In [35], a model is proposed based on Faster R-CNN [36]. The model is divided into two parts. First, the regions of interest are detected by a selective search, and then these regions are extracted, classified and modified by CNNs. However, the algorithm also treats traffic lights as traffic signs, so the accuracy of the algorithm is reduced, and the effect is not ideal.

In [37], the small object sensitive (SOS)-CNN model is proposed; this algorithm is used to detect traffic signs through an image pyramid. The algorithm first samples the input original image and divides it into a small image with a fixed size. These small images are then input into the SOS-CNN model as input images to build the image pyramid. The algorithm has good performance in small target detection. However, the algorithm adopts the sliding window strategy, which is very time-consuming and may not be able to satisfy the requirement of the real-time detection time.

In [38], the cascade R-CNN algorithm with multiscale concerns and unbalanced samples is proposed. The algorithm first obtains the multiscale features of the image pyramid by a cascaded R-CNN. Then, a multiscale attention algorithm is proposed, which is used for the dot product and soft-max processing of the input multiscale features to obtain the similarity measurement. Finally, the features are weighted and summed to refine the features and improve the accuracy of the object detection. The algorithm has good performance, but it is not a good choice for mobile and embedded devices because of its high hardware requirements.

In [39], the framework of the faster R-CNN and the structure of MobileNet [40] are used to design and implement a traffic sign detection algorithm. In the algorithm, colour and shape information is used to refine the location of small traffic signs. Finally, an efficient CNN with an asymmetric kernel is used to classify traffic signs. The algorithm can be trained end-to-end, can be applied in real time, and can detect traffic signs beyond the limits of colour and shape. The detection rate is better than the most advanced algorithm, and the detection speed also meets the real-time detection requirements. However, the localization refinement method of the boundary box of this algorithm is designed for specific traffic signs, which is not robust enough and needs to be improved.

In [41], the combination of several publicly popular target detection frameworks (faster R-CNN, R-FCN [42], single-shot detection (SSD) [43] and YOLO V2 [44]) with various feature extractors (ResNet-V1-50 [45], ResNet-V1-101, Inception-V2 [46], Inception-ResNet-V2 [47], MobileNet-V1 and Darknet-19) is analysed. The objective of the author is to explore the properties of these modified target detection models that are particularly applicable to the problem area of traffic sign detection by means of transfer learning. Experiments show that the combination of the faster R-CNN and Inception-ResNet-V2 achieves the best average accuracy (mAP), while the combination of the R-FCN and ResNet-101 achieves the best balance between the detection rate and detection time. YOLO V2 is a framework with the second highest detection speed and high detection rate. The combination of SSD and MobileNet is the fastest and lightest in terms of memory consumption, making it the best choice for mobile and embedded device deployment.

## 3 Traffic Sign Recognition Algorithm

Traffic sign recognition is the explanation of the meaning of traffic signs and is the goal of traffic sign detection. An unmanned driving system locates the area of traffic signs in the complex traffic scene and then recognizes the types of traffic signs by the algorithm. As shown in Table 2, we list the several exciting traffic sign recognition algorithms.

**Table 2.** Comparison of traffic sign recognition approaches

| Name of Technique | Cat | Theory | Advantage | Possible Issue(s) | Ref | Dev. Year | Performance | Dataset |
|---|---|---|---|---|---|---|---|---|
| Feature Matching or Template Matching | T* | All signs to be recognized are stored in a database. Each potential sign has a normalized size and is compared to every template of the same shape. | Fast and efficient. | Image distortion or blocking may pose a problem. | [28]<br>[29] | 2012<br>2012 | [28] has a recognition rate of 93.77%. No detailed recognition test results are given in [29]. | Both use a self-created dataset. |
| Edge Analysis | T* | Based on the input image edge information to carry out geometric analysis to recognize the type of traffic signs. | Will not be affected by changes in the shooting distance or shooting angle. | In large datasets, the convergence rate will be very slow, and many experiments must be carried out. | [51] | 2012 | [51] has a recognition rate of 93%. | Tests in real driving conditions. |
| Support Vector Machine | T* | Binary classification of data according to supervised learning. | High performance. | In the complex traffic scene, may not be able to achieve real-time recognition. | [54] | 2013 | [54] has a recognition rate of 94.49%. | The self-created dataset is acquired by a mono-camera mounted on a moving vehicle. |
| Artificial Neural Network | T* | Train neurons to recognize colour, shape or pictogram on traffic sign. | High performance and high speed. | High computational complexity and high cost. | [56]<br>[57] | 2016<br>2009 | [56] has a recognition rate of 100% in daylight and 94.7% in shadow. [57] has a recognition rate of 95.29%. | The self-created dataset is acquired by a mono-camera mounted on a moving vehicle. |

**Table 2.** Comparison of traffic sign recognition approaches (continue)

| Name of Technique | Cat | Theory | Advantage | Possible Issue(s) | Ref | Dev. Year | Performance | Dataset |
|---|---|---|---|---|---|---|---|---|
| Deep Learning | D* | Using CNN for feature extraction. | High speed and high performance. | High cost. | [58]<br>[59]<br>[60]<br>[61] | 2011<br>2012<br>2014<br>2018 | The recognition rate of [58] is 98.5%, and the number of parameters is 1.54 M. The recognition rate of [59] is 99.46%, and the number of parameters is 38.5 M. The recognition rate of [60] is 99.63%, and the number of parameters is 23.2 M. The recognition rate of [61] is 98.0%, and the number of parameters is 0.5 M. | Both use the GTSRB dataset. |

*Note.* T*: Traditional algorithm; D*: Deep learning algorithm.

## 3.1 Traditional Traffic Sign Recognition Algorithm

In the traditional traffic sign recognition, the feature matching algorithm, edge analysis algorithm and machine learning algorithm are mostly used.

### 3.1.1 Feature Matching or Template Matching Algorithm

Some researchers have used a point of interest detector based on the speeded up robust features (SURF) [48] algorithm to recognize traffic signs. First, SURF is used to detect the points of interest in the image, such as spots and corners. SURF is then used as a descriptor to form an eigenvector to represent the detected points of interest. Then, similarity matching is obtained by comparing the feature vectors of the target traffic signs with the feature vectors stored in the database.

Some algorithms based on scale- and rotation-invariant binary robust invariant scalable keypoints (BRISK) [49] are used to recognize traffic signs. The two main stages of BRISK are key point detection and binary bit string extraction. The first stage is the identification of points of interest, and the second stage involves pixel comparisons of the intensity to form descriptor vectors. The maximum number of key-value matches provided by the template image is considered the target class. BRISK algorithms have an advantage over SURF algorithms in terms of speed of detection and descriptor computation. However, for high-speed vehicles, the recognition time is still long and may not meet the requirements of real-time recognition.

### 3.1.2 Edge Analysis Algorithm

The edge analysis algorithm is based on input image edge information to carry out a geometric analysis to recognize the type of traffic signs. In 1962, Hu proposed the Hu-moment invariant [50]. Using the nonlinear combination of moments, the invariance of translation, rotation and scale can be obtained. These three features facilitate good image recognition performance due to the change in the shooting distance or shooting angle. However, this algorithm requires considerable computation time.

In [51], a traffic sign recognition algorithm is proposed based on the Hough transform depending on the histogram function and the information received from every candidate contour. However, this method lacks a common evaluation framework for traffic sign detection systems. In [52], the BP neural network is adopted to classify the image with the Hu-moment as the feature. However, when the algorithm is trained on a large-scale training set, the convergence speed is very slow, and many experiments are needed.

### 3.1.3 Machine Learning Algorithm

Machine learning-based algorithms are popular in recognition tasks, where they are trained to mine the differences between different objects and then classify them. The most commonly used algorithms are SVM-based and neural network-based algorithms.

SVM is a kind of generalized linear classifier that classifies a data binary according to supervised learning, and its decision boundary is the maximum distance hyperplane for solving the learning sample. In [53], the detected traffic signs are coded by extracting the histogram of oriented gradients (HOG) feature of the image to generate feature vectors. The algorithm takes the generated feature vector as the input of the SVM and has high recognition accuracy. In [54], Chinese traffic signs are recognized by using the uniform rotation-invariant local binary pattern (LBP) and SVM. The algorithm has a recognition accuracy of 94.49%. In [55], the use of global features and Zemike-moment features of binary images is discussed for training and analysis. The experimental results of four different kernel SVMs are compared and satisfactory results are obtained. The recognition accuracy of SVM is very good, but it only applies to the case of linear separability, and the interference degree is poor. The computation time is too long, and in a complex traffic scene, real-time recognition may not be achieved.

An artificial neural network (ANN) is proposed on the basis of the research results of modern neuroscience. This model tries to process information by simulating brain neural network processing and memory information. In [56], the author uses a self-associative neural network to recognize traffic signs in the image, and the recognition rates in sunlight and shadow environments are 100% and 94.7%, respectively. In [57], a radial basis function (RBF) neural network is used to recognize traffic signs. Before the traffic signs are input into the RBF neural network, principal component analysis (PCA) is used to extract the characteristics of traffic signs and reduce the dimensionality of the original image. The recognition rate of the algorithm is more than 95%. The ANN has good real-time performance and high recognition accuracy, so it is a good choice for traffic sign recognition. However, this algorithm demonstrates high computational complexity, which also means a corresponding increase in cost.

### 3.2 Traffic Sign Recognition Algorithm Based on Deep Learning

An ANN is a fully connected structure. This structure has considerable model parameters, which deepens the training difficulty. During training, gradient disappearance or gradient explosion may occur. ANNs are not widely used because they are more difficult to train when using larger datasets. To solve these problems, CNNs came into being and entered the era of deep learning to solve visual problems.

Over the past 10 years, researchers in the field of traffic sign recognition have proposed various high-performance network models. At the beginning, most of them studied the depth and width of CNNs to improve the accuracy of the model in the task. In recent years, many researchers have tried to reduce the number of parameters while maintaining the accuracy of the model.

Ciresan et al. [58] proposed a fast and fully parameterized algorithm based on a CNN using a graphics processing unit (GPU). The algorithm does not require elaborate preconnected feature extractors, which are learned in a supervised manner to render the system insensitive to contrast and light changes. The average recognition rate reached 98.5%, while the number of parameters was 1.54 M. One year later, Ciresan et al. [58] proposed an improved algorithm that combined various deep neural networks (DNNs) trained by different preprocessing data into multicolumn DNNs (MCDNN) to replace the CNN. Moreover, the recognition rate of the algorithm reached 99.46%. This algorithm was the champion of the identification benchmark of traffic sign detection and recognition contest held in Germany in 2011. However, the number of parameters of this algorithm is increased to 38.5 M, and the calculation amount and time increase accordingly.

In [60], a recognition algorithm is proposed to train a CNN through hinge loss stochastic gradient descent (HLSGD). The network structure of the algorithm is composed of three layers, and the number of parameters is 23.2 M; the algorithm recognition rate reached 99.63%. The computational complexity of this algorithm is lower than that of the MCDNN, and the recognition accuracy is higher than that of the MCDNN, which suggests an excellent traffic sign recognition algorithm.

In [61], a traffic sign recognition algorithm (MicronNet) is proposed that is very suitable for embedded scenes. The network architecture of the algorithm is a highly compact deep CNN. The number of parameters is only 0.51 M, which is very friendly for embedded devices. The recognition rate is 98.0% in terms of the full precision and 98.9% in half precision.

## 4 Technical Difficulties and Future Prospects

Some progress has been made in the research of traffic sign detection and recognition, but a gap still exists between research and practical applications. The existing difficulties in the current research are listed as follows:

- There are many kinds of traffic signs. China, for example, has 204 traffic signs. Among them, there are 116 kinds of signs in three categories directly related to road traffic safety and some derived signs (such as speed limit signs and other cases with numbers). Furthermore, this number is growing. The wide variety of traffic signs requires scalability of the classification algorithm.
- Most of the current studies emphasize theoretical aspects, while few are application-oriented. Many theories and methods have been put forward based on standard diagrams or some local conditions without considering the need of practical applications.
- Most of the experimental subjects are based on standard diagrams, and there are few studies on real scene diagrams. A practical traffic sign detection and recognition system should take the standard map as the reference sample and the real scene map as the recognition object.

- Regarding accuracy and real-time performance, traffic sign detection and recognition algorithms require high accuracy and high processing speed. If the accuracy and processing speed do not meet the requirements, not only will the expected role of auxiliary driving be unrealized but also traffic accidents will be easily caused.
- In terms of the practical application of economic costs, the following are considerations. In practical applications, it is necessary to overcome various difficulties, among which economic cost is also an important point. There is little research based on mobile devices, and many good models are expensive.

With the improvement in the structure and learning methods of deep learning models, the research on traffic sign detection and recognition has made great progress. Therefore, establishing implementations of deep learning methods to improve the time and efficiency of traffic sign detection and recognition and reducing the economic cost are areas in need of further research. Moreover, few research studies have been conducted on road videos. Therefore, research that considers real-world scenarios is also one of the future directions.

## 5 Conclusion

In this paper, the challenges of traffic sign detection and recognition are discussed, and the traditional algorithms and deep learning algorithms commonly used in traffic sign detection and recognition are briefly reviewed.

In traditional detection and recognition algorithms, the most common method is to locate and segment traffic signs in the image based on colour or shape and then classify them using the SVM model. Some studies have used algorithms based on ANNs, but the speed of these algorithms may not be suitable for real-time tasks. Furthermore, the datasets utilized in many of the related papers are not publicly available benchmark datasets, and some papers lack some evaluation features, such as the accuracy, true positive (TP) value, false positive (FP) value, recall or area under the curve (AUC).

In recent years, with the development of deep learning in many other fields, many researchers have begun to apply CNNs to traffic sign detection and recognition tasks. The traffic sign detection and recognition algorithm based on the CNN has the advantages of high precision and fast processing speed. However, due to the large amount of computation, the hardware requirements are also high. Therefore, using this approach in the ITSs does not seem easy.

Most of the algorithms introduced in this paper are tested on a static image dataset, and only a few are tested on a real road video record. While traffic sign detection and recognition techniques have improved over the years, more research and testing need to be done on different datasets and real-time videos. Thus, we think that intelligent computations based on the Internet of Things (IoT) and Internet of Vehicles (IoV) [62-63] with 5G technology should be promoted.

## References


[1] Z. Zhang, T. Tan, K. Huang, Y. Wang, Practical Camera Calibration from Moving Objects for Traffic Scene Surveillance, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 23, No. 3, pp. 518-533, March, 2013.

[2] Z. Zhang, K. Huang, Y. Wang, M. Li, View Independent Object Classification by Exploring Scene Consistency Information for Traffic Scene Surveillance, *Neurocomputing*, Vol. 99, pp. 250-260, January, 2013.

[3] C. Yao, X. Bai, W. Liu, L. J. Latecki, Human Detection Using Learned Part Alphabet and Pose Dictionary, *13th European Conference on Computer Vision (ECCV 2014)*, Zurich, Switzerland, 2014, pp. 251-266.

[4] C. Hu, X. Bai, L. Qi, X. Wang, G. Xue, L. Mei, Learning Discriminative Pattern for Real-Time Car Brand Recognition, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 16, No. 6, pp. 3170-3181, December, 2015.

[5] Y. Xia, W. Xu, L. Zhang, X. Shi, K. Mao, Integrating 3D Structure into Traffic Scene Understanding with RGB-D Data, *Neurocomputing*, Vol. 151, pp. 700-709, March, 2015.

[6] Q. Ling, J. Yan, F. Li, Y. Zhang, A Background Modeling and Foreground Segmentation Approach Based on the Feedback of Moving Objects in Traffic Surveillance Systems, *Neurocomputing*, Vol. 133, pp. 32-45, June, 2014.

[7] R. Narasimhan, J. P. Fornango, Some Further Experiments in the Parallel Processing of Pictures, *IEEE Transactions on Electronic Computers*, Vol. EC-13, No. 6, pp. 748-750, December, 1964.

[8] A. de la Escalera, L. Moreno, M. Salichs, J. M. Armingol, Road Traffic Sign Detection and Classification, *IEEE Transactions on Industrial Electronics,* Vol. 44, No. 6, pp. 848 -859, December, 1997.

[9] D. Ghica, S. Lu, X. Yuan, Recognition of Traffic Signs by Artificial Neural Network, *ICNN'95- International Conference on Neural Networks*, Perth, WA, Australia, 1995, pp. 1444-1449.

[10] H. Gomez-Moreno, S. Maldonado-Bascon, P. Gil-Jimenez, S. Lafuente-Arroyo, Goal Evaluation of Segmentation Algorithms for Traffic Sign Recognition, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 11, No. 4, pp. 917-930, December, 2010.

[11] S. Vitabile, G. Pollaccia, G. Pilato, F. Sorbello, Road Signs Recognition Using a Dynamic Pixel Aggregation Technique in the HSV Color Space, *11th International Conference on Image Analysis and Processing*, Palermo, Italy, 2001, pp. 572-577.

[12] M. J. Swain, D. H. Ballard, Indexing via Color Histograms, *Third International Conference on Computer Vision*, Osaka,

Japan, 1990, pp. 390-393.

[13] M. J. Swain, D. H. Ballard, Color Indexing, *International Journal of Computer Vision*, Vol. 7, no. 1, pp. 11-32, November, 1991.

[14] B. V. Funt, G. D. Finlayson, Color Constant Color Indexing, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 5, pp. 522-529, May 1995.

[15] G. Healey and D. Slater, Global Color Constancy: Recognition of Objects by Use of Illumination-invariant Properties of Color Distributions, *Journal of the Optical Society of America A*, Vol. 11, No. 11, pp. 3003-3010, November, 1994.

[16] X. Gao, N. Shevtsova, K. Hong, S. Batty, L. Podladchikova, A. Golovan, D. Shaposhnikov, V. Gusakova, Vision Models Based Identification of Traffic Signs, *First European Conference on Color in Graphics, Imaging, and Vision (CGIV 2002)*, Poitiers, France, 2002, pp. 47-51.

[17] D. Shaposhnikov, L. N, E. Golovan, N. Shevtsova, K. Hong, X. Gao, *Road Sign Recognition by Single Positioning of Space-Variant Sensor*, https://core.ac.uk/display/22601125, 2002.

[18] H. X. Liu, B. Ran, Vision-Based Stop Sign Detection and Recognition System for Intelligent Vehicles, *Transportation Research Record*, Vol. 1748, No. 1, pp. 161-166, January, 2001.

[19] H. Fleyeh, Shadow and Highlight Invariant Colour Segmentation Algorithm For Traffic Signs, *2006 IEEE Conference on Cybernetics and Intelligent Systems*, Bangkok, Thailand, 2006, pp. 1-7.

[20] A. de la Escalera, J. M. Armingol, M. Mata, Traffic Sign Recognition and Analysis for Intelligent Vehicles, *Image and Vision Computing*, Vol. 21, No. 3, pp. 247-258, March, 2003.

[21] B. Cyganek, Color Image Segmentation with Support Vector Machines: Applications to Road Signs Detection, *International Journal of Neural Systems*, Vol. 18, No. 4, pp. 339-345, August, 2008.

[22] C. G. Kiran, L. V. Prabhu, V. Abdu Rahiman, K. Rajeev, A. Sreekumar, Support Vector Machine Learning Based Traffic Sign Detection and Shape Classification Using Distance to Borders and Distance from Center Features, *TENCON 2008*, Hyderabad, India, 2008, pp. 1-6.

[23] O. Ennio, C. Paolo, P. Giulia, N. Gianluca, Automatic Road Sign Detection and Classification from Color Image Sequences, *7th International Conference on Image Analysis and Processing*, Monopoli, Italy, 1993, pp. 1-23.

[24] L. Priese, V. Rehrmann, On Hierarchical Color Segmentation and Applications, *IEEE Conference on Computer Vision and Pattern Recognition*, New York, NY, USA, 1993, pp. 633-634.

[25] P. Paclík, J. Novovicova, Road Sign Classification without Color Information, *6th Annual conference of the Advanced School of Imaging and Computing*, Lommel, Belgium, 2000, pp. 1-7.

[26] M. A. Garcia-Garrido, M. A. Sotelo, E. Martin-Gorostiza, Fast Traffic Sign Detection and Recognition under Changing Lighting Conditions, *2006 IEEE Intelligent Transportation Systems Conference*, Toronto, Ont., 2006, pp. 811-816.

[27] D. M. Gavrila, V. Philomin, Real-time Object Detection for "Smart" Vehicles, *The Seventh IEEE International Conference on Computer Vision*, Kerkyra, Greece, 1999, pp. 87-93.

[28] Z. Zheng, H. Zhang, B. Wang, Z. Gao, Robust Traffic Sign Recognition and Tracking for Advanced Driver Assistance Systems, *The 15th International IEEE Conference on Intelligent Transportation Systems*, Anchorage, AK, 2012, pp. 704-709.

[29] E. Oruklu, D. Pesty, J. Neveux, J. Guebey, Real-time Traffic Sign Detection and Recognition for In-car Driver Assistance Systems, *IEEE 55th International Midwest Symposium on Circuits and Systems (MWSCAS)*, Boise, ID, 2012, pp. 976-979.

[30] R. Timofte, V. A. Prisacariu, L. V. Gool, I. Reid, Combining Traffic Sign Detection with 3D Tracking Towards Better Driver Assistance, in: C. H. Chen (Ed.), *Emerging Topics in Computer Vision and Its Applications*, World Scientific Publishing Co. Pte. Ltd., 2011, pp. 425-446.

[31] H. Fleyeh, M. Dougherty, Traffic Sign Classification Using Invariant Features and Support Vector Machines, *2008 IEEE Intelligent Vehicles Symposium*, Eindhoven, Netherlands, 2008, pp. 530-535.

[32] R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation, *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, 2014, pp. 580-587.

[33] Y. Wu, Y. Liu, J. Li, H. Liu, X. Hu, Traffic Sign Detection Based on Convolutional Neural Networks, *The 2013 International Joint Conference on Neural Networks* (*IJCNN*), Dallas, TX, 2013, pp. 1-7.

[34] C. Xiong, C. Wang, W. Ma, Y. Shan, A Traffic Sign Detection Algorithm Based on Deep Convolutional Neural Network, *2016 IEEE International Conference on Signal and Image Processing (ICSIP)*, Beijing, China, 2016, pp. 676-679.

[35] Z. Zuo, K. Yu, Q. Zhou, X. Wang, T. Li, Traffic Signs Detection Based on Faster R-CNN, *2017 IEEE 37th International Conference on Distributed Computing Systems Workshops (ICDCSW)*, Atlanta, GA, 2017, pp. 286-288.

[36] S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, No. 6, pp. 1137-1149, June, 2017.

[37] Z. Meng, X. Fan, X. Chen, M. Chen, Y. Tong, *Detecting Small Signs from Large Images,* https://arxiv.org/abs/1706.08574, 2017.

[38] J. Zhang, Z. Xie, J. Sun, X. Zou, J. Wang, A Cascaded R-CNN With Multiscale Attention and Imbalanced Samples for Traffic Sign Detection, *IEEE Access*, Vol. 8, pp. 29742-29754, February, 2020.

[39] J. Li, Z. Wang, Real-Time Traffic Sign Recognition Based on Efficient CNNs in the Wild, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 20, No. 3, pp. 975-984, March, 2019.

[40] A. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T.

Weyand, M. Andreetto, H. Adam, *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*, https://arxiv.org/abs/1704.04861, 2017.

[41] Á. Arcos-García, J. Alvarez-Garcia, L. Soria-Morillo, Evaluation of Deep Neural Networks for Traffic Sign Detection Systems, *Neurocomputing*, Vol. 316, pp. 332-344, November, 2018.

[42] J. Dai, Y. Li, K. He, J. Sun, R-FCN: Object Detection via Region-based Fully Convolutional Networks, *The 30th International Conference on Neural Information Processing Systems*, Barcelona, Spain, 2016, pp. 379-387.

[43] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, A. Berg, SSD: Single Shot MultiBox Detector, *15th European Conference on Computer Vision (ECCV 2016)*, Amsterdam, Netherlands, 2016, pp. 21-37.

[44] J. Redmon, A. Farhadi, YOLO9000: Better, Faster, Stronger, *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 2017, pp. 6517-6525.

[45] K. He, X. Zhang, S. Ren, J. Sun, Deep Residual Learning for Image Recognition, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, 2016, pp. 770-778.

[46] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the Inception Architecture for Computer Vision, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, 2016, pp. 2818-2826.

[47] C. Szegedy, S. Loffe, V. Vanhoucke, A. A. Alemi, Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning, *The Thirty-First AAAI Conference on Artificial Intelligence (AAAI)*, San Francisco, California, USA, 2017, pp. 4278-4284.

[48] H. Bay, T. Tuytelaars, L. V. Gool, SURF: Speeded up Robust Features, *European Conference on Computer Vision (ECCV 2006)*, Graz, Austria, 2006, pp. 404-417.

[49] S. Leutenegger, M. Chli, R. Y. Siegwart, BRISK: Binary Robust Invariant Scalable Keypoints, *2011 International Conference on Computer Vision*, Barcelona, Spain, 2011, pp. 2548-2555.

[50] M.-K. Hu, Visual Pattern Recognition by Moment Invariants, *IRE Transactions on Information Theory*, Vol. 8, No. 2, pp. 179-187, February, 1962.

[51] M. Garcia-Garrido, M. Ocaña, D. Fernández-Llorca, E. Arroyo, J. Pozuelo, M. Gavilan, Complete Vision-Based Traffic Sign Recognition Supported by an I2V Communication System, *Sensors (Basel)*, Vol. 12, No. 2, pp. 1148-1169, January, 2012.

[52] N. Barnes, A. Zelinsky, L. S. Fletcher, Real-Time Speed Sign Detection Using the Radial Symmetry Detector, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 9, No. 2, pp. 322-332, June, 2008.

[53] N. B. Romdhane, H. Mliki, M. Hammami, An Improved Traffic Signs Recognition and Tracking Method for Driver Assistance System, *2016 IEEE/ACIS 15th International Conference on Computer and Information Science* (*ICIS*), Okayama, Japan, 2016, pp. 1-6.

[54] H. Zhang, B. Wang, Z. Zheng, Y. Dai, A Novel Detection and Recognition System for Chinese Traffic Signs, *The 32nd Chinese Control Conference*, Xi'an, China, 2013, pp. 8102-8107.

[55] F. Boi, L. Gagliardini, A Support Vector Machines Network for Traffic Sign Recognition, *The 2011 International Joint Conference on Neural Networks*, San Jose, CA, 2011, pp. 2210-2216.

[56] S. H. S. G, C. M. Patil, An Approach towards Efficient Detection and Recognition of Traffic Signs in Videos Using Neural Networks, *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, Chennai, India, 2016, pp. 456-459.

[57] K. H. Lim, L.-M. Ang, K. P. Seng, New Hybrid Technique for Traffic Sign Recognition, *2008 International Symposium on Intelligent Signal Processing and Communications Systems*, Bangkok, Thailand, 2009, pp. 1-4.

[58] D. Ciresan, U. Meier, J. Masci, J. Schmidhuber, A Committee of Neural Networks for Traffic Sign Classification, *The International Joint Conference on Neural Networks*, San Jose, CA, 2011, pp. 1918-1921.

[59] D. Cireşan, U. Meier, J. Masci, J. Schmidhuber, Multi-Column Deep Neural Network for Traffic Sign Classification, *Neural networks*, Vol. 32, pp. 333-338, August, 2012.

[60] J. Jin, K. Fu, C. Zhang, Traffic Sign Recognition with Hinge Loss Trained Convolutional Neural Networks, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 15, No. 5, pp. 1991-2000, October, 2014.

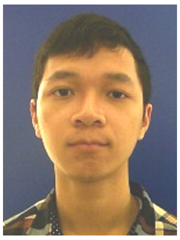
[61] A. Wong, M. J. Shafiee, M. St. Jules, MicronNet: A Highly Compact Deep Convolutional Neural Network Architecture for Real-Time Embedded Traffic Sign Classification, *IEEE Access*, Vol. 6, pp. 59803-59810, October, 2018.

[62] H. Gao, C. Liu, Y. Li, X. Yang, V2VR: Reliable Hybrid-Network-Oriented V2V Data Transmission and Routing Considering RSUs and Connectivity Probability, *IEEE Transactions on Intelligent Transportation Systems (T-ITS)*, pp. 1-14, April, 2020, DOI: 10.1109/TITS.2020.2983835.

[63] H. Gao, W. Huang, X. Yang, Applying Probabilistic Model Checking to Path Planning in an Intelligent Transportation System Using Mobility Trajectories and Their Statistical Data, *Intelligent Automation And Soft Computing (Autosoft)*, Vol. 25, No.3, pp. 547-559, 2019.

## Biographies

**Shuanghu Luo** received a master's degree in computer software and theory in 2008 from Shanghai University, China, where he is currently pursuing a Ph.D. degree in computer applications with the School of Computer Engineering and Science. His research interests include computer networks and security, data analysis, and machine learning.



**Ling Yu** received a bachelor's degree in software engineering from Shanghai University of Electric Power, China. He is currently pursuing a graduate degree in computer science at Shanghai University of Electric Power, China. His research interests include computer vision and object detection.

**Zhongqin Bi** received a Ph.D. in system analysis and integration from East China Normal University, China, in 2009. He is currently an assistant professor at the College of Computer Science and Technology, Shanghai University of Electric Power, China. Dr. Bi has published more than 30 articles in refereed journals and conference proceedings. His main research interests are cloud computing, data processes and quality control in smart grids.

**Yongbin Li** received his MA. Sc in computer application technology from Shanghai Jiaotong University, China, in 2008. He is currently an assistant professor in the College of Computer Science and Technology, Shanghai University of Electric Power, China. Dr. Li has published more than 10 articles in refereed journals and conference proceedings. His main research interests are data processing and software engineering, and he has won the third prize of Anhui provincial science and technology progress.