Machine Vision Based Traffic Sign Detection Methods: Review, Analyses and Perspectives

CHUNSHENG LIU[1], **SHUANG LI**[1], **FALIANG CHANG**[1], **AND YINHAI WANG**[2]
[1]School of Control Science and Engineering, Shandong University, Jinan 250061, China
[2]Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195, USA

Corresponding author: Faliang Chang (flchang@sdu.edu.cn)

ABSTRACT Traffic signs recognition (TSR) is an important part of some advanced driver-assistance systems (ADASs) and auto driving systems (ADSs). As the first key step of TSR, traffic sign detection (TSD) is a challenging problem because of different types, small sizes, complex driving scenes, and occlusions. In recent years, there have been a large number of TSD algorithms based on machine vision and pattern recognition. In this paper, a comprehensive review of the literature on TSD is presented. We divide the reviewed detection methods into five main categories: color-based methods, shape-based methods, color- and shape-based methods, machine-learning-based methods, and LIDAR-based methods. The methods in each category are also classified into different subcategories for understanding and summarizing the mechanisms of different methods. For some reviewed methods that lack comparisons on public datasets, we reimplemented part of these methods for comparison. The experimental comparisons and analyses are presented on the reported performance and the performance of our reimplemented methods. Furthermore, future directions and recommendations of the TSD research are given to promote the development of the TSD.

INDEX TERMS Traffic sign detection (TSD), traffic sign recognition (TSR), object detection, neural networks (NN), support vector machine (SVM), AdaBoost.

I. INTRODUCTION

Computer vision and pattern recognition based traffic sign detection, tracking and classification methods have been studied for several purposes, such as Advanced Driver Assistance Systems (ADAS) and Auto Driving Systems (ADS). Generally, traffic sign recognition (TSR) systems consist of two phases of detection and classification; for some TSR systems, a tracking phase is designed between detection and classification for dealing with video sequences [1]. For TSR, camera and LIDAR are two most popular used sensing devices. In this paper, we review the literature on traffic sign detection (TSD) based on camera or LIDAR, and do comparison and analysis of the reviewed methods based on the reported performance and the performance of our reimplemented methods.

For a TSR system, traffic sign detection (TSD) usually is the first key process. TSD is a process of detecting and locating signs. Then, the detected traffic signs are utilized

as inputs of the following tracking or classification methods; hence, the accuracy of the traffic sign detection and locating results has a great influence on the following tracking or classification algorithms.

Though the structures and appearances of traffic signs are different across the world, the distinct color and shape characteristics of traffic signs provide important cues to design detection methods. In the past decades, many detection methods were designed based on detecting special colors such as blue, red and yellow [2]; these methods were commonly used for preliminary reduction of the search space, followed by some other detection methods. Shape or edge detection methods are also popular in the detection literature. Different shape detection methods are designed to detect circle, triangle or octagon. Shape and edge detection methods can also be used to extract the accurate position of a traffic sign.

In recent years, with the development of machine learning methods especially deep learning methodologies, the machine learning based detection methods have gradually become the mainstream algorithms. There are three main traffic sign detection structures: AdaBoost based detection [3],

The associate editor coordinating the review of this manuscript and approving it for publication was Bora Onat.

Support Vector Machine (SVM) based detection [4], and Neural Networks (NN) based detection [5]. These detection structures have many derivatives with different input features, different training methods or different detection processes. The machine learning based detection methods have achieved the-state-of-the-art results in some aspects [6].

In some TSR systems, a tracking method is needed. The goal of traffic sign tracking is usually designed for boosting classification performance, fine-positioning or predicting positions for detection in the next frame.

After traffic sign detection or tracking, traffic sign recognition is performed to classify the detected traffic signs into correct classes. The main classification methods include binary-tree-based classification, SVM, NN and Sparse Representation Classification (SRC), etc. The binary-tree-based classification method usually classify traffic signs according to the shapes and colors in a coarse-to-fine tree process. As a binary-classification method, SVM classifies traffic signs using one-vs-one or one-vs-others classification process. SRC and NN belong to binary-classification methodology and can recognize multiclass traffic signs directly.

In the past decade, there are some surveys on TSR. Fu and Huang [7] reviews part of the TSD methods before 2010; most of the reviewed methods in [7] are out of date. Møgelmose *et al.* [1] presents a comprehensive survey for TSD, which covers popular detection methods before 2012. Gudigar *et al.* [8] and Saadna and Behloul [9] present reviews for both detection and recognition. These two reviews list limited reported results for detection and lack comprehensive comparisons and summaries of their reviewed detection methods. Furthermore, all previous surveys do not review the LIDAR based methods. Distinguished from these previous surveys, we classify the reviewed methods into fine categories, reimplement part of the TSD methods for comprehensive comparisons of these methods, and also review the LIDAR based TSD methods. In this survey, we mainly review the TSD methods in last five years, and give analyses and future research suggestions.

This paper is organized as follows. Section II presents the introduction of traffic signs, influence to human driving safety, machine vision based TSR system and its applications, and benchmarks for TSR. Section III shows overview of traffic sign detection; traffic sign detection methods are classified into five categories: color based methods, shape based methods, color and shape based methods, machine learning based methods, and LIDAR based methods. From Section IV to Section VIII, the methods in these five categories are reviewed. Section IX gives the conclusions and perspectives.

II. TRAFFIC SIGN

Traffic signs are placed along the roads with the function of informing drivers about the front road conditions, directions, restrictions or text information. Though traffic signs have different structures and appearances in different countries, the most essential types of traffic signs are prohibitory, danger, mandatory and text-based signs. The prohibitory, danger

FIGURE 1. Different types of traffic signs from germany, china and america. (a) German signs, (b) Chinese signs, (c) American signs. Signs from Germany and China are classified into prohibitory signs, danger signs, mandatory signs and other types of signs. American signs are classified into regulatory signs, warning signs, guide signs and other signs according to Wikipedia. More signs from these three countries can be found in German GTSDB dataset [6], Chinese TT100K dataset [10], and American LISA dataset [11].

or mandatory signs often have standard shapes, such as circle, triangle and rectangle, and often have standard colors such as red, blue and yellow. The text-based signs usually do not have fixed shapes and contain informative text. In Fig. 1, we list

some types of German signs, Chinese signs, and American signs. Signs from Germany and China are classified into prohibitory signs, danger signs, mandatory signs and other types of signs. American signs are classified into regulatory signs, warning signs, guide signs and other signs. More signs from these three countries can be found in German GTSDB dataset [6], Chinese TT100K dataset [10], and American LISA dataset [11]. In this section, we firstly describe the importance of traffic signs for human driving safety and then describe the machine vision based TSR systems and their applications; lastly, benchmarks for TSR are listed.

A. TRAFFIC SIGNS FOR HUMAN DRIVING SAFETY

Though traffic signs play an important role in traffic safety and regulating drivers' behavior, they are often unattended. In the study of [12], Costa *et al.* show that different types of signs have different ability to capture the attention of drivers. During gazing, the drivers may not remember the content of a sign or may miss some other important signs.

During driving, traffic signs with different distances and different presentation times have different influences on the accuracy of sign identification for human drivers [12]; the study in [12] shows the drivers have 75% accuracy with less than 35 ms presentation time and 100% accuracy with 130 ms presentation time; this study also shows the drivers need enough time to correctly recognize the signs in front.

According to [13], the sign context and drivers' age have effect on traffic sign comprehension; their experiments show that younger drivers perform better than older drivers on both accuracy and response time, and that the sign context increase the comprehension time.

B. MACHINE VISION BASED TSR SYSTEMS AND THEIR APPLICATIONS

Based on some types of sensing devices, such as on-board cameras and LIDAR, different TSR systems can be designed for traffic sign detection, classification and result presentation. For a TSR system, the key stages are detection and classification. The detection stage can detect and locate traffic signs; the detection and localization accuracy largely affects the following processing. Then, the classification stage can classify the detected traffic signs into different types and output the results of TSR. In some systems, a tracking stage is needed for processing consecutive frames.

Some structures of TSR are shown in Fig. 2. Fig. 2 (a) is the most popular camera based TSR structure without tracking; this structure can detect and recognize traffic signs in a single frame without using any temporal information from videos. Fig. 2 (b) is a camera based structure with tracking described in [1]; this structure can consecutively confirm the tracking results in consecutive frames to boost classification performance. Fig. 2 (c) is a camera based TSR structure with tracking for fine-positioning [14]; in this structure, the tracking results are used for fine-positioning and classification. Fig. 2 (d) is a camera based structure with tracking for position prediction [15]; the multi-ROI tracking
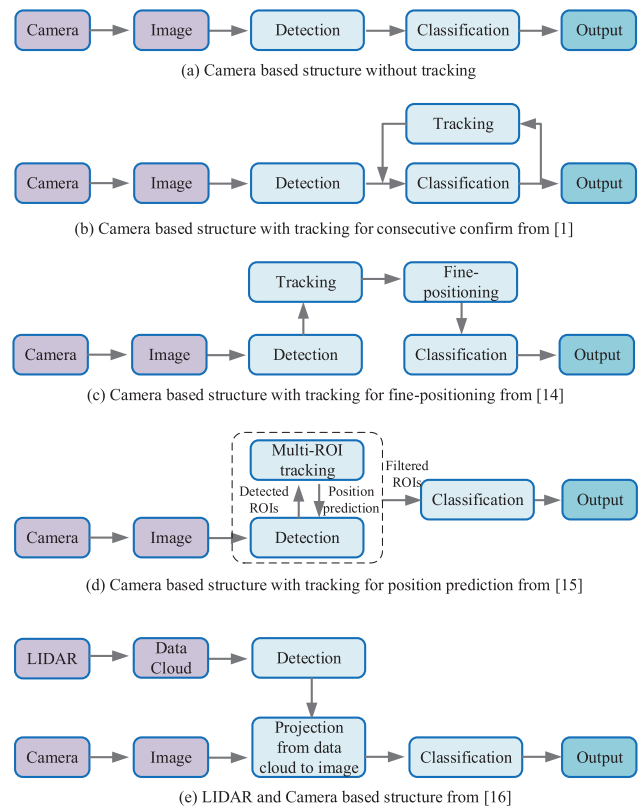

FIGURE 2. Different structures of traffic sign recognition systems.

process in this structure is utilized for position prediction and getting filtered ROIs for classification. Fig. 2 (e) is a common LIDAR and camera based TSR structure [16]; the data cloud from laser scanning is utilized for traffic sign detection; the detection results in data clouds are projected into images captured by camera; then, classification is processed with the detected signs in the projected images.

TSR systems have various well-defined applications. We summarize some reported TSR applications in recent years.

1) Driver-assistance systems. In the literature on TSR, a large proportion of methods are for assisted driving. A TSR driver-assistance system can assist the driver by informing the contents of traffic signs ahead, including restrictions, warnings, and limits. There have been some commercial products for assisted driving.

2) Autonomous vehicles. In the past decade, many companies and research labs focused on designing their autonomous vehicles. The TSR system is a very important part for autonomous vehicles, making the autonomous vehicle know the current traffic regulations in public roads.

3) Maintenance of traffic signs. TSR systems can be used for maintenance of traffic signs or roads. In [17] and [18], TSR systems were utilized to check the condition of traffic signs along the major roads. Wen *et al.* [19] utilized mobile laser scanning data for spatial-related traffic sign inspection. The luminance and reflectivity of traffic signs were evaluated with a camera to fulfill the purpose of automatic recognizing

TABLE 1. Publicly available datasets.

Dataset	Details of the public datasets					
	Purpose	Classes	Total images	Total Signs	Image Sizes	Country
GTSDB	Detection	43	900	1, 213	$1,360 \times 800$	Germany
GTSRB	Recognition	43	50, 000+	50, 000+	15×15 to 250×250	Germany
BTSD	Detection	62	25, 634	4, 627	$1,628 \times 1,236$	Belgium
BTSC	Recognition	62	7, 125	7, 125	26×26 to 527×674	Belgium
TT100K	Detection/Recognition	45	100, 000	30, 000	$2,048 \times 2,048$	China
LISA	Detection/Recognition	49	7, 855	6, 610	640×480 to $1,924 \times 522$	United States
STS	Detection/Recognition	7	20, 000	3, 488	$1,280 \times 960$	Sweden
RUG	Detection/Recognition	3	48	60	360×270	Netherland
Stereopolis	Detection/Recognition	10	847	251	$960 \times 1,080$	France
FTSD	Detection/Recognition	N/A	4, 239	N/A	640×480	Sweden
MASTIF	Detection/Recognition	N/A	4, 875	13,036	720×576	Croatia
ETSD	Recognition	164	82, 476	82,476	6×6 to 780×780	Europe

deteriorated reflective sheeting material of which the traffic signs were made [20].

4) Engineering measurements. In [21], detection and recognition of traffic signs in Google Street View (GSV) were used to automatically extract traffic sign locations for engineering measurements.

5) Vehicle-to-X (V2X) communication. Traffic sign is an important scatterer for vehicle-to-X (V2X) communication scenarios, and can affect the propagation channel appreciably. Guan *et al.* [22] presented an integration of the full-wave simulation, analytical models, measurement, and validation of the bistatic radar cross section of three types of representative traffic signs for V2X communication.

6) Reducing fuel consumption. Based on detecting some certain types of signs ahead, Muñoz-Organero *et al.* [23] implemented and validated an expert system that can reduce fuel consumption by detecting optimal deceleration traffic signs, minimizing the use of braking.

C. BENCHMARKS FOR TSR
We describe the public benchmarks for TSR. Because signs from different countries are usually different, it is difficult to compare the TSR methods designed for different countries. The public datasets provide benchmarks for comparison. Table 1 provides a summary of these publicly available datasets. The detailed descriptions of these public datasets are as follows.

1) GERMAN TRAFFIC SIGN BENCHMARK (GTSB)
The GTSB has two datasets including German Traffic Sign Detection Benchmark (GTSDB) [6] and German Traffic Sign Recognition Benchmark (GTSRB) [24]. The GTSDB and GTSRB datasets were created for the competition of detection and recognition of German traffic signs. Both GTSDB and GTSRB are large comprehensive datasets, which have been widely used in training and testing of different TSR methods. The GTSDB includes 600 images for training and 300 images for testing. The GTSRB includes more than

50, 000 traffic signs with different illuminations, sizes and directions, for training and testing.

2) BELGIUMTS (BTS) DATASET [25]
The BelgiumTS dataset has two datasets including the BelgiumTS detection dataset (BTSD) and the BelgiumTS classification dataset (BTSC). The BTSD and BTSC are large comprehensive datasets for detection and classification respectively. The BTSD dataset provides only partially annotated positive images. There are total 25, 634 images in BTSD including 5, 905 annotated training images and 3, 101 annotated testing images. The BTSC includes 4, 591 training images and 2, 534 testing images.

3) TSINGHUA-TENCENT 100K (TT100K) [10]
Appeared in 2016, the TT100K dataset includes 100,000 images with 30,000 traffic signs. Each traffic sign in this benchmark is annotated with a class label, its bounding box and pixel mask. The images in this dataset are with 2048 × 2048 resolution and cover relative large variations in illuminance changes and weather conditions.

4) LISA DATASET [11]
The LISA dataset is a large dataset for American signs, which includes video tracks of all the annotated signs. There are 7, 855 images with 6, 610 signs. This dataset can be used to verify detection or tracking methods.

5) SWEDISH TRAFFIC SIGNS (STS) DATASET [26]
The STS dataset includes more than 20,000 images with 3,488 annotated signs. The STS includes all frames from the videos, which means that both detection and tracking methods can be test on this dataset.

6) RUG DATASET [27]
This small dataset contains 48 images with 360 × 270 pixels, each representing a traffic scene. The images are grouped in 3 classes including pedestrian crossing, intersection and compulsory for bikes.

7) STEREOPOLIS DATABASE [28]
The Stereopolis dataset is made of 847 images with 960 × 1080 resolution of complex urban scenes in France.

8) FLEYEH TRAFFIC SIGNS DATASET (FTSD) [29]
This dataset consists of 4, 239 image of traffic scenes with 640 × 480 resolution. It is collected in different parts of Sweden.

9) MAPPING AND ASSESSING THE STATE OF TRAFFIC INFRASTRUCTURE (MASTIF) DATASET [18]
This dataset consists of three small datasets collected in 2009, 2010 and 2011, respectively. The dataset-2009 is a classification dataset containing 6, 423 signs from 97 classes. The dataset-2010 has 3, 862 images with resolution of 720 × 576 pixels; there are 5, 184 signs from 88 classes. The dataset-2011 contains 1, 013 images with resolution of 720 × 576 pixels; there are 1, 429 signs from 53 classes.

10) EUROPEAN TRAFFIC SIGN DATASET (ETSD) [91]
The ETSD dataset is composed with different traffic signs datasets that were captured from different European countries, including GTSRB [24], BTS [25], STS [26], RUG [27], Stereopolis [28], FTSD [29], and MASTIF [18]. The ETSD dataset has 82, 476 signs with 164 classes. The signs in ETSD have sizes varying between 6 and 780 pixels. The ETSD dataset contains many images with different lighting conditions, occlusions, motion blur, human made artifacts and perspectives.

III. OVERVIEW OF TRAFFIC SIGN DETECTION
In [1] and [9], the traffic sign detection (TSD) methods are classified into two categories including shape based methods and color based methods. Now, it have been commonly accepted that the machine learning methods have some superiorities over the traditional color or shape based methods in some aspects. The machine learning methods often need a large amount of training samples with informative both color and shape information. Besides machine learning methods, there also some methods designed based on both color and shape characteristics. It is not appropriate to classify these TSD methods into color or shape based methods. Furthermore, LIDAR based methods have developed rapidly in recent years, and previous review methods did not review LIDAR based TSD methods.

In this review, we divide the traffic sign detection methods into five categories: color based methods, shape based methods, color and shape based methods, machine learning based methods, and LIDAR based methods. The color based methods are the methods mainly designed with color information. The shape detection based methods are the methods mainly designed with shape information. The color and shape based detection methods are the methods designed with both color and shape information. Though some color or shape based methods are designed with machine learning methods;

in this review, we still classify these methods into the category of color based methods or the category of shape based methods. The category of machine learning based methods in this classification means the TSD methods directly designed for detecting signs not for detecting colors or shapes. The category of LIDAR based methods means the TSD methods designed to deal with point cloud data captured by LIDAR. The methods in these five categories are reviewed and analyzed in the following sections.

IV. COLOR BASED DETECTION METHODS
The distinct color characteristics of traffic signs can attract drivers' attention and can also provide important cues to design color based detection methods. In the past decades, a large amount of detection methods are designed to detect distinct traffic sign colors such as blue, red and yellow. These methods can be directly used for traffic sign detection, and can also be used for preliminary reduction of the search space, followed by other detection methods. This section reviews and gives comparisons of the color based detection methods.

A. REVIEW OF COLOR BASED DETECTION METHODS
In this subsection, different color based detection methods are classified into five categories. Color based detection methods are summarized in Table 2. The details are reviewed as follows.

TABLE 2. Color based detection methods.

	Category	Paper	Year	Method	Detected colors
Color Based Detection Methods	RGB based thresholding	[2]	2010	Normalized RGB thresholding	Red, blue, yellow
		[30]	2010	Color Enhancement	Red, blue, yellow
		[31]	2015	Color Enhancement	Red, blue, yellow
	Hue and saturation thresholding	[2]	2010	Hue and saturation thresholding	Red, blue, yellow
		[33]	2004	LUTs based HS thresholding	Red, blue, yellow
	Thresholding on other spaces	[2]	2010	Ohta thresholding	Red, blue, yellow
		[34]	2015	Lab thresholding	Red, blue, yellow, green
	Chromatic/Achromatic Decomposition	[2]	2010	RGB, HIS, Ohta decomposition	white
		[34]	2015	RGB based achromatic segment	white
	Pixel classification	[2]	2010	SVM classification	Red, blue, yellow
		[36]	2012	Probabilistic neural networks	Red, blue, yellow

1) RGB BASED THRESHOLDING
Using the channels in some color space to do thresholding is the most intuitive way to segment some special colors. Selecting a suitable color space is a key point to these methods. The RGB space is the most basic color space for images and videos captured by cameras. Though RGB can be used with no transformation, the R, G and B channels have high correlation and are sensitive to illumination changes. It is difficult to robustly segment a special color with some fixed thresholds in RGB space.

One popular solution is the use of a normalized version of RGB (NRGB) with respect to $R + G + B$. In the NRGB space, different illuminations have little effect on the pixel values; and two channels are enough to perform classification because the rest channel can be obtained with these two channels. The masks for each color can be obtained as $Red(i, j)$,

Blue(x, y) and *Yellow(i, j)* [2]:

$$Red(i,j) = \begin{cases} \text{True,} & \text{if } r(i,j) \geq ThR \\ & \text{and } g(i,j) \leq ThG \\ \text{False,} & \text{otherwise} \end{cases}$$

$$Blue(i,j) = \begin{cases} \text{True,} & \text{if } b(i,j) \geq ThB \\ \\ \text{False,} & \text{otherwise} \end{cases}$$

$$Yellow(i,j) = \begin{cases} \text{True,} & \text{if } (r(i,j) + g(i,j)) \geq ThY \\ \\ \text{False,} & \text{otherwise.} \end{cases} \quad (1)$$

where, *r*, *g* and *b* are the normalized red, green and blue channels; *ThR*, *ThG*, *ThB* and *ThY* are the fixed thresholds, which can be found in [2].

Ruta *et al.* [30] enhanced colors with maximum and minimum operations using RGB values. For each RGB pixel $i = [i_R, i_G, i_B]$ and $s = i_R + i_G + i_B$, a set of transformations [30] is,

$$f_R(i) = max(0, min(i_R - i_G, i_R - i_B)/s),$$
$$f_B(i) = max(0, min(i_B - i_R, i_B - i_G)/s),$$
$$f_Y(i) = max(0, min(i_R - i_B, i_G - i_B)/s). \quad (2)$$

After the transformation in formula (2), the red, blue and yellow colors can be enhanced in their corresponding enhanced images. Thresholds can be used in an enhanced image to extract a special color. Yet, the blue mandatory signs with very dark or bright illumination have similar values in the blue and green channels, which may result in failure in extracting blue color with formula (2). Salti *et al.* [31] did not consider the strength of the blue with respect to the green and changed the enhanced blue channels accordingly to,

$$f_B'(i) = max(0, i_B - i_R)/s). \quad (3)$$

2) HUE AND SATURATION THRESHOLDING

The hue and saturation channels in HSV color space or HSI color space are more immune to illumination changes than RGB [32]. The hue and saturation channels can be calculated using RGB, which increases the processing time. The hue and saturation channels based methods are usually simple and partly immune to illumination changes. One drawback is that the instability of hue may result in unsatisfactory results in different scenes [2].

The output for extracting different colors using hue and saturation thresholding are as [2],

$$Red(i,j) = \begin{cases} \text{True,} & \text{if } H(i,j) \leq ThR_1 \\ & \text{or } H(i,j) \geq ThR_2 \\ \text{False,} & \text{otherwise} \end{cases}$$

$$Blue(i,j) = \begin{cases} \text{True,} & \text{if } H(i,j) \geq ThB_1 \\ & \text{and } H(i,j) \leq ThB_2 \\ \text{False,} & \text{otherwise} \end{cases}$$

$$Yellow(i,j) = \begin{cases} \text{True,} & \text{if } H(i,j) \geq ThY_1 \\ & \text{and } H(i,j) \leq ThY_2 \\ & \text{and } H(i,j) \leq ThY_3 \\ \text{False,} & \text{otherwise.} \end{cases} \quad (4)$$

where, *H* and *S* are the hue and saturation channels; ThR_i, ThB_i and ThY_i are the fixed thresholds, which can be found in [2].

In order to avoid rigid thresholding [33], a soft threshold method based on two lookup tables (LUTs) was presented to extract red and blue colors in the Hue and Saturation channels. In the method described in [33], color extraction was achieved by using three LUTs and then thresholds are applied to get extracted results.

3) THRESHOLDING ON OTHER SPACES

There are some methods that are designed based on some other color spaces, such as Ohta, L*a*b* and XYZ. With the purpose finding uncorrelated color components, the Ohta space was used to extract red, blue and yellow colors [2]. In [34], a K-means clustering method was used for detecting blue, red, yellow, and green colors on the L*a*b* space.

4) CHROMATIC/ACHROMATIC DECOMPOSITION

Most color based detection methods are designed for significant colors including red, blue and yellow. The chromatic/achromatic decomposition methodology tries to find the pixels with no color information. A detailed description of these methods with five categories [2] is: chromatic/achromatic index method, RGB differences method, normalized RGB differences method, saturation and intensity based method and Ohta components based method. In each category, different thresholds are adopted on different color spaces to extract white traffic sign color. Lillo-Castellano *et al.* [34] combined L*a*b* space, HSI space and RGB space to detect white color.

5) PIXEL CLASSIFICATION

The thresholding methods based on some color spaces often have some thresholds to be adjusted. The adjustment of these thresholds depends on the trained images and usually does not have enough generalization ability. Some authors tried to transfer the color extraction problem into pixel classification problem, and used classification methods to classify each pixel in the input image.

The SVM classification method was used to classify color pixels from background pixels in [35] and [2]. In [36], the input pixel values were used to train a neural network for color pixel classification. These methods first get color vectors from some color spaces and then use the color vectors to train a SVM based or a NN based classifier. With a process to classify every pixel in the input image, the pixel classification algorithms are often slower than other color extraction methods.

B. ANALYSIS OF THE COLOR BASED METHODS

In this subsection, we compare the color based detection methods. Because most of the reviewed color extraction methods gave their results on some datasets that are not publicly available, we reimplemented some methods with detailed reported steps and parameters, and tested them on the public GTSDB dataset to give a comprehensive comparison.

In this comparison, the reimplemented color extraction methods includes NRGB thresholding method [2], HSI thresholding method [2], Ohta thresholding method [2], and HSV thresholding method [37]. The parameters we used were the same as those used in the original references. The detailed parameters can be found in the corresponding references. The test dataset is the GTSDB dataset. The results are reflected in several parameters including detection rate (DR) and extraction rate (ER). DR is the ratio of the number of detected objects to the number of all objects. ER is the ratio of the pixel number of extracted regions to the pixel number of the input image. A detection result is considered true if the IoU (Intersection over Union) is more than 50%. Comparison results are shown in Table 3.

TABLE 3. Comparison of color based detection methods.

Methods	Blue		Red	
	DR (%)	ER (%)	DR (%)	ER (%)
HSI	97.14%	35.35%	97.23%	34.57%
HSV	N/A	N/A	88.92%	30.20%
Ohta	84.76%	23.60%	95.56%	3.50%
NRGB	92.38%	9.02%	90.03%	1.15%

The HSI based method achieves the best DRs of 97.14% and 97.23% on detecting blue and red colors respectively. Yet, the HSI based method achieves more than 30% ERs which are usually too large for a ROI extraction process. The Ohta based method achieves good results on detecting red color with a low 3.50% ER and a high 95.56% DR, yet fails to extract blue color with a high 23.60% ER and a relative low DR of 84.76%. The NRGB thresholding method can keep a relative good balance of DR and ER, achieving 92.38% DR and 9.02% ER on blue color detection, and 90.03% DR and 1.15% ER on red color detection. The experiment results in Table 3 show that the methods using fixed thresholds on some color spaces can not achieve good performance in both DR and ER. One reason of this poor performance is that color thresholding is sensitive to various factors, such as illumination changes, different colors, time of day and reflection of signs' surface. The other reason is that the thresholds used by the methods in Table 3 are published for color extraction in some special datasets from different countries and may result in relative bad results when dealing with other datasets.

V. SHAPE BASED DETECTION METHODS

Common standard shapes of traffic signs are triangle, circle, rectangle, and octagon. Shape characteristics used for shape detection include standard shapes, boundaries, texture, key points, etc.

A. REVIEW OF SHAPE BASED DETECTION METHODS

In this subsection, we classify the shape based detection methods into four categories and review them as follows. Shape based detection methods are summarized in Table 4.

TABLE 4. Shape based detection methods.

	Category	Paper	Year	Method	Detected shapes
Shape Based Detection Methods	Shape detection	[38]	2015	Hough	Circle and triangle
		[39]	2008	Radial symmetry transform	Circle
		[86]	2004	Radial symmetry transform	Polygons
	Shape analysis and matching	[41]	2003	Complex shape models	Circle, polygons
		[42]	2008	Shape decomposition	Circle, square, triangle
	Fourier transformation	[26]	2011	Fourier descriptors	Circle, square, triangle
		[43]	2008	Fast Fourier Transformation	Circle, square, triangle
	Key points detection	[45]	2014	SIFT	Circle, square, triangle, octagon
		[15]	2014	Harris corner	Circle, triangle
		[46]	2014	Interest points clustering	Different shapes

1) SHAPE DETECTION

Shape detection methods are usually designed for traffic sign detection with standard shapes. The shape detection techniques such as Hough detection [38] are utilized to detect special shapes. The Hough based methods are usually slow to compute over large images.
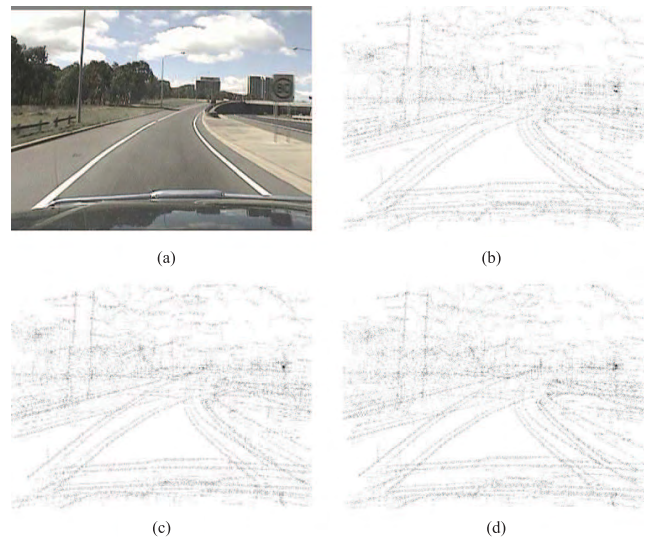

(a) (b)

(c) (d)

FIGURE 3. Radial symmetry voting method [39]. (a) is the input image with a road sign. (b), (c) and (d) are sample radial symmetry images for the three largest radii. Detection peaks appear in (c).

Derived from Hough method, Barnes *et al.* [39] designed a more efficient method for speed sign detection, called fast radial symmetry. The fast radial symmetry utilizes radial symmetry voting mechanism to detect symmetry shapes, which is robust to un-occluded shapes and run as a detector faster than Hough. The fast radial symmetry was also utilized to detect traffic signs with polygonal shapes [40]. The radial symmetry voting results in [39] are shown in Fig. 3.

2) SHAPE ANALYSIS AND MATCHING

The analysis and matching of different shapes can be used to detect signs with significant edges. Fang *et al.* [41] designed different complex shape models for circular signs, triangular signs and octagonal signs. The manually designed shape models reply on distinct edges and are often sensitive to noises and shape changes.

A decomposition method was designed in [42] to represent complex shapes using multiple simpler components. The decomposition of complex shapes are determined by maximal supported convex arcs, which can partition several connected traffic signs and remove the internal contents.

3) FOURIER TRANSFORMATION

Fourier transformation provides a useful way to represent traffic sign shapes. Larsson and Felsberg [26] utilized Fourier descriptors to express traffic signs and then combined locally segmented contours to detect different traffic signs.

Larsson *et al.* [45] designed traffic sign detection method based on Fourier Transformation. Arroyo [43] utilized Fast Fourier Transformation (FFT) analysis to express different shapes of traffic signs and then adopted a triangle normalization and reorientation algorithm to locate sign positions.

4) KEY POINTS DETECTION

Singularities or angular edges of traffic signs can be detected by key points detection methods to represent signs. Scale-invariant feature transform (SIFT) local descriptor [44] is a popular scale-invariant and rotation-invariant key point description. Boumediene *et al.* [15] utilized Harris corner detector to detect corners of traffic signs. For every corner, a candidate ROI can be selected according to the shapes in the corresponding corner neighborhood. Khan *et al.* [46] utilized Gabor filter to extract stable local features of the detected interest points, and then designed a clustering method to detect traffic signs.

B. ANALYSIS OF THE SHAPE BASED METHODS

In this subsection, we do analysis of the shape based detection methods. Four shape based methods that have reported detailed performance are listed in Table 5. These four methods include the method based on Fourier descriptors and spatial models (F&S) [26], the method based on correlating Fourier descriptors of local patches (FL) [45], Gabor based method [46] and Harris based method [15]. The performance is measured by recall and precision for the methods in [26], [45] and [46], and measured by detection rate (DR) and false positives per frame (FPPF) for the method in [15].

The methods in [26], [45] and [46] have recall values of 94.23%, 95.37% and 96.33% respectively, and have precision values of 78.54%, 95.40% and 89.75% respectively; these results mean that the methods in [26], [45] and [46] did not have good results on their own small-size datasets. The Harris method [15] has a DR value of 89.92% and an FPPF value of 0.16 on a video dataset with 2, 850 signs.

TABLE 5. Comparison of some shape based detection methods.

Methods	Performance	Dataset
F&S [26]	Recall: 94.23%, Precision: 78.54%	641 signs
FL [45]	Recall: 95.37%, Precision: 95.40%	216 signs
Gabor [46]	Recall: 96.33%, Precision: 89.75%	300 signs
Harris [15]	DR: 89.92%, FPPF: 0.16	2,850 signs

The results in Table 5 show that these shape based methods did not achieve high performance on their own datasets. There are two main reasons. The shape based detection methods [26], [45] often rely on distinguished edges and may fail when detecting signs with small size or vague edges. Without using edge information, the key point detection methods [46], [15] need distinguished points or corners, which may bring instability when detecting vague signs.

Besides these four listed methods in Table 5, the other shape based methods often have similar advantages and shortcomings. Compared with the color based detection methods, the shape based detection methods are often robust to color changes. The main shortcoming is that the shape based detection methods are often sensitive to small-size and vague signs. For example, the Hough transform method [38] and the shape matching method [42] need edge detection first, the performance of which highly relies on distinguished edges. Hence, the shape based detection methods are often used to deal with traffic signs with distinguished edges.

In some methods, the shape detection methods can be combined with color based methods to fulfill the traffic sign detection work. For instance, the methods described in [41] and [42] utilized color based methods to extract candidates and then designed shape based methods to detect signs from the extracted candidates. The method in [46] utilized Harris to extract features and then designed data association method to associate the features in consecutive frames.

VI. COLOR AND SHAPE BASED METHODS

In this section, we review the detection methods using both color and shape characteristics. A large number of TSD structures are combined with some phases; the method in each phase is designed based on color or shape. The color and shape based methods in this review mean the methods designed based on both color and shape characteristics instead of logical combination of different phases.

A. REVIEW OF COLOR AND SHAPE BASED DETECTION METHODS

We classify the color and shape based detection methods into three categories and review them as follows. Color and shape based detection methods are summarized in Table 6.

1) EXTREME REGIONS BASED DETECTION

The Maximally stable extremal regions (MSERs) method detects high-contrast regions of approximately uniform

TABLE 6. Color and shape based detection methods.

	Category	Paper	Year	Method	Detection
Color and Shape Based Detection Methods	Extreme regions based detection	[47]	2015	MSERs, and Hue and saturation thresholding	Text-based sign plates
		[48]	2012	Enhanced red/blue image and MSERs	Red and blue signs
		[49]	2016	Enhanced colors with Gaussian distribution, and MSERs	Signs with different colors
	High contrasted margin detection	[50]	2016	High Contrast Region Extraction	Red, blue and yellow signs
	Saliency detection	[52]	2017	Center-surround saliency test	Different traffic signs
		[87]	2018	Channel-wise feature responses	Different traffic signs

gray tone and arbitrary shape, and is therefore likely to extract colored regions within traffic signs. Greenhalgh and Mirmehdi [47] utilized MSERs to locate a large number of candidate regions and then utilized hue, saturation, and value color thresholding to detect the text-based traffic sign regions.

Instead of detecting MSERs from a grayscale image, Greenhalgh and Mirmehdi [48] designed a MSERs extraction method based on color enhancement images. This method firstly transforms the RGB space into normalized red/blue image, and then utilizes the MSERs to extract red and blue regions. This method is named MSERs_NRB in this review. The greater of the pixel values of the normalized red and blue channels are used to form a red/blue enhanced

image Ω_{RB},

$$\Omega_{RB} = max(\frac{R}{R+G+B}, \frac{B}{R+G+B}). \quad (5)$$

Then, MSERs method is utilized to extract extreme regions on the image Ω_{RB}. This method has robust results on regular red and blue colors, and is not designed for other colors. The extracted results of our re-implemented MSERs_NRB are shown in Fig. 4.

Yang *et al.* [49] designed a color probability model to enhance traffic sign colors using Ohta space and Gaussian distribution; then, the MSERs method is utilized to extract ROIs. The extraction results of red color and blue color are shown in Fig. 5. Unlike MSERs_NRB [48], this method can enhance different colors and get enhanced image for each color.

Salti *et al.* [31] designed color enhancement methods to enhance red, blue and yellow colors, then utilized MSERs and Wave-based Detector (WaDe) to extract traffic sign regions. We re-implemented the color enhancement and MSERs based extraction method in [31]. The extracted results are shown in Fig. 6

These MSERs based methods in [48], [49] and [31] rely on the color enhancement results and suitable parameters of MSERs. Hence, the color enhancement methods and

(a) Original image (b) Enhanced red and blue (c) MSERs detection

FIGURE 4. Re-implemented MSERs_NRB detection method in [48]. (a) is the original input image. (b) red and blue enhanced image of MSERs_NRB. (c) detection result of MSERs_NRB. The results were gotten with a threshold of 1.5.
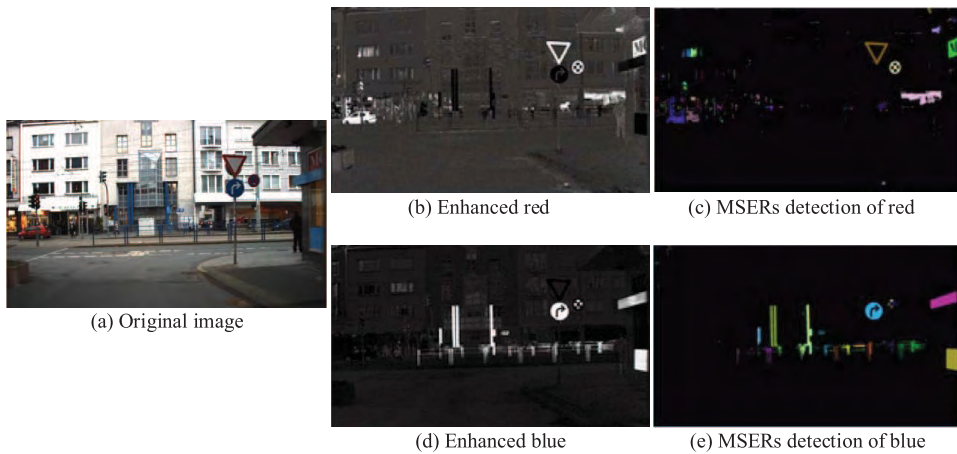

(a) Original image

(b) Enhanced red (c) MSERs detection of red

(d) Enhanced blue (e) MSERs detection of blue

FIGURE 5. Re-implemented color enhancement and MSERs based detection in [49]. (a) is the original input image, (b) is the red enhanced image with the method in [49], (c) is the MSERs detection result of red color, (d) is the image of enhanced blue with the method in [49], (e) is the MSERs detection result of blue. The results were gotten with a threshold of 5.0.
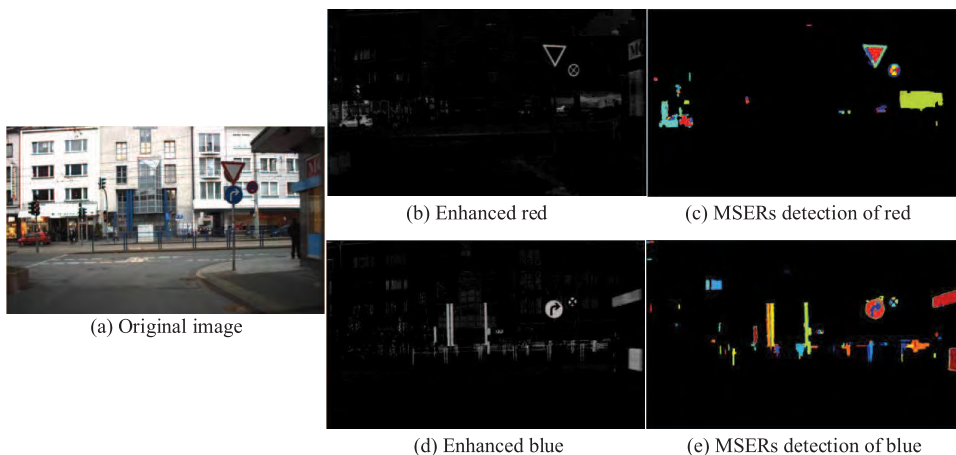
(a) Original image

(b) Enhanced red

(c) MSERs detection of red

(d) Enhanced blue

(e) MSERs detection of blue

FIGURE 6. Color enhancement and MSERs based detection in [31]. (a) is the original input image, (b) is the red enhanced image with the method in [31], (c) is the MSERs detection result of red color, (d) is the image of enhanced blue with the method in [31], (e) is the MSERs detection result of blue.
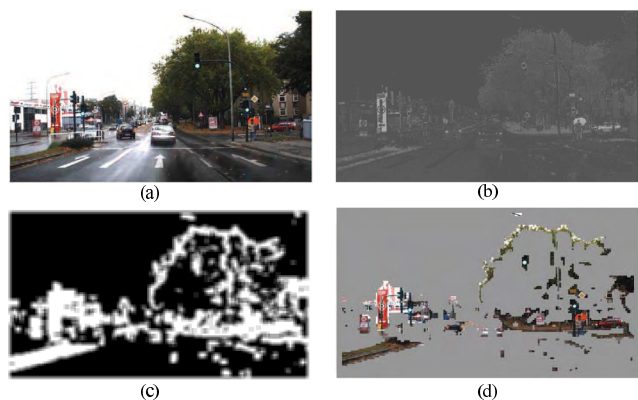

(a)

(b)

(c)

(d)

FIGURE 7. The process of the ROI extraction of HCRE [50]. (a) is the input RGB image. (b) is the image after color enhancement. (c) is the saliency image. (d) is the result image that is to show the extraction results; the gray region in (d) is background.

suitable parameters of MSERs are very important for these methods.

2) HIGH CONTRASTED MARGIN REGIONS BASED DETECTION

Unlike MSERs that can detect high-contrast regions of approximately uniform gray tone and arbitrary shape, High Contrast Region Extraction (HCRE) was designed to detect high contrasted margin regions [50]. This method first enhances the red, blue and yellow colors using RGB space transformation to enhance the contrast between these colors and their surrounding regions; and then a high contrast region detection method is designed to detect the margin regions with high contrast. The extraction process of HCRE is shown in Fig. 7.

3) SALIENCY DETECTION

In [51], the center-surround saliency method was designed to extract saliency regions of traffic signs based on the assumption that saliency can be reflected by local contrast.

This method calculates two cell-level saliency maps based on two types of features including compressed Histograms of Oriented Gradients (HOG) and non-normalized HOG (without block-based normalization). In this method, the HOG features can be extracted from gray images or different color channels.

Yuan *et al.* [52] proposed an algorithm that combines the information of color, saliency, spatial, and contextual relationship. In this method, a graph was firstly designed to represent images; then, a ranking algorithm was designed to exploit the intrinsic manifold structure of the graph nodes, giving each node a ranking score according to its saliency; finally, a multithreshold segmentation approach was proposed to segment traffic sign candidate regions.

B. ANALYSIS OF THE COLOR AND SHAPE BASED DETECTION METHODS

Because most of the color and shape based detection methods are tested on different datasets, it is useful to compare these methods on the same public datasets. In this review, two MSERs based detection methods with details in their published papers were reimplemented, and were compared in detailed curves on GTSDB in Fig. 8. The thresholds ranging from 0 and 50 were used to get these curves.

MSERs presents a good technique to get color ROIs. The methods in [48], [49] and [31] first enhance color using some methods and then utilize MSERs to detect color candidates on the enhanced images; lastly, classification methods are utilized to classify these candidates resulting traffic sign detection results. Combined with MSERs, the WaDe was used for color ROIs extraction in [31]. We reimplemented the MSERs_NRB in [48] and the color enhancement and MSERs (MSER_NRGB_Enhancement) method in [31]. In the reimplemented MSER_NRGB_Enhancement, we did not combine WaDe because the method in [31] lacking details about the combination of WaDe and MSERs. The 300 test images in GTSDB are used for testing. In Fig. 8, the curves show the
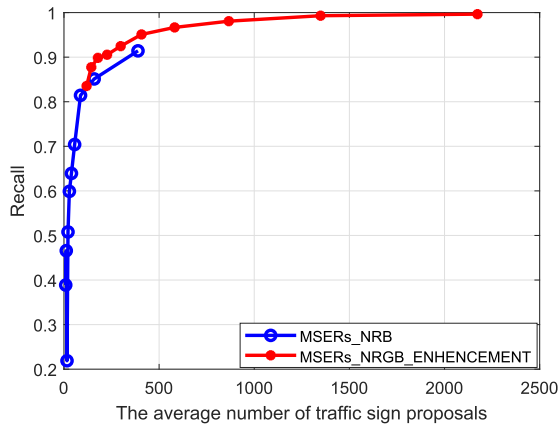
FIGURE 8. Comparison of different MSERs based methods.

TABLE 7. Comparison of different color based methods and color and shape based methods from [50].

Methods	Recall (%)			ER (%)
	Red	Blue	Yellow	
RGBE [30]	70.56%	78.21%	36.84%	16.35%
RGBNT [2]	85.43%	87.90%	52.63%	6.40%
HST [2]	88.34%	92.34%	54.74%	50.91%
MSERs_NRB [48]	92.34%	93.53%	62.30%	9.68%
HCRE [50]	97.65%	98.80%	95.21%	16.90%

relation of the average number of traffic sign proposals (AN) and recall. The traffic sign proposals are the extracted regions of MSERs.

The curves show MSERs_NRB achieves the highest recall of 91% with an AN value of 389. With a similar AN value of 407, MSER_NRGB_Enhancement has a recall of 92%. As MSER_NRGB_Enhancement achieves higher recall values, the AN becomes larger, which means that there are more backgrounds that are wrongly classified as ROI regions.

In Table 7, the color and shape based methods of MSERs_NRB [48] and HCRE [50] are compared with three color based methods including RGB enhancement (RGBE) [30], RGB normalized thresholding (RGBNT) [2] and hue and saturation thresholding [2]. The comparison results on red, blue and yellow colors are shown in Table 7. Recall is the ratio of the number of detected objects to the number of all objects. Extraction rate (ER) is the ratio of the pixel number of extracted regions to the pixel number of the input image. Compared with the color based methods, the MSERs based method [48] and the HCRE method [50] can achieve higher performance in ROI extraction. The color and shape based methods often rely on good color enhancement results and suitable parameters. Furthermore, the extracted candidates of these methods may be incomplete, which brings difficulty to the following classification methods; without directly using the candidates as inputs, scanning the regions with classification methods is a good way to overcome this problem.

VII. MACHINE LEARNING BASED METHODS
In recent years, with the development of machine learning methods, the machine learning based detection methods have gradually become the mainstream algorithms and achieved the-state-of-the-art results in some aspects.

A. REVIEW OF MACHINE LEARNING BASED DETECTION METHODS
The machine learning based TSD methods are reviewed according to their adopted machine learning methods including AdaBoost, SVM and NN. Machine learning based detection methods are summarized in Table 8.

TABLE 8. Machine learning based detection methods.

Category	Paper	Year	Features	Training method and detection structure
AdaBoost based methods	[6]	2012	Haar-like	Cascade
	[53]	2009	Dissociated dipoles	Parallel cascade
	[88]	2014	MN-LBP	Split-flow cascade tree
	[57]	2017	ACF	Cascade of ACF
	[58]	2016	ACF, LBP, Spatially Pooled LBP	Cascade
	[25]	2013	ICF (or ChnFtrs)	Cascade of ICF
	[11]	2015	ICF (or ChnFtrs)	Cascade of ICF
	[11]	2015	ACF	Cascade of ACF
	[59]	2016	Decision tree on image pyramid	Boosted decision trees
SVM based methods	[48]	2012	HOG	Linear SVM
	[89]	2013	Color HOG	IK-SVM, LDA
	[67]	2013	HOG	SVM with RBF kernel
	[68]	2013	HOG	Liblinear SVM
	[62]	2015	PHOG	SVM
	[63]	2016	HSI-HOG, LSS features	SVM with RBF kernel
	[52]	2017	Integral, compressed, color HOG	Linear SVM
	[64]	2014	Color global LOEMP	Linear SVM
	[65]	2013	Gabor	SVM with polynomial kernel
	[66]	2016	HOG, LBP and Gabor	SVM
CNN based methods	[69]	2013	SVM+CNN	
	[70]	2015	RGB thresholding+RCNN	
	[76]	2018	Convolutional neural network (CNN)	
	[71]	2016	fully convolutional network (FCN) and deep CNN	
	[72]	2018	AN (Attention Network) and Faster-RCNN	
	[10]	2016	Convolutional neural network (CNN)	
	[74]	2018	Cascaded segmentation detection networks	
	[73]	2016	Cascaded convolutional neural networks	
	[75]	2017	YOLOv2	

(Machine Learning Based Methods)

1) ADABOOST BASED METHODS
Viola and Jones' AdaBoost and cascade based detection structure (VJ) [3] has been proved very efficient in some object detection problems, such as face detection, car detection, license plate detection, etc. This structure has also been successfully applied in different TSD applications.

Combined with some types of rectangular features, an AdaBoost based learning method and a cascade structure, the VJ structure can select features with the AdaBoost method for object expression and then detect objects in a cascade process.

The selection of features is crucial for AdaBoost based TSD detectors. The Haar-like feature [3] is the most popular feature used in different detection problems. The Haar-like feature can express the gray level difference of traffic signs. Considering that Haar-like features have connected dipoles, Baró *et al.* [53] proposed the dissociated dipoles feature, which is a more general rectangular feature. Using unconnected two dipoles, the dissociated dipoles feature can produce more features to express traffic signs. Multi-Block Local

Binary Pattern (MB-LBP) feature [54] is another popular used rectangular features. Liu *et al.* [88] designed multi-block normalization LBP (MN-LBP) features to express different types of features. The designed MN-LBP feature can be trained to find the common features of different types of traffic signs.

Without of using one type of feature, the Integral Channel Features (ICF, sometimes also abbreviated as ChnFtrs) can extract features such as local sums, histograms and Haar-like features from multiple registered image channels. ICF was first presented for pedestrian detection [55] and was repurposed to achieve good TSD results in different traffic sign detection problems [11], [25].

Unlike traditional AdaBoost based structures that use gray level features, the Aggregate Channel Features (ACF) detection [56] is based on a cascade of boosted week tree classifiers which are trained using 10 channel features. The ACF based detection methods have been used in some detection problems, and have also been successfully applied in TSD applications [11], [57].

Hu *et al.* [58] utilized some features to fulfill the detection work; the features include ACF, LBP and Spatially Pooled LBP, etc. These different types features can generate a large amount of training features; yet, the training and detection processes are often more complex than using one type of feature.

The structures of the Haar-like features, dissociated dipoles, MN-LBP, ICF and ACF are shown in Fig. 9.

The common AdaBoost based training methods include Real AdaBoost, Gentle AdaBoost, Discrete AdaBoost and other derived Boosting methods. These AdaBoost training methods can select powerful features as weak classifiers, which can form a strong classifier for object detection.

The design of cascade structures also plays an important role in different TSD applications. The cascade structure [3] is the most popular structure for AdaBoost based detectors. This structure can reject background in a coarse-to-fine process saving processing time. Yet, the classical structure often can only handle traffic signs with similar appearances and structures. Baró *et al.* [53] designed a parallel cascade with some detectors working in parallel to detect different types of traffic signs. The detectors in this parallel cascade need to process an image several times, which is more time-consuming and has more false alarms than using one cascaded detector. Liu *et al.* [88] proposed a split-flow cascade tree (SFC-tree) structure to detect different types of traffic signs. Combined with MN-LBP features, the SFC-tree structure can detect traffic signs in a coarse-to-fine process. Compared with the parallel cascade, the SFC-tree structure just needs to scan the image once saving processing time.

Though AdaBoost based detection is very fast, scanning a high-resolution image is still time-consuming. Some methods utilized color extraction or other ROI extraction methods to give ROIs for the AdaBoost detection process [88]. In some applications, AdaBoost based detectors can also be utilized for coarse detection followed by some other detection methods, such as SVM or CNN [59], [25].

2) SVM BASED METHODS

The SVM and Histograms of Oriented Gradients (HOG) [4] based detection structure was first proposed to detect pedestrians and has been commonly used in different detection problems in the past decade. This structure utilizes HOG-like features to express the objects and treats the object detection problem as an SVM classification problem, in which each candidate is classified into objects or backgrounds. The SVM based detection structure has been successfully applied in TSD problems.

The introduction of HOG-like features is the key of the success of SVM based detection methods. The HOG feature [4] is the most popular feature used in different detection problems.

Using classical HOG features, the HOG+SVM based detection methods [31], [60] can achieve high detection results. Different features have been derived from HOG features.

The pyramid histogram of oriented gradients (PHOG) feature proposed in [61] has been used in some object detection problems including TSD. As a pyramid scaled version of HOG, PHOG can represent the global and local shape information, making it more effective for object detection [62].

In [63], the HOG features were extended to the HIS color space and then combined with the local self-similarity (LSS) features to get the descriptor for TSD. A derivative feature of HOG, called Color Global and Local Oriented Edge Magnitude Pattern (Color Global LOEMP). The LOEMP utilizes HOG to express objects and then uses LBP histogram codes of each orientation to get a texture vector for SVM classification [64].

Yuan *et al.* [52] used several different HOG variants, including HOG, color HOG, integral HOG and its compressed version. Using different HOG features can generate more vectors for SVM classification. If different HOG features can express objects well, the performance may be improved.

Without using HOG-like features, Park and Kim [65] utilized edge-adaptive Gabor filter and SVM classification for traffic sign detection. Berkaya *et al.* [66] proposed an ensemble of different features including LBP, HOG, and Gabor, and then utilized SVM for classification.

Different HOG features including HOG descriptor [4], PHOG [61], HIS-HOG [63] and Color Global LOEMP [64] are shown in Fig. 10.

An exhaustive scanning process is used in the SVM-based detection process, which is a time-consuming process for scanning a high-resolution image. Hence, most SVM based detection methods have a ROI extraction process, which can largely reduce the scanning regions saving detection time. For example, in [67] and [68], color and shape based ROI extraction methods were utilized to provide candidates for SVM classification. In [31], color enhancement and MSERs based
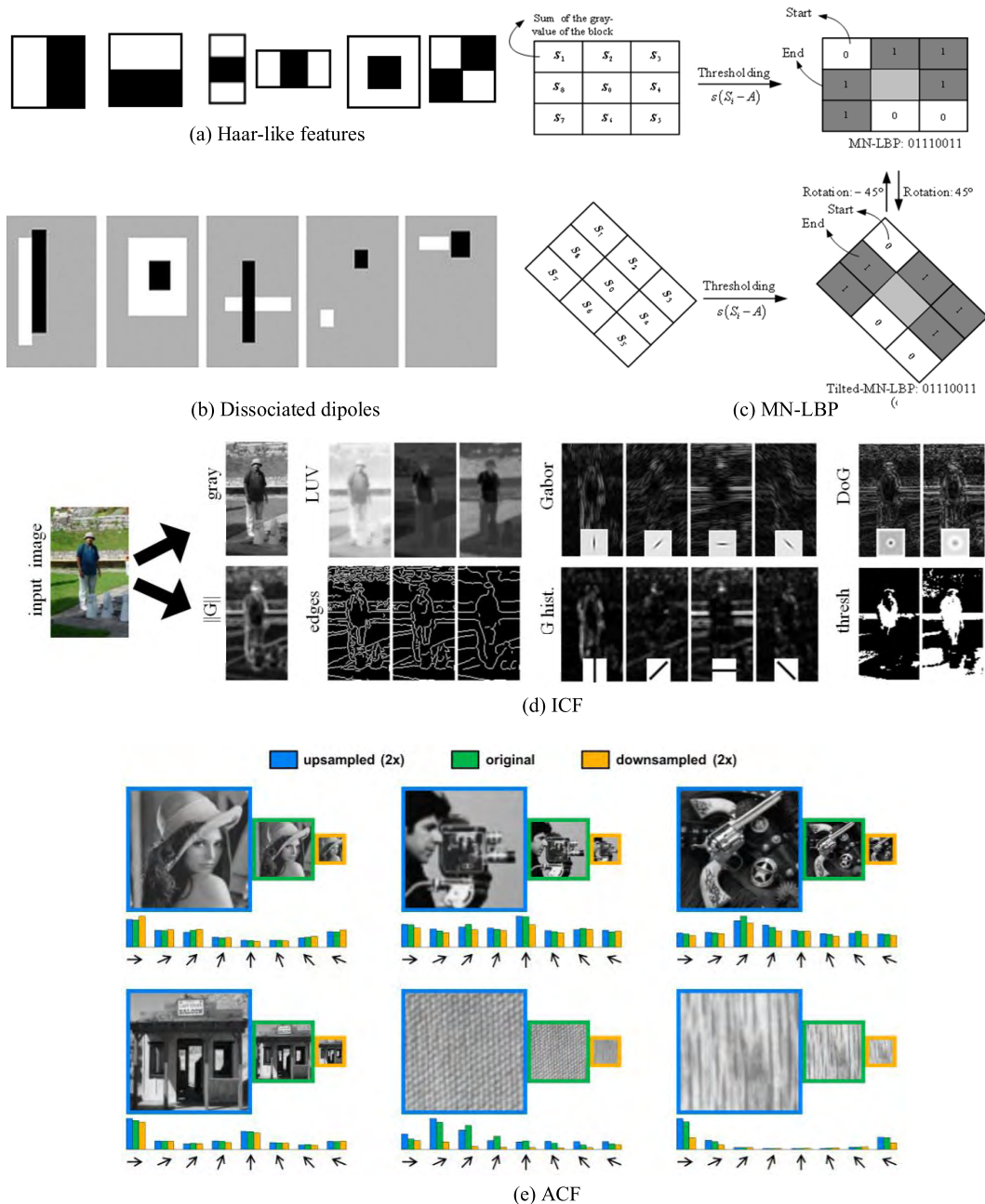
(a) Haar-like features

(b) Dissociated dipoles

(c) MN-LBP

(d) ICF

(e) ACF

FIGURE 9. Features for AdaBoost. (a) Haar-like features [3], (b) Dissociated dipoles [53], (c) MN-LBP [88], (d) ICF [55], (e) ACF [56]. The detailed description of these features can be found in their corresponding papers.

method was utilized to extract ROIs and applied SVM+HOG detector to classify the ROIs as objects or backgrounds. Many ROI extraction methods for SVM detectors have been described in the color or shape based detection parts in this review. Furthermore, methods based on SVM and HOG also play an important role in the classification of shapes, normal signs and occluded signs [93].

3) CNN BASED METHODS

Most of the AdaBoost or SVM based detection methods rely on handcrafted features to identify signs. Distinguished from

these methods, the Convolutional Neural network (CNN) based detection methods learn features through convolutional network. In recent years, with the development of deep learning, many different deep neural network structures have appeared and made breakthrough in different detection areas.

The use of CNN for the TSD problem started in [69] and [70]. These works use a CNN classifier to classify objects from backgrounds and need ROIs extraction methods to get candidates. Zang *et al.* [73] utilized an AdaBoost classifier to extract ROIs for the following CNN based detector. Zhu *et al.* [74] proposed a text-based detection method with

(a) HOG descriptor

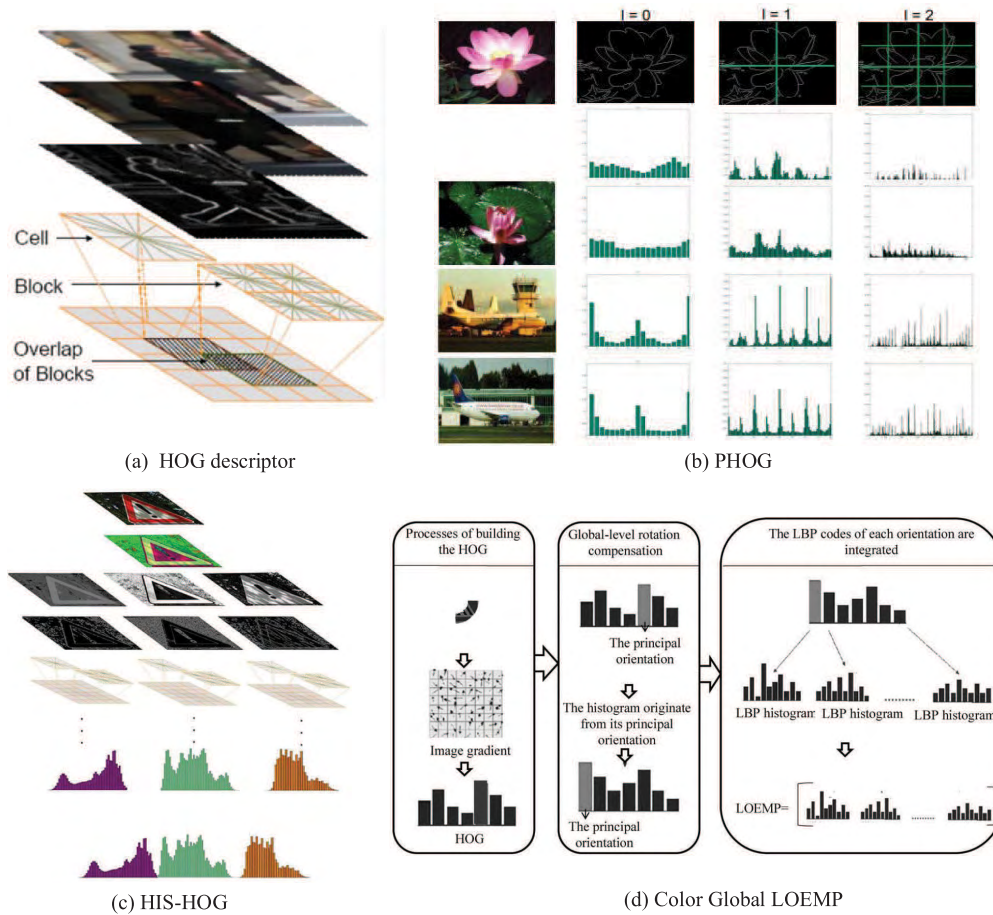(b) PHOG

(c) HIS-HOG

(d) Color Global LOEMP

FIGURE 10. Different HOG features used for TSD. (a) HOG descriptor [4], (b) PHOG [61], (c) HIS-HOG [63], (d) Color Global LOEMP [64]. The detailed description of these features can be found in their corresponding papers.

two NN components including a ROI extraction network and a fast detection network. The accuracy and efficiency of these methods are affected by the accuracy of the designed ROIs extraction methods.

Instead of using some ROI methods, some CNN based methods have their own ROI extraction net. Zhu *et al.* [71] proposed a method based on a fully convolutional network (FCN) and a deep CNN for classification. The FCN is used for detecting traffic sign proposals and the CNN is used to classify traffic sign proposals.

Yang *et al.* [72] proposed an end-to-end deep network that extracts region proposals by a two-stage adjusting strategy. The first stage is an Attention Network (AN), designed to find potential ROIs and roughly classifying them into three categories. The second stage is a Fine Region Proposal Network (FRPN) that generates the final region proposals.

Without ROI extraction methods or nets, some networks use one net to fulfill the detection task. Zhu *et al.* [10] proposed a robust end-to-end CNN that can simultaneously detect and classify traffic signs. Most CNN based detection networks are slow to detect signs. There are some networks that have fast performance such as You only look

once (YOLO) net. Zhang *et al.* [75] utilized YOLOv2 to design their real-time traffic sign detection method.

Though most CNN detection methods provide accurate bounding boxes, processes dealing with bounding boxes may follow to obtain more precise bounding boxes. Lee and Kang [76] built their boundary estimation CNN based on the Single Shot MultiBox structure which can predict bounding boxes across multiple feature levels. Besides TSD, many deep learning methods are designed for traffic sign classification [92].

B. ANALYSIS OF THE MACHINE LEARNING BASED METHODS

After review of the AdaBoost, SVM and CNN based methods, a brief analysis of these methods including advantages and shortcomings is presented in this subsection. Comparison results on some popular public datasets are also listed.

Comparison results on the public GTSDB, BTSD, TT100k and LISA datasets are listed in Table 9. AUC (Area Under Curve), AP (Average Precision), recall and accuracy are used for evaluation. In this table, ''Small'', ''Med'' and ''Large'' mean the test sets with small, medium and large size signs.

TABLE 9. Performance of different machine learning based methods on different datasets.

Dataset	Methods	Prohibitive (AUC)	Danger (AUC)	Mandatory (AUC)	Time (s)
GTSDB	HOG+LDA [6]	70.33%	35.94%	12.01%	N/A
	Hough-like [6]	26.09%	30.41%	12.86%	N/A
	Viola-Jones [6]	90.81%	46.26%	44.87%	N/A
	HOG+LDA+SVM [89]	100%	99.91%	100%	3.533
	ChnFtrs [25]	100%	100%	96.98%	N/A
	HOG+SVM [67]	99.98%	98.72%	95.76%	3.032
	SVM+Shape [68]	100%	98.85%	92.00%	0.4-1
	SVM+CNN [69]	N/A	99.78%	97.62%	12-32
	SFC-tree [88]	100%	99.20%	98.57%	0.192 (3.19 GHz CPU)
	CNN [E-53]	99.89%	99.93%	99.16%	0.162 (Titan X GPU)
	ACF+SPC+LBP+AdaBoost [58]	100%	98.00%	97.57%	N/A
	AdaBoost+SVR [59]	100%	100%	99.87%	N/A
	AdaBoost+CNN+SVM [73]	99.45%	98.33%	96.50%	N/A
BTSD	ChnFtrs [25]	94.44%	97.40%	97.96%	1~3 (Intel Core i7 870 CPU, GTX 470 GPU)
	AdaBoost+SVR [59]	93.45%	99.88%	97.78%	0.05~0.5 (Intel Core-i7 4770 CPU)
	AN+FRPN [72]	AP(%): 50.82%(Small), 88.05%(med), 96.82%(large)			0.128 (Tesla K20 GPU)
	Faster-RCNN in [72]	AP(%): 43.93%(Small), 97.8%(medium), 98.31%(large)			0.165 (Tesla K20 GPU)
TT100k	Fast R-CNN in [10]	Recall: 56%; Accuracy: 50% Curves can be found in [10]			N/A
	Multi-class Network [10]	Recall: 91%; Accuracy: 88% Curves can be found in [10]			N/A
	AN+FRPN [72]	AP(%): 49.81%(Small), 86.9%(med), 96.05%(large)			0.128 (Tesla K20 GPU)
	Faster-RCNN in [72]	AP(%): 31.22%(Small), 77.17%(med), 94.05%(large)			0.165 (Tesla K20 GPU)
LISA	ICF in [11]	87.32% (Diamond)	96.03% (Stop)	91.09% (NoTurn)	N/A
	ACF in [11]	98.98% (Diamond)	96.11% (Stop)	96.17% (NoTurn)	N/A

GTSDB is the most commonly used dataset. For GTSDB, it is convenient to get a large amount of training samples from GTSDB and GTSRB to train a detector. The problem is that there is little room to improve the performance on GTSDB. Some AdaBoost based, SVM based or CNN based methods can achieve nearly 100% AUC values on prohibitive, danger or mandatory sign detection. According to the published results in Table 9, the HOG+LDA+SVM detector [89] and the AdaBoost+SVR [59] method achieved the highest AUCs.

Three papers published their results on BTSD and two papers published their results on TT100k. There is still much room to improve the performance on BTSD and TT100k. For BTSD, the AdaBoost+SVR [59] method achieved the highest AUCs, while the Faster-RCNN [72] achieved the best APs in detecting medium and large signs. For TT100k, the Multi-class Network [10] achieved a highest recall of 91% and a highest accuracy of 88%; and the AN+FRPN [72] achieved the best APs on detection of small, medium and large signs. The American traffic signs from LISA seem different from the signs from GTSDB and TT100k. The performance of ICF and ACF was tested on LISA [11]. ACF has a better performance on diamond, stop and no-turn signs.

Compared with the SVM and CNN based methods, the AdaBoost based methods are often much faster and do not need a ROI extraction process; yet, some AdaBoost based methods often have weaker generalization when handling samples with large differences in shape and appearance.

There are some existing methods, such as ICF, ACF and SCF-tree, that can partly overcome this shortcoming.

The SVM based methods achieved the-state-of-the-art results in some aspects during the competition of GTSDB in 2013. The processing speed of SVM methods is often faster than CNN based methods and slower than AdaBoost based methods. The SVM based methods often need a ROI extraction process to get ROI regions for scanning or candidate regions for classifying, which has a great effect on the performance of the SVM based TSD detectors.

With the fast development of deep learning methodologies, the deep CNN based methods have achieved significant improvements in different detection problems. Compared with SVM and AdaBoost based methods, the CNN based methods do not need manually designed features and can handle a larger amount of training samples to generate a detector with better generalization ability. Yet, only GTSDB and TT100K have enough training samples, while BTSD and LISA have limited training and test samples. From the test results in TT100k in Table 9, it can be seen that the deep learning methods including Fast RCNN [10], Faster-RCNN [72] and AN+FRPN [72] did not achieve a promising performance. Especially in detecting small size signs, the methods of Faster-RCNN [72] and AN+FRPN [72] achieved low APs of 31.22% and 49.81% respectively. For deep learning based TSD methods, there is still much room for improvement in the field of detecting license plates, especially small-size license plates.

Note that, though we use machine learning methods to name the methods listed in Table 9, a part of these methods have some assisted methods, such as color based methods and color and shape based methods.

VIII. LIDAR BASED METHODS

Mobile laser scanning technology of the LIDAR has experienced significant growth in recent years, and has been a key solution in many Advanced Driver Assistance Systems (ADAS) and Auto Driving Systems (ADS). With a mobile LIDAR system, 3D urban objects can be detected and classified for different purposes. A review of the methods for detection, segmentation and classification of 3D urban objects can be found in [77]. Recently, many road management systems have utilized laser scanning to provide point cloud information for infrastructure inventory analysis and routine inspections [78].
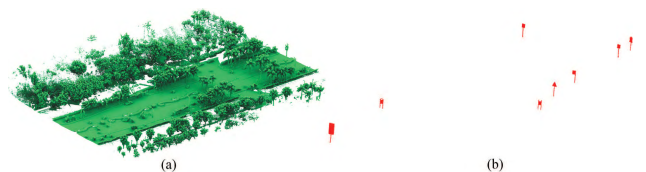

FIGURE 11. LIDAR based traffic sign detection from [79]. (a) shows the point cloud. (b) is the detection results of traffic signs.

A schematic of the LIDAR based traffic sign detection from [79] is shown in Fig. 11. In Fig. 11, the point cloud

captured by LIDAR is processed to detect traffic signs based on the retro-reflective properties; then, the detected signs in point cloud are associated with their corresponding positions in RGB images; lastly, a classification process is performed to classify the types of traffic signs from the filtered set of RGB images.

In this work we propose the next methodology: initially our vehicle equipped with LIDAR and RGB cameras gathers information (3D point cloud and 2D imagery). Then, the point cloud is processed to automatically detect traffic signs based on their retro- reflective properties. Furthermore,

A. REVIEW OF LIDAR BASED DETECTION METHODS

In this subsection, we classify the LIDAR based detection methods into two categories, including data cloud based detection methods, and data cloud and RGB image based detection methods. The methods are listed in Table 10 and reviewed as follows.

TABLE 10. LIDAR based methods.

	Paper	Year	Equipment	Process	Methods	Purpose
Laser Based TSR Methods	[77]	2016	LIDAR	Clustering of the point cloud, shape recognition	DBSCAN clustering, shape classification	Traffic sign detection and shape recognition
	[82]	2017	LIDAR	LIDAR based occlusion detection	Correctness and completeness based detection	Traffic sign occlusion detection
	[16]	2016	LIDAR and camera	LIDAR based detection and camera based recognition	WSMLR	Multiview detection and recognition
	[83]	2018	LIDAR and camera	LIDAR based detection and camera based recognition	SVM and CSA+LR detection, SVM+HOG classification	Traffic sign detection and recognition
	[85]	2018	LIDAR and camera	LIDAR based detection and camera based recognition	Voxel-based detection, deep learning classification	Traffic sign detection and recognition
	[79]	2016	LIDAR and camera	LIDAR based detection and camera based recognition	Bag-of-visual-phrases detection, Boltzmann classification	Traffic sign detection and recognition
	[90]	2016	LIDAR and camera	LIDAR based detection and camera based recognition	DBSCAN clustering, SVM+HOG classification	Traffic sign detection and recognition
	[84]	2017	LIDAR and camera	LIDAR based detection and camera based recognition	Gaussian Mixture Models, DNN	Traffic sign detection and recognition

1) DATA CLOUD BASED DETECTION

Pu *et al.* presented a laser scanning based detection method for traffic signs, trees, building walls and barriers [80]. In this study, poles were recognized for up to 86%; this method need to integrate with images to classify the poles into further categories.

Considering the radiometric and geometric information generated by laser scanning, Riveiro *et al.* [81] presented a method for detection of retro-reflective traffic signs. This method first creates an intensity map of point cloud and uses thresholding method to select pixels of highly reflective surfaces. Then a fine intensity thresholding process is used to get a filtered point cloud. Clustering method based on Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is designed to get a clustered point cloud for feature recognition.

For maintenance of traffic signs, Huang *et al.* [82] presented an occluded traffic sign detection method based on point cloud data and trajectory data acquired by a LIDAR system. Traffic signs with various occlusions can be detected and analyzed.

2) DATA CLOUD AND RGB IMAGE BASED DETECTION

For ITS, camera and LIDAR are two most popular sensing devices. Camera can capture planar information like color, shape, and texture. RGB images captured with cameras are quite sensitive to illumination changes and viewpoint. As an active vision, LIDAR can provide 3D geometric information, such as point clouds; yet, LIDAR cannot capture visual planar information. Data association of LIDAR and camera gives a promising direction for detection and recognition of traffic signs [83]. The structure of LIDAR and camera based system from [84] is shown in Fig. 12. In this structure,
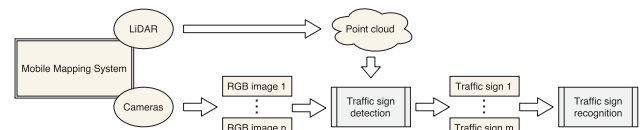

FIGURE 12. Structure of LIDAR and camera based system from [84].

For the data association of a laser point and its corresponding image pixel, parameters of both camera and LIDAR must be precomputed as a prerequisite [83]. Point clouds can provide geometric and localization information; whereas, digital images can provide detailed color, shape and texture information. By fusing images and point clouds, a promising framework is to use LIDAR point clouds for traffic sign detection and digital images for classification. Using this technique, a mobile mapping system based on laser scanners and digital cameras was designed in [85].

In [16], the 3D LIDAR points were assisted to detect 2D multiview signs from the images. Then, the multi-view sign recognition method was developed based on a metric-learning-based template matching approach.

Yu *et al.* [79] presented a structure for detection and recognition of traffic signs based on both LIDAR and camera. This method uses bag-of-visual-phrases representations for traffic sign detection based on 3D point clouds. For the recognition task, a deep Boltzmann machine-based hierarchical classifier was designed based on 2D images.

Arcos-García *et al.* [84] presented an efficient two-stage TSR system. The retro-reflective material of traffic signs was used to design the detection process based on 3D point clouds. Then, a deep neural network is designed for classification on RGB images projected from point cloud data.

B. ANALYSIS OF THE LIDAR BASED TSD METHODS

Because of the large improvement of LIDAR technology, there are some LIDAR based methods appeared in the recent years. The process of laser scanning generate 3D point cloud data. Some methods used point cloud data to detect poles including traffic signs based on the pole structure of traffic signs. Some methods use geometric and radiometric information to detect retro-reflective traffic signs. LIDAR can provide 3D point clouds, while cameras can provide 2D RGB images.

Data association of LIDAR and camera is a promising way for detection and recognition of traffic signs.

Because the specificity of 3D point cloud, most of the traditional deep learning methods are hard to be directly applied in the detection problem with 3D point cloud. Different processes of point cloud preprocessing, thresholding, filtering, clustering are commonly used in LIDAR based TSD methods. Unlike deep learning methods, these processes usually still rely on manually designed models and thresholds, which may result in low generalization ability.

Though there are many studies for LIDAR based TSD, different previous studies used their own LIDAR systems. There is no public dataset for testing LIDAR based TSD methods. Hence, it is hard to compare and do analysis the performance of different studies.

IX. CONCLUSIONS AND PERSPECTIVES

In this review, we divide the traffic sign detection methods into five categories: color based methods, shape based methods, color and shape based methods, machine learning based methods, and LIDAR based methods. Conclusions and perspectives are given in this section.

The color based methods are often fast and relatively simple. Though most of the previous color based detection methods have been out of date, they are still important ways to extract ROIs for the following fine detection process. Building robust color enhancement methods or color extraction methods for other detection methods is an assisted way to achieve fast detection in real applications.

The shape based methods have not been widely studied in recent years. Relying on edge detection, most shape based methods are often not suitable for detecting traffic signs with small size or vague edges, yet have potential on traffic sign extraction in some applications.

The color and shape based methods such as MSERs based methods and HCRE based methods can achieve high performance for ROI extraction; these methods usually need a good color enhancement process. In future, robust color enhancement and extraction methodologies may be developed to further improve the performance of these methods.

The machine learning methods have achieved the-state-of-the-art results. When dealing with high resolution images and small vague traffic signs, some machine learning methods are still hard to keep a good balance of the consuming time and accuracy. A large portion of these methods need some assisted methods to achieve fast and accurate detection.

Mobile laser scanning technology has experienced significant growth in recent five years, and has been a key solution in many ADAS systems. There are many methods published using different laser scanning devices and their own dataset. It is hard to compare the performance of these methods.

The previous TSD methods tested on some public datasets for traffic sign detection have reported high performance. For example, the methods tested on GTSDB have achieved nearly 100% AUC. There is little room to improve the performance of different methods on GTSDB. Released in 2016,

the TT100K dataset is a promising dataset for future comparison. Because signs from different countries usually have different appearances and structures, some new public datasets are needed for evaluating TSD methods designed for traffic signs from different countries.

The previous TSD methods and public datasets mainly involved the challenging problems of small sizes, occlusions, complex driving scenes, rotation in or out the plane, illumination changes, etc. These variations belong to classical TSD problems and have been researched for many years. Rare methods focused on the traffic sign detection problem at night which has some difficulties to deal, such as headlight reflection, street lighting and dark illumination. Extreme weather has a great impact on the quality of the images captured by cameras. Extreme weather conditions such as heavy fog, heavy rain and heavy snow were also not considered in previous methods. In future, new methods and new datasets that can handle night and extreme weather conditions are needed to improve the ability of camera based TSD methods to deal with these conditions. The LIDAR based methods have large potential to handle these conditions. Yet, a small part of the researchers have their own on-board LIDAR to collect data and do research. Some public LIDAR datasets for TSD may be released in future.

REFERENCES

[1] A. Mogelmose, M. M. Trivedi, and T. B. Moeslund, "Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 4, pp. 1484–1497, Dec. 2012.

[2] H. Gómez-Moreno, S. Maldonado-Bascón, P. Gil-Jiménez, and S. Lafuente-Arroyo, "Goal evaluation of segmentation algorithms for traffic sign recognition," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 4, pp. 917–930, Dec. 2010.

[3] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, 2004.

[4] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. CVPR*, San Diego, CA, USA, Jun. 2005, pp. 886–893.

[5] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. CVPR*, Columbus, OH, USA, Jun. 2014, pp. 580–587.

[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," in *Proc. Int. Joint Conf. Neural Netw.*, Dallas, TX, USA, Aug. 2013, pp. 1–8.

[7] M.-Y. Fu and Y.-S. Huang, "A survey of traffic sign recognition," in *Proc. ICWAPR*, Qingdao, China, Jul. 2010, pp. 119–124.

[8] A. Gudigar, S. Chokkadi, and U. Raghavendra, "A review on automatic detection and recognition of traffic sign," *Multi. Tools Appli.*, vol. 75, no. 1, pp. 333–364, Jan. 2016.

[9] Y. Saadna and A. Behloul, "An overview of traffic sign detection and classification methods," *Int. J. Multimedia Inf. Retr.*, vol. 6, no. 3, pp. 193–210, Sep. 2017.

[10] Z. Zhu, D. Liang, S. Zhang, X. Huang, B. Li, and S. Hu, "Traffic-sign detection and classification in the wild," in *Proc. CVPR*, Las Vegas, NV, USA, Jun. 2016, pp. 2110–2118.

[11] A. Mogelmose, D. Liu, and M. M. Trivedi, "Detection of U.S. traffic signs," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 6, pp. 3116–3125, Sep. 2015.

[12] M. Costa, A. Simone, V. Vignali, C. Lantieri, and N. Palena, "Fixation distance and fixation duration to vertical road signs," *Appl. Ergonomics*, vol. 69, pp. 48–57, May 2018.

[13] T. Ben-Bassat and D. Shinar, "The effect of context and drivers' age on highway traffic signs comprehension," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 33, pp. 117–127, Aug. 2015.

[14] M. Meuter, C. Nunn, S. M. Gormer, S. Müller-Schneiders, and A. Kummert, "A decision fusion and reasoning module for a traffic sign recognition system," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 4, pp. 1126–1134, Dec. 2011.

[15] M. Boumediene, J.-P. Lauffenburger, J. Daniel, C. Cudel, and A. Ouamri, "Multi-ROI association and tracking with belief functions: Application to traffic sign recognition," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 6, pp. 2470–2479, Dec. 2014.

[16] M. Tan, B. Wang, Z. Wu, J. Wang, and G. Pan, "Weakly supervised metric learning for traffic sign recognition in a LIDAR-equipped vehicle," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 5, pp. 1415–1427, May 2016.

[17] Á. Gonzalez, M. Á. Garrido, D. F. Llorca, M. Gavilan, J. P. Fernandez, P. F. Alcantarilla, I. Parra, F. Herranz, L. M. Bergasa, M. Á. Sotelo, and P. R. de Toro, "Automatic traffic signs and panels inspection system using computer vision," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 485–499, Jun. 2011.

[18] S. Šegvić, K. Brkić, Z. Kalafatić, V. Stanisavljević, M. Ševrović, D. Budimir, and I. Dadić, "A computer vision assisted geoinformation inventory for traffic infrastructure," in *Proc. Conf. ITS*, Funchal, Portugal, Sep. 2010, pp. 66–73.

[19] C. Wen, J. Li, H. Luo, Y. Yu, Z. Cai, H. Wang, and C. Wang, "Spatial-related traffic sign inspection for inventory purposes using mobile laser scanning data," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 1, pp. 27–37, Jan. 2015.

[20] P. Siegmann, R. Lopez-Sastre, P. Gil-Jimenez, S. Lafuente-Arroyo, and S. Maldonado-Bascón, "Fundaments in luminance and retroreflectivity measurements of vertical traffic signs using a color digital camera," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 3, pp. 607–615, Mar. 2008.

[21] W. Y. Yan, A. Shaker, and S. Easa, "Potential accuracy of traffic signs' positions extracted from Google street view," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 2, pp. 1011–1016, Jun. 2013.

[22] K. Guan, B. Ai, M. Nicolás, R. Geise, A. Möller, Z. Zhong, and T. Kürner, "On the influence of scattering from traffic signs in vehicle-to-X communications," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 5835–5849, Aug. 2016.

[23] M. Munoz-Organero and V. C. Magana, "Validating the impact on reducing fuel consumption by using an EcoDriving assistant based on traffic sign detection and optimal deceleration patterns," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 2, pp. 1023–1028, Jun. 2013.

[24] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition," *Neural Netw.*, vol. 32, pp. 323–332, Aug. 2012.

[25] R. Timofte, M. Mathias, R. Benenson, and L. Van Gool, "Traffic sign recognition—How far are we from the solution?" in *Proc. IJCNN*, Dallas, TX, USA, Aug. 2013, pp. 1–8.

[26] F. Larsson and M. Felsberg, "Using Fourier descriptors and spatial models for traffic sign recognition," in *Image Analysis* (Lecture Notes in Computer Science), vol. 6688. Berlin, Germany: Springer, May 2011, pp. 238–249.

[27] C. Grigorescu and N. Petkov, "Distance sets for shape filters and shape recognition," *IEEE Trans. Image Process.*, vol. 12, no. 10, pp. 1274–1286, Oct. 2003.

[28] R. Belaroussi, P. Foucher, J.-P. Tarel, B. Soheilian, P. Charbonnier, and N. Paparoditis, "Road sign detection in images: A case study," in *Proc. Conf. ICPR*, Istanbul, Turkey, Aug. 2010, pp. 484–488.

[29] H. Fleyeh, "Traffic and Road Sign Recognition," Ph.D. dissertation, Dept. Comput. Eng., Edinburgh Napier Univ., Edinburgh, Scotland, 2008.

[30] A. Ruta, Y. Li, and X. Liu, "Real-time traffic sign recognition from video by class-specific discriminative features," *Pattern Recognit.*, vol. 43, no. 1, pp. 416–430, 2010.

[31] S. Salti, A. Petrelli, F. Tombari, N. Fioraio, and L. D. Stefano, "Traffic sign detection via interest region extraction," *Pattern Recognit.*, vol. 48, no. 4, pp. 1039–1049, Apr. 2015.

[32] S. Lafuente-Arroyo, P. García-Díaz, F. Acevedo-Rodríguez, P. Gil-Jiménez, and S. Maldonado-Bascón, "Traffic Sign Classification invariant to rotations using support vector machines," in *Proc. Adv. Concepts Intell. Vis. Syst.*, Brussels, Belgium, 2004, pp. 1–6.

[33] A. de la Escalera, J. M. Armingol, J. M. Pastor, and F. J. Rodríguez, "Visual sign information extraction and identification by deformable models for intelligent vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 5, no. 2, pp. 57–68, Jun. 2004.

[34] J. M. Lillo-Castellano, I. Mora-Jiménez, C. Figuera-Pozuelo, J. L. Rojo-Álvarez, "Traffic sign segmentation and classification using statistical learning methods," *Neurocomputing*, vol. 153, pp. 286–299, Apr. 2015.

[35] H. Gómez-Moreno, P. Gil-Jiménez, S. Lafuente-Arroyo, R. Vicen-Bueno, and R. Sánchez-Montero, "Color images segmentation using the support vector machines," *Bioinformatics*, vol. 1, no. 3, pp. 1–28, 2000.

[36] K. Zhang, Y. Sheng, and J. Li, "Automatic detection of road traffic signs from natural scene images based on pixel vector and central projected shape feature," *IET Intell. Transp. Syst.*, vol. 6, no. 3, pp. 282–291, Sep. 2012.

[37] P. Yakimov, "Traffic signs detection using tracking with prediction," in *Proc. Conf. E-Bus. Telecommun.*, Colmar, France, 2015, pp. 454–467.

[38] P. Yakimov and V. Fursov, "Traffic signs detection and tracking using modified Hough transform," in *Proc. Conf. E-Bus. Telecommun.*, Colmar, France, Jul. 2015, pp. 1–7.

[39] N. Barnes, A. Zelinsky, and L. S. Fletcher, "Real-time speed sign detection using the radial symmetry detector," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 2, pp. 322–332, Jul. 2008.

[40] N. Barnes and G. Loy, "Real-time regular polygonal sign detection," in *Proc. Conf. Field Service Robot.*, 2006, pp. 55–66.

[41] C.-Y. Fang, S.-W. Chen, and C.-S. Fuh, "Road-sign detection and tracking," *IEEE Trans. Veh. Technol.*, vol. 52, no. 5, pp. 1329–1341, Sep. 2003.

[42] S. Xu, "Robust traffic sign shape recognition using geometric matching," *IET Intell. Transp. Syst.*, vol. 3, no. 1, pp. 10–18, Mar. 2009.

[43] P. Gil-Jiménez, S. Maldonado-Bascón, H. Gómez-Moreno, S. Lafuente-Arroyo, and F. López-Ferreras, "Traffic sign shape classification and localization based on the normalized FFT of the signature of blobs and 2D homographies," *Signal Process.*, vol. 88, no. 12, pp. 2943–2955, 2008.

[44] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. CVPR*, Kerkyra, Greece, Sep. 1999, pp. 1150–1157.

[45] F. Larsson, M. Felsberg, and P. E. Forssen, "Correlating Fourier descriptors of local patches for road sign recognition," *IET Comput. Vis.*, vol. 5, no. 4, pp. 244–254, Jul. 2011.

[46] J. Khan, S. Bhuiyan, and R. Adhami, "Hierarchical clustering of EMD based interest points for road sign detection," *Opt. Laser Technol.*, vol. 57, no. 4, pp. 271–283, Apr. 2014.

[47] J. Greenhalgh and M. Mirmehdi, "Recognizing text-based traffic signs," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 3, pp. 1360–1369, Jun. 2015.

[48] J. Greenhalgh and M. Mirmehdi, "Real-time detection and recognition of road traffic signs," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 4, pp. 1498–1506, Dec. 2012.

[49] Y. Yang, H. Luo, H. Xu, and F. Wu, "Towards real-time traffic sign detection and classification," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 7, pp. 2022–2031, Jul. 2016.

[50] C. Liu, F. Chang, Z. Chen, and D. Liu, "Fast traffic sign recognition via high-contrast region extraction and extended sparse representation," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 1, pp. 79–92, Jan. 2016.

[51] D. Wang, X. Hou, J. Xu, S. Yue, and C.-L. Liu, "Traffic sign detection using a cascade method with fast feature extraction and saliency test," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 12, pp. 3290–3302, Dec. 2017.

[52] X. Yuan, J. Guo, X. Hao, and H. Chen, "Traffic sign detection via graph-based ranking and segmentation algorithms," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 45, no. 12, pp. 1509–1521, Dec. 2015.

[53] X. Baró, S. Escalera, J. Vitri, O. Pujol, and P. Radeva, "Traffic sign recognition using evolutionary Adaboost detection and forest-ECOC classification," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 1, pp. 113–126, Mar. 2009.

[54] L. Zhang, R. Chu, S. Xiang, S. Liao, and S. Z. Li, "Face detection based on multi-block LBP representation," in *Proc. Int. Conf. Biometrics*, Seoul, South Korea, 2007, pp. 11–18.

[55] P. Dollár, Z. Tu, P. Perona, and S. Belongie, "Integral channel features," in *Proc. BMVC*, London, U.K., 2009, pp. 1–11.

[56] P. Dollár, R. Appel, S. Belongie, and P. Perona, "Fast feature pyramids for object detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 8, pp. 1532–1545, Aug. 2014.

[57] Y. Yuan, Z. Xiong, and Q. Wang, "An incremental framework for video-based traffic sign detection, tracking, and recognition," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 7, pp. 1918–1929, Jul. 2017.

[58] Q. Hu, S. Paisitkriangkrai, C. Shen, A. van den Hengel, and F. Porikli, "Fast detection of multiple objects in traffic scenes with a common detection framework," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 4, pp. 1002–1014, Apr. 2017.

[59] T. Chen and S. Lu, "Accurate and efficient traffic sign detection using discriminative AdaBoost and support vector regression," *IEEE Trans. Veh. Technol.*, vol. 65, no. 6, pp. 4006–4015, Jun. 2016.

[60] F. Zaklouta and B. Stanciulescu, "Real-time traffic sign recognition in three stages," *Robot. Auto. Syst.*, vol. 62, no. 1, pp. 16–24, 2014.

[61] A. Bosch, A. Zisserman, and X. Munoz, "Representing shape with a spatial pyramid kernel," in *Proc. CIVR*, Amsterdam, The Netherlands, Jul. 2007, pp. 401–408.

[62] H. Li, F. Sun, L. Wang, and L. Liu, "A novel traffic sign detection method via color segmentation and robust shape matching," *Neurocomputing*, vol. 169, no. 2, pp. 77–88, Dec. 2015.

[63] A. Ellahyani, M. El Ansari, and I. El Jaafari, "Traffic sign detection and recognition based on random forests," *Appl. Soft Comput.*, vol. 46, pp. 805–815, Sep. 2016.

[64] X. Yuan, X. Hao, H. Chen, and X. Wei, "Robust traffic sign recognition based on color global and local oriented edge magnitude patterns," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 4, pp. 1466–1477, Apr. 2014.

[65] J.-G. Park and K.-J. Kim, "Design of a visual perception model with edge-adaptive Gabor filter and support vector machine for traffic sign detection," *Expert Syst. Appl.*, vol. 40, no. 9, pp. 3679–3687, Jul. 2013.

[66] S. K. Berkaya, H. Gunduz, O. Ozsen, C. Akinlar, and S. Gunal, "On circular traffic sign detection and recognition," *Expert Syst. Appl.*, vol. 48, pp. 67–75, Apr. 2016.

[67] S. Salti, A. Petrelli, F. Tombari, N. Fioraio, and L. Di Stefano, "A traffic sign detection pipeline based on interest region extraction," in *Proc. Int. Joint Conf. Neural Netw.*, Dallas, TX, USA, Aug. 2013, pp. 723–729.

[68] M. Liang, M. Yuan, X. Hu, J. Li, and H. Liu, "Traffic sign detection by ROI extraction and histogram features-based recognition," in *Proc. Int. Joint Conf. Neural Netw.*, Dallas, TX, USA, Aug. 2013, pp. 739–746.

[69] Y. Wu, Y. Liu, J. Li, H. Liu, and X. Hu, "Traffic sign detection based on convolutional neural networks," in *Proc. Int. Joint Conf. Neural Netw.*, Dallas, TX, USA, Aug. 2013, pp. 1–7.

[70] R. Qian, B. Zhang, Y. Yue, Z. Wang, and F. Coenen, "Robust Chinese traffic sign detection and recognition with deep convolutional neural network," in *Proc. Int. Conf. Natural Comput.*, Zhangjiajie, China, Aug. 2015, pp. 791–796.

[71] Y. Zhu, C. Zhang, D. Zhou, X. Wang, X. Bai, and W. Liu, "Traffic sign detection and recognition using fully convolutional network guided proposals," *Neurocomputing*, vol. 214, pp. 758–766, Nov. 2016.

[72] T. Yang, X. Long, A. K. Sangaiah, Z. Zheng, and C. Tong, "Deep detection network for real-life traffic sign in vehicular networks," *Comput. Netw.*, vol. 136, no. 8, pp. 95–104, May 2018.

[73] D. Zang, J. Zhang, M. Bao, J. Cheng, K. Tang, and D. Zhang, "Traffic sign detection based on cascaded convolutional neural networks," in *Proc. IEEE/ACIS Int. Conf. Softw. Eng., Artif. Intell., Netw. Parallel/Distrib. Comput.*, Shanghai, China, May/Jun. 2016, pp. 201–206.

[74] Y. Zhu, M. Liao, W. Liu, and M. Yang, "Cascaded segmentation-detection networks for text-based traffic sign detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 1, pp. 209–219, Jan. 2018.

[75] J. Zhang, M. Huang, X. Li, and X. Jin, " A real-time Chinese traffic sign detection algorithm based on modified YOLOv2," *Algorithms*, vol. 10, no. 4, p. 127, Nov. 2017.

[76] H. S. Lee and K. Kim, "Simultaneous traffic sign detection and boundary estimation using convolutional neural network," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 5, pp. 1652–1663, May 2018.

[77] A. Serna and B. Marcotegui, "Detection, segmentation and classification of 3D urban objects using mathematical morphology and supervised learning," *ISPRS J. Photogram. Remote Sens.*, vol. 93, pp. 243–255, Jul. 2014.

[78] M. Varela-González, H. González-Jorge, B. Riveiro, and P. Arias, "Performance testing of 3D point cloud software," *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. II-5/W2, pp. 307–312, Oct. 2013.

[79] Y. Yu, J. Li, C. Wen, H. Guan, H. Luo, and C. Wang, "Bag-of-visual-phrases and hierarchical deep models for traffic sign detection and recognition in mobile laser scanning data," *ISPRS J. Photogram. Remote Sens.*, vol. 113, pp. 106–123, Mar. 2016.

[80] D. Pei, F. Sun, and H. Liu, "Supervised low-rank matrix recovery for traffic sign recognition in image sequences," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 241–244, Mar. 2013.

[81] B. Riveiro, L. Díaz-Vilariño, B. Conde-Carnero, M. Soilán, and P. Arias, "Automatic segmentation and shape-based classification of retro-reflective traffic signs from mobile LiDAR data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Ses.*, vol. 9, no. 1, pp. 295–303, Jan. 2017.

[82] P. Huang, M. Cheng, Y. Chen, H. Luo, C. Wang, and J. Li, "Traffic sign occlusion detection using mobile laser scanning point clouds," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 9, pp. 2364–2376, Sep. 2017.

[83] Z. Deng and L. Zhou, "Detection and recognition of traffic planar objects using colorized laser scan and perspective distortion rectification," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 5, pp. 1485–1495, May 2018.

[84] Á. Arcos-García, M. Soilán, J. A. Álvarez-García, and B. Riveiro, "Exploiting synergies of mobile mapping sensors and deep learning for traffic sign recognition systems," *Expert Syst. Appl.*, vol. 89, no. 15, pp. 286–295, Dec. 2017.

[85] H. Guan, W. Yan, Y. Yu, L. Zhong, and D. Li, "Robust traffic-sign detection and classification using mobile LiDAR data with digital images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 5, pp. 1715–1724, May 2018.

[86] G. Loy and N. Barnes, "Fast shape-based road sign detection for a driver assistance system," in *Proc. Conf. IROS*, Sendai, Japan, Sep./Oct. 2005, pp. 70–75.

[87] C. Li, Z. Chen, Q. M. J. Wu, and C. Liu, "Deep saliency with channel-wise hierarchical feature responses for traffic sign detection," *IEEE Trans. Intell. Transp. Syst.*, to be published.

[88] C. Liu, F. Chang, and Z. Chen, "Rapid multiclass traffic sign detection in high-resolution images," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 6, pp. 2394–2403, Dec. 2014.

[89] G. Wang, G. Ren, Z. Wu, Y. Zhao, and L. Jiang, "A robust, coarse-to-fine traffic sign detection method," in *Proc. Int. Joint Conf. Neural Netw.*, Dallas, TX, USA, Aug. 2013, pp. 754–758.

[90] M. Soilán, B. Riveiro, J. Martínez-Sánchez, and P. Arias, "Traffic sign detection in MLS acquired point clouds for geometric and image-based semantic inventory," *ISPRS J. Photogramm. Remote Sens.*, vol. 114, pp. 92–101, Apr. 2016.

[91] C. G. Serna and Y. Ruichek, "Classification of traffic signs: The European dataset," *IEEE Access*, vol. 6, pp. 78136–78148, 2018.

[92] A. Wong, M. J. Shafiee, and M. S. Jules, "MicronNet: A highly compact deep convolutional neural network architecture for real-time embedded traffic sign classification," *IEEE Access*, vol. 6, pp. 59803–59810, 2018.

[93] Y.-L. Hou, X. Hao, and H. Chen, "A cognitively motivated method for classification of occluded traffic signs," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 2, pp. 255–262, Feb. 2017.