Impact of Traffic Sign Diversity on Autonomous Vehicles
A Literature Review

Ziyad N. Aldoski¹,²*, Csaba Koren²

¹ Department of Highway and Bridge, Technical College of Engineering, Duhok Polytechnic University, 61 Zakho Road, 1006 Mazi Qr Duhok, Kurdistan-Iraq
² Department of Transport Infrastructure and Water Resources Engineering, Faculty of Architecture, Civil Engineering and Transportation Sciences, Széchenyi István University, Egyetem tér 1, 9026 Győr, Hungary

* Corresponding author, e-mail: ziyad.nayef@dpu.edu.krd

Received: 08 November 2022, Accepted: 08 August 2023, Published online: 18 August 2023

Abstract
Traffic sign classification is indispensable for road traffic systems, including automated ones. There is a fundamental difference in the visual appearance of traffic signs from one country to another. Each dataset has its design standards and regulations based on shape, color, and information content, making implementing classification and recognition techniques more difficult. This paper aims to assess the influence of traffic sign diversity on autonomous vehicles (AVs) by reviewing several previous studies, comparing, summarizing their results, and focusing on classifying and detecting traffic sign datasets based on color, shape, and deep learning spaces using various methods and applications. Furthermore, it covers the main challenges facing road designers and planners considering changes to road safety infrastructure. It will be argued that compiling and standardizing a comprehensive global database of traffic signs is very difficult because it is costly and complex in application. However, it is still one of the possible solutions for the coming decades. Recommendations for future developments are also presented in this study.

Keywords
autonomous vehicles, traffic sign classification, traffic sign detection

1 Introduction
Traffic signs are visual engineering equipment located above or on the side of the roadways to communicate with road users. They are usually designed with excellent optical characteristics; accordingly, drivers can quickly notice and identify them. In addition, they are also one of the most critical elements of modern road infrastructure, serving to control traffic and provide guidance to drivers on road conditions to improve road safety and navigation (Almutairy et al., 2021; Ben-Bassat et al., 2019; Chen et al., 2022). The consistency and standardization of traffic signs became critical to the safe and efficient transport of all types of road users with the onset of the automobile industry (Almutairy et al., 2021). Nowadays, many different traffic sign datasets exist worldwide, such as European and USA traffic sign datasets. Rapid economic growth and technology have led to an increase in the use of vehicles, especially in developing countries, and with the advent of autonomous driving systems technology that can assist or even independently complement the process of moving vehicles, the importance of identifying and classifying traffic signs to improve road safety has increased. However, until now, it has been difficult for Advanced Driver Assistance Systems (ADAS) to classify them successfully due to their wide variety today (Gámez Serna and Ruichek, 2018).

This investigation aims to study the effects of traffic sign diversity on self-driving vehicles by reviewing several previous studies, comparing, summarizing their results, and evaluating their reach and limitations. An additional aim is to introduce the latest deep learning techniques from convolutional neural networks (CNNs) and their ability to identify, recognize and classify these signs. This also covers the main challenges facing road designers and planners considering changes to the road safety infrastructure. However, compiling and standardizing a comprehensive global database of traffic signs and imposing them on the countries of the world is problematic because it is uneconomic and challenging in terms of operation; nonetheless,
it is one of the possible solutions for the coming decades. The architecture of this paper has been divided into four parts. The first part investigates the importance of traffic signs for autonomous vehicles. The second part deals with classifying traffic signs, while the third presents traffic signs' variety, detecting methods, and design challenges in sections four and five, respectively.

2 Traffic signs for AVs
Traffic signs are imperative for road traffic systems, including automated driving systems. It is a means of visual communication with road users, controlling traffic, and improving road safety (Ben-Bassat et al., 2019; Magnussen et al., 2020). These signs support suitable driving conditions on the road by demonstrating crucial visual information through their categories (warning, prohibition, obligation, and informative) as well as by their shape, color, context, and location (Gámez Serna and Ruichek, 2018; Lengyel and Szalay, 2018a), for instance, information about drivable lanes, speed limits, temporary closures, restrictive areas, roadway directions, parking, etc. Therefore, particular attention should be paid to the physical and occlusive conditions surrounding these signs to be free of anomalies and to make it easier for the Traffic Sign Recognition (TSR) systems to identify them. Since drivers receive much information from these signs while traveling, human drivers can overcome these challenging conditions by making independent decisions. At the same time, recognition systems address only the problems they have been trained for (Lengyel and Szalay, 2018b). Therefore, comprehending traffic signs by the vehicle's driving control systems, especially in the case of automated vehicles, is critical to traffic flow and safety on the road.

In the past decades, the importance of traffic signs increased with the advent of autonomous vehicles. While the automotive industry has moved to the production of automated vehicles, and with the rapid development of automotive vehicle systems, such as the ADAS, highway designers and planners have had to consider changes in road safety infrastructure, including traffic signs and how to recognize them (Gámez Serna and Ruichek, 2018; Mohammed and Horváth, 2021). Consequently, traffic sign recognition is a challenging real-world problem faced by autonomous vehicle designers due to the tremendous variety in the visual appearance of traffic signs, making it difficult for recognition systems to classify them successfully (Almutairy et al., 2021; Gámez Serna and Ruichek, 2018). However, there could also be drastic modifications to the content and physicality of these signs in the future to facilitate their quick identification by the machine learning algorithms of autonomous vehicles.

3 Classification of traffic signs
Meanwhile, since autonomous vehicles will coexist in traffic with regular cars, traffic signs are inevitable for modern road infrastructure. Consequently, categorizing traffic signs is essential for autonomous driving systems (Gámez Serna and Ruichek, 2018; Lengyel and Szalay, 2018a; Zhu et al., 2016). Various classes of traffic signs have been studied in a wide range of literature depending on their importance and priority. De la Escalera et al. (1997) discussed that there are four different kinds of traffic signs in the traffic system based on their color and shape design:

- Warning signs: Designed in the shape of an equilateral triangle with one vertex upward with a white background and bounded by a red border.
- Prohibition signs: Designed in the shape of circles with a white or blue background and bounded by a red border.
- Obligation signs: They are circular with a blue background.
- Informative signs: Rectangular panels with a blue or green background indicating public places and facilities (Mohammed and Horváth, 2021).

Although the octagonal stop signs and yield signs, which have an inverted triangle shape, are essential traffic signs, the authors pointed out them as exceptions that do not fall into the above categories. Similarly, the classification of traffic signs used by many European countries is based on the Vienna Road Traffic Sign Convention (Convention on Road Signs and Signals 1968 was mentioned in the document of United Nations Economic Commission for Europe Tansport Division (2006)), which divides classes into four primary groups and subclasses, as shown in Table 1. At the same time, the Manual on Uniform Traffic Control Devices (MUTCD) (Federal Highway Administration, 2009), along with the Standard Highway Signs and Markings (SHSM) (Federal Highway Administration, 2012), is the guideline that standardizes the design and installation of traffic signs in the USA (Almutairy et al., 2021). Besides, the MUTCD has an impact on many nations in North and South America, as well as many countries worldwide, that have adopted rules similar to those described in the MUTCD (for example, Japan, Australia, New Zealand, Indonesia, Thailand, and...
Malaysia (Almutairy et al., 2021). However, there is a substantial difference in the visual appearance of traffic signs from country to country; each type has a distinctive shape and color depending on the sign's category; the standards and regulations for their design differ from one country to another, making the implementation of classification schemes more difficult (Ben-Bassat et al., 2019; Gámez Serna and Ruichek, 2018; Lu et al., 2022).

Almutairy et al. (2021) claim that an inadequate number of USA traffic sign datasets have been made publicly available in the past decade. They need to be more comprehensive to include all basic types of road signs. Only two open-source datasets are available for TSR tasks: Telenav.AI and LISA. Although the Telenav.AI dataset covers the most significant quantity of images, it contains only a small number of traffic sign classes. It concentrates on traffic lights, speed limit signs, give-way signs, turn restrictions, and stop signs. In comparison, the LISA dataset covers more types than Telenav.AI. Nevertheless, it contains slightly fewer images in the areas covered by both. In addition to the two dataset categories mentioned, the authors have introduced the Automotive Traffic Signs Repository (ARTS), a new dataset for USA Traffic Signs that covers a wide range of sign types, including warning, guiding, regulatory, and temporary signs as defined in the (MUTCD).

There are several public benchmarks for TSR worldwide and applying TSR approaches created for a specific country in other countries is challenging because traffic signs are usually different. The available datasets assist as comparison points. These publicly accessible datasets are listed below and summarized in Table 2.

- The GTSDB dataset (Saadna and Behloul, 2017) is the German Traffic Sign Detection Benchmark. It is a single-image sign detection system divided into three categories: mandatory, warning, and prohibitive. It contains 900 images with a resolution of 1360 × 800 pixels, divided into 600 training images and 300 evaluation images.
- The BTSD (Liu et al., 2019) is the Belgium Traffic Sign dataset. It is divided into three categories: mandatory, cautionary, and prohibitive, consisting of more than 10,000 annotations, and includes four Belgian videos that can be used in tracking investigations.
- LISA dataset (Magelmose et al., 2015) from the Laboratory for Intelligent and Safe Automobiles. It consists of 7855 photos with 47 types of traffic signs, only 6610 of which are annotated, and images ranging in size from 640 × 480 to 1024 × 522. It also includes videos and annotated frames.
- Swedish Traffic Signs Dataset (Saadna and Behloul, 2017) (STSD Dataset): A sequence of approximately 20,000 images was compiled using recordings made on more than 350 km of Swedish roads and manually annotated on every fifth frame from the series.
- Data Set of Italian Traffic Signs or DITS data (Saadna and Behloul, 2017) is a dataset consisting of 43,289 images taken from 14 hours of video (1280 × 720 at 10 frames per second) recorded in Italy under various conditions. The detection dataset consists of 471 test images and 1416 training images.
- Mapping and Assessment of the State of Traffic Infrastructure (MASTIF) dataset (Zang et al., 2018). The dataset contains video frames from vehicle cameras recorded in Croatia and includes three datasets of TS2009, TS2010, and TS2011; there are 6430 traffic sign images in the TS2009, the TS2010 dataset contains annotated video of 3891 traffic signs, and the TS2011 dataset contains four annotated videos of 1015 traffic signs.
4 Traffic signs diversity

There is a wide variety of traffic sign systems in the world. Despite efforts to standardize them, such as the Vienna Convention on Road Signs and Signals (United Nations Economic Commission for Europe Tansport Division, 2006), this disparity in specific traffic sign categories between countries is still a significant problem (Aggar et al., 2021; Gámez Serna and Ruichek, 2018). The USA standard is a diamond-shaped sign with a black border and a yellow background for warning signs, while a triangle with a red border and a white or yellow background is the European standard (Ben-Bassat et al., 2019), as shown in Fig. 1. Even in countries that have agreed to a standard convention for traffic signs and signals, such as common sizes, shapes, and shared colors, as in Europe, there is an apparent variation in the inscriptions and symbols of traffic signs, where each country is allowed to use its own marks (Gámez Serna and Ruichek, 2018). Fig. 2 illustrates a few examples of intra-class diversity, demonstrating that symbols differ between countries and within each of them. In Germany, two different mandatory pass-right characters are used. At the same time, France utilizes two symbols in the danger category for pedestrian crossings, with yellow and white colors for danger and prohibitory signs, although other countries use only one. There are speed limit signs in Belgium with and without the notation Km (Gámez Serna and Ruichek, 2018). This distinction is more evident in no-parking signs, which are blue round ones with a red diagonal line in most European countries. Nevertheless, some countries add a white borderline (e.g., Ukraine and Italy). Together, in Ireland, the central emblem is the letter P and white background with a red diagonal line, as shown in Fig. 3 (Ben-Bassat et al., 2019).

While in the USA, distinctions in text, rather than shape, color, or character, are used to separate sign types (Almutairy et al., 2021). Warning signs in the USA are often diamond-shaped, octagonal signs indicate a complete stop, and several regulatory and instructional signs are rectangular, with color used to identify their functions. Speed limit signs in the USA are rectangular in shape, but in most of Europe, they are circular. Fig. 1 emphasizes these similarities and differences. Thus, it is difficult for traffic sign detectors in European countries to determine the type of sign used in countries that use the USA guide. Therefore, they must be retrained if the same sensor is used in the USA. To conclude, due to the significant development in the autonomous vehicle industry in recent times and international travel in general, there must be as few differences as possible among the traffic signs regulations in various countries of the world. Achieving this will require breaking many barriers, to overcome many political obstacles, and to resolve some uncertainties that must be achieved through further research.

5 Traffic signs detection approaches

Traffic sign recognition (TSR) technology is crucial for driver assistance systems and potentially autonomous driving in identifying and tracking road signs and exposing information about them in the vehicle. In Japan, the initial study on road sign recognition was conducted in 1984, and subsequently, numerous proposals have been made to address the problem using various techniques (Gámez Serna and Ruichek, 2018). However, investigating
traffic signs classification and recognition works can only be compared after proposing the (GTSRB) and (GTSDB) Benchmarks (Gámez Serna and Ruichek, 2018). This section presents various methods and applications that have been widely discussed by researchers in the essential areas to understand the detection, recognition, and classification process of different traffic sign classes.

In the SLAIN project, European Road Assessment Programme (EuroRAP) (2021) measured the readiness of European roads for connected and automated vehicles (CAVs) in line markings and signage. The investigation included 2,000 km of roads among four countries (Italy, Croatia, Spain, and Greece). They evaluated the CAV readiness of traffic signs focusing on Greek and Croatian road segments. The method adopted was to examine TomTom’s MN-R database of sign locations and how many signs had been detected using CAV-compatible computer vision techniques by TomTom vehicles. Approximately 5.4% and 10.6% of the five key sign types (primarily speed signs) were not detected in Greece and Croatia, respectively. Signs that were high on light poles, low and attached to barriers, complex, angled to the carriageway in one direction, multilane speed signs, and cluttered with low-speed signs were among the reasons why they were not detected.

5.1 Detection according to the color space

The dominant traffic sign colors are red, blue, and yellow, with ideographs primarily black. A prevalent color-based segmentation technique is used to detect regions of interest. The RGB (Red-Green-Blue) and HSV (Hue-Saturation-Value) color spaces are the most frequent ones. However, these characteristics show sensitivity to several variables, including the variation of light, weather conditions, and the sign's retroreflectivity, making segmentation difficult. Authors are working on a variety of color spaces to address this issue, including the following:

- RGB color space,
- HSV color space
- and YCbCr color space.

Horak et al. (2016) summarized the characteristics of RGB and HSV color spaces, as shown in Fig. 4. The RGB space is in the shape of a cube, as seen in Fig. 4 (a). The limits of the minimum and maximum values along each of the three axes corresponding to the R, G, and B channels define the color-based segmentation in this space. The limits may be a straightforward orthogonal block in the RGB color space or a more complex form like an octahedron, ellipsoid, sphere, etc.

The hue saturation value (HSV) space is cone-shaped, as shown in Fig. 4 (b), which is closer to how the human eye perceives color naturally. Hue is the primary shade of the color and ranges in an angle from 0° to 360°. The other two elements are numbers from 0 to 1, which denote saturation (purity) and value (brightness).

They performed an assessment of various familiar color spaces such as RGB, HSV, and YCbCr to detect traffic lights in the European Union. The segmentation process was best suited for the HSV color space, and color-based models for the traffic light representatives were created. To recognize circles, triangles, and squares as the primary geometric shapes of traffic signs, the fast radial symmetry (FRS) approach and the Harris corner detector were utilized. The recognition method’s total accuracy reached approximately 93%.

Lai et al. (2018) introduced A CNN-SVM-based identification and classification system. The YCbCr color space is used in this technique to divide the color channels and extract features. The components of the blue and red differences are Cb and Cr, respectively, where Y is the illumination factor. This color space was mainly used for the ongoing processing of images and videos that were taken inside the vehicle. The effectiveness of this strategy was 98.6%.

De la Escalera et al. (1997) adopted the intuitive RGB color space because the HSI formulas are non-linear, with the three elements, red, green, and blue, determining the color of each pixel. The authors used the relationship between these components of the color threshold as the following expression:

\[
R_g \leq f(x, y) \leq R_b \\
\begin{align*}
g(x, y) &= k_i \\
T_G &= \frac{f(x, y)}{f(x, y)} \leq T_G \\
TB &= \frac{f(x, y)}{f(x, y)} \leq TB_b \\
g(x, y) &= k_i \quad \text{in any further case.}
\end{align*}
\]
Where the functions that provide the red, green, and blue levels of each point in the image are $f_r(x, y)$, $f_g(x, y)$, and $f_b(x, y)$, respectively.

Several researchers used the RGB color space threshold (Saadna and Behloul, 2017). However, their techniques are related to the selected points, making comparing their performances a problematic task. Other authors (Ellahyani et al., 2016; Mohammed and Horváth, 2021; Saadna and Behloul, 2017; Soheilian et al., 2013) have argued that several factors influence color segmentation algorithms with different parameters, making it challenging to detect traffic signs in authentic images by utilizing a clear boundary directly in the RGB space. These factors include lighting differences, inclined road signs, challenging weather conditions, the impact of other objects on the street that have the same color as the signs, and many different constraints. This prompted many researchers to experiment with other spaces. Yakimov (2015) emphasized HSV color space as the most suitable for extracting red color in images using an improved generalized Hough transform algorithm to detect traffic signs. This enables detecting and recognizing signs in Full HD 1920 × 1080 images from a real-time video sequence.

Cao et al. (2019) utilized the HSV color space to determine the optimal values for the threshold segmentation, which has a faster detection speed, less impact by illumination, and a better segmentation advantage compared to RGB and HSI color spaces. Fig. 5 illustrates the image converted from RGB to HSV using an inverted cone converter.

### 5.2 Detection according to the shape space

In this method, many authors do not consider color segmentation a discriminatory property due to its susceptibility to many conditions, including atmospheric conditions, target distance, sign reflection, and illumination intensity. In contrast, sign shape-based methods are more reliable than chromatic procedures because they can process grayscale images and manipulate their gradations. However, the processing rate is highly dependent on the number of edges detected; it is time-consuming and uneconomical.

Chincholkar and Kumar (2019) used the Hough Transform (HT) technique to recognize traffic sign panels from video sequences. Before using the Canny edge detector to verify the shape of the signs, the video frame was first processed by transforming RGB color images into grayscale images using a preprocessing method. Next, the Hough Transform algorithm measured the video's attributes and image regions for further analysis. The SVM classifier categorizes the traffic sign board into different classes on MATLAB.

Vishwanathan et al. (2017) compared three different edge detection methods. Comparisons were made on still images and video of the octagonal stop sign under various situations. The three methods used are:

- **Canny method**: It is a frequently used method for edge detection (Canny, 1986). This technique's primary conditions are to lower the error proportion and reduce the gap between the points marked by the detector and the center of the actual edge.
- **Zhang method**: This edge detection procedure follows the principle of linear prediction (Zhang et al., 2010). The main idea behind this linear prediction method is to optimize filter coefficients to minimize prediction errors.
- **Sobel method**: This edge detection technique is based on the idea of the image gradient (Sobel, 1970). It uses two kernels, each of order (3, 3), convolved with the original image to approximate the derivatives along the horizontal and vertical directions.

They concluded that Zhang’s method (linear prediction) is less sensitive to the conditions of the original images than the Canny and Sobel methods. Wherein it accurately locates the edges of the STOP sign under various conditions, such as during snow or rain.

Lu et al. (2022) focused on a two-stage structure for detecting and recognizing traffic signs. As a first step Mask R-CNN was applied to identify the position of traffic signs and their equivalent shapes, which were then divided into 23 categories based on their data. The second step seeks to solve the classification problem by training a second CNN model to learn more about the content of
the detected object using a model called Xception. Finally, they were able to achieve an accuracy of 99.73% for the circular shape categories and 98.45% for the triangular classes, reasoning that this slight variance in precision could be attributed to the fact that the number of triangular traffic signs in the dataset being higher than the circular ones, which makes classifying the model more complex.

Behloul and Saadna (2014) propose another method to recognize traffic sign shapes; by comparing the detected pattern with the BoxOut rectangle that includes it. Fig. 6 demonstrates the calculation of the degree of intersections between the perimeter of the model and the four BoxOut lines. Although 2.17% of failed alarms are due to the weakness of this approach in noise and occlusions, this method can detect 95.65% of sign shapes from the dataset consisting of 48 images per a resolution of 360 × 270 pixels covering three various traffic signs.

Junaid et al. (2021) focused on Mask R-CNN to detect objects (pedestrians) while the vehicle is traveling on the road, and for image manipulation, the inverse gamma correction method was used, which is directly related to the intensity of illumination. Six backbone models of Mask R-CNN were tested on the Penn-Fudan database in the process of feature extraction and bounding box identification. The comparison results (Fig. 7) showed that the best model for real-time detection systems is Mask R-CNN ResNet50, which performs well in different illumination conditions, whether dark or bright.

Recently, a new algorithm (EDCircles) was developed by Akinlar and Topal (2013), and Kaplan Berkaya et al. (2016) began using this algorithm to detect circular traffic signs. The first step in this technique is to use Edge Drawing Parameter Free (EDPF) algorithm to find edges in grayscale images. Then the circular arcs are extracted from the edges, arcs of similar radius are combined, and the contender circles are confirmed. The authors used this method on the GTSDB dataset; the detection accuracy rate for prohibitive signs achieved 93.78% with only 0.99% false alarms, while for mandatory signs, 75.51% accuracy with 2.04% false positive detections.

5.3 Detection based on deep learning

The previous techniques had weaknesses in terms of various factors, for example, scale change, changes in illumination, occlusions, translations, and rotations. However, machine learning may potentially solve these issues, although this necessitates a sizable collection of annotated data (Saadna and Behloul, 2017). Convolutional Neural Networks (CNNs) are a type of artificial neural network (ANN) used to evaluate visual images in deep learning, and it has been in use since the late 1980s. Several researchers have used CNN to study the identification classification of traffic signs.

Gámez Serna and Ruichek (2018) have reasoned by examining several research papers in the field of traffic sign recognition that CNNs can solve traffic sign recognition problems and ranked five networks with the best performance based on rating traffic signs, which are:

1. LeNet-5: They are CNNs mainly used for handwriting recognition. They consist of seven sheets: three convolutional layers, sub-sample layers (excluding the final layer), one fully linked layer, and a final output layer composed of Euclidean RBF units.
2. IDSIA Model: A multi-column deep convolutional neural network (MCDNN) is created by combining the number of columns of a deep convolutional neural network DNN. Its network consists of two convolutional layers followed by Max pooling layers. Two fully connected hidden layers transmit the output
to a fully connected final layer with six neurons to achieve classification.

3. URV Model: This CNN consists of three convolution-pooling layers, two completely connected layers, and a dropout layer to prevent overfitting; after each convolutional layer and the first fully connected layer, Rectified Linear Unit (ReLU) activations are performed.

4. CNN with Asymmetric Kernels: Three convolutional layers with symmetric kernels, six convolutional layers with asymmetric kernels, and two fully connected layers framework this CNN design. Except for the final layer (Softmax classifier), each layer is followed by Batch Normalization and ReLu activations.

5. CNN 8-Layers: Although it is not considered a deep network, it is highly accurate compared to the most up-to-date technologies. The network architecture with numerous hidden layers produced the best results. However, it is complex and takes a long time to compute. As a result, primary networks are utilized for information preprocessing and data augmentation if the dataset contains more examples for the learning phase.

The CNN models, as mentioned above, have been trained on the GTSDB dataset and the proposed European dataset (Germany, Netherlands, Sweden, France, Croatia, and Belgium) for traffic signs with intraclass variability; the results are shown in Table 3.

Almutairy et al. (2021) focused on Faster R-CNN, YOLOv3, and RetinaNet to assess the performance of the latest deep learning detectors on datasets in the USA destined for traffic sign recognition tasks such as (LISA) and their new dataset, the Automotive Repository of Traffic Signs (ARTS). They emphasized that all models have been tested and trained on the LISA dataset (the parent and additive together) and on the new dataset (ARTS). The results show that image resolution is the most important key factor for performance/accuracy assessment among these deep learning algorithms. RetinaNet-50 performs poorly on the LISA dataset because it has a high proportion of low-resolution images. Another critical factor is the variance of the dataset. The experiment found that on the challenging dataset, which includes most classes among the three categories, RetinaNet-50 outperformed YOLOv3 and Faster-RCNN. Through harnessing the power of the Focal Loss function, RetinaNet evolves better with more extensive and significantly scattered datasets; the three algorithms did reasonably well-configured on the Easy dataset, which was discarded in order to make the task simpler. Tables 4 and 5 illustrate the results.

Zang et al. (2018) applied the Quaternion Convolutional Neural Network (QCNN) by integrating Spatiotemporal features into a single frame. They used the (MASTIF) traffic dataset. The fastest R-CNN (including 12 convolutional layers and four pooling layers) was applied to identify the traffic sign areas; then, the detected traffic signs were tracked in three frames using the motion constraint method of Mean Shift. Finally, three QCNNs, as shown in Table 6, were applied, two of them to extract the Spatiotemporal features. The feature maps acquired in both domains were combined as inputs to the third QCNN to achieve the final identification results. In their tests, they reached 99.31% in sign detection with 1.19% false positive and 99.15% accuracy in classification with 0.17% false alarms.

Farag (2018) developed the (CNN) based classifier "WAF-LeNet" as a comprehensive classifier for recognizing and identifying traffic signs. The applied structural design is a fifteen-layer deep network selected after extensive testing to be high-speed and was trained using Adam’s optimization algorithm as a modification of the Stochastic Gradient Descent (SGD) method. WAF-LeNet performed well in recognizing 43 different categories of traffic signs selected from the (GTSRB) dataset with an accuracy of 96.5% and identifying 100% in the robustness test. Finally, the author summarized several recommendations for improving the outcomes, including deepening CNN by having additional layers, switching to RGB color images from grayscale ones, and including "Skip Connections" in CNN.

### Table 3 Results of accuracy in percentages were obtained for the European and GTSRB test groups (Gámez Serna and Ruichek, 2018)

<table>
<thead>
<tr>
<th>Model</th>
<th>Input size</th>
<th>European Parameters (millions)</th>
<th>Accuracy %</th>
<th>European Parameters (millions)</th>
<th>Accuracy %</th>
<th>GTSRB Parameters (millions)</th>
<th>Accuracy %</th>
<th>Time milliseconds (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LeNet-5</td>
<td>32 × 32 × 1</td>
<td>0.35</td>
<td>89.8</td>
<td>0.13</td>
<td>89.1</td>
<td></td>
<td></td>
<td>0.0067</td>
</tr>
<tr>
<td>IDSIA</td>
<td>48 × 48 × 3</td>
<td>1.58</td>
<td>95.82</td>
<td>1.54</td>
<td>94.62</td>
<td></td>
<td></td>
<td>0.6</td>
</tr>
<tr>
<td>URV</td>
<td>48 × 48 × 3</td>
<td>1.16</td>
<td>96.53</td>
<td>1.12</td>
<td>96.1</td>
<td></td>
<td></td>
<td>0.61</td>
</tr>
<tr>
<td>CNN asymmetric</td>
<td>48 × 48 × 3</td>
<td>2.95</td>
<td>98.48</td>
<td>2.92</td>
<td>97.88</td>
<td></td>
<td></td>
<td>0.39</td>
</tr>
<tr>
<td>CNN 8-Layers</td>
<td>48 × 48 × 3</td>
<td>1.51</td>
<td>97.88</td>
<td>1.48</td>
<td>98.52</td>
<td></td>
<td></td>
<td>0.15</td>
</tr>
</tbody>
</table>
In the Arab region, in the paper of Alghmgham et al. (2019) convolutional neural networks were used to create the Automatic Arabic Traffic Signs Recognition System (AATS); the Saudi Arabian Traffic and Road Signs (SA-TRS-2018) database was used for Testing and Training. The final Deep CNN architecture recommended in their work includes two convolutional layers, two max-pooling layers, and three dense layers, with 100% accuracy achieved for epoch 150 for all batch sizes.

Tabernik and Skočaj (2020), in their study, focused on the challenge of detecting and recognizing a wide variety of traffic sign classes that could be used to automate traffic sign inventory management. To address the entire pipeline of detection and recognition through end-to-end self-learning, they used a convolutional neural network Mask R-CNN technology called ResNet-50 on 200 classes of traffic signs represented in their new dataset. After a comprehensive analysis of their deep learning method for traffic sign detection, the average error rate was about 2–3% of the actual detections.

Tables 7, 8, and 9 offer a comprehensive summary of research evaluating different detection modalities for traffic signs, including color-based, shape-based, and deep learning-based approaches. Each table provides an overview of traffic signs’ classification, detection, and recognition, along with the corresponding processing speed achieved in the studies. Table 7 focuses on evaluating the color-based detection modality, while Table 8 highlights the evaluation of the shape-based detection modality. Lastly, Table 9 summarizes the research conducted on deep learning-based detection, encompassing the detection, recognition, and categorization of traffic signs alongside the attained processing speed.

### 6 Traffic sign design challenges

Even though the TSR has advanced significantly in recent years, the vast diversity of traffic signs worldwide indicates that the problem of consistent and accurate detection and identification remains indefinable. Professional road designers and planners face a considerable challenge. They must reflect on appropriate solutions to changes to

---

**Table 4** Configuration of the experiment (Almutairy et al., 2021)

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Classes</th>
<th>Annotations</th>
<th>Images</th>
</tr>
</thead>
<tbody>
<tr>
<td>LISA 2012+2015</td>
<td>56</td>
<td>6644</td>
<td>5557</td>
</tr>
<tr>
<td>Validation</td>
<td>56</td>
<td>2088</td>
<td>1886</td>
</tr>
<tr>
<td>Testing</td>
<td>56</td>
<td>2795</td>
<td>2481</td>
</tr>
<tr>
<td>ARTS Easy</td>
<td>62</td>
<td>5058</td>
<td>3828</td>
</tr>
<tr>
<td>Validation</td>
<td>62</td>
<td>1739</td>
<td>1277</td>
</tr>
<tr>
<td>Testing</td>
<td>62</td>
<td>2209</td>
<td>1702</td>
</tr>
<tr>
<td>ARTS Challenging</td>
<td>175</td>
<td>15198</td>
<td>9012</td>
</tr>
<tr>
<td>Training</td>
<td>175</td>
<td>5024</td>
<td>3005</td>
</tr>
<tr>
<td>Testing</td>
<td>175</td>
<td>6959</td>
<td>4006</td>
</tr>
</tbody>
</table>

**Table 5** Traffic sign recognition benchmark (Almutairy et al., 2021)

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Mean Average Precision m.AP50</th>
<th>YOLOv3-416</th>
<th>RetinaNet-50</th>
<th>Faster-RCNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>LISA 2012+2015</td>
<td>81.9%</td>
<td>54.9%</td>
<td>84.0%</td>
<td></td>
</tr>
<tr>
<td>ARTS Easy</td>
<td>90.5%</td>
<td>81.09%</td>
<td>86.9%</td>
<td></td>
</tr>
<tr>
<td>ARTS Challenging</td>
<td>65.4%</td>
<td>67.3%</td>
<td>36.9%</td>
<td></td>
</tr>
</tbody>
</table>

**Table 6** Quaternion convolutional neural network (QCNN) parameter settings (Zang et al., 2018)

<table>
<thead>
<tr>
<th>Layer</th>
<th>Method</th>
<th>Parameters</th>
<th>Class</th>
<th>Neural network</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>QCNN 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Layer 1</td>
<td>Convolution</td>
<td>8 kernels (5 × 5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Layer 2</td>
<td>Convolution</td>
<td>12 kernels (3 × 3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Layer 3</td>
<td>Pooling</td>
<td>Kernel size (2 × 2)</td>
<td></td>
<td>QCNN 2</td>
</tr>
<tr>
<td>Layer 1</td>
<td>Convolution</td>
<td>8 kernels (5 × 5)</td>
<td></td>
<td>QCNN 3</td>
</tr>
<tr>
<td>Layer 2</td>
<td>Convolution</td>
<td>12 kernels (3 × 3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Layer 3</td>
<td>Pooling</td>
<td>Kernel size (2 × 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Layer 1</td>
<td>Convolution</td>
<td>8 kernels (5 × 5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Layer 2</td>
<td>Convolution</td>
<td>12 kernels (3 × 3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Layer 3</td>
<td>Pooling</td>
<td>Kernel size (2 × 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Layer 4</td>
<td>Full connection</td>
<td>6320 neurons</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 7** Color space detection approach

<table>
<thead>
<tr>
<th>Authors</th>
<th>Application</th>
<th>Method</th>
<th>Color space</th>
<th>Detection rate %</th>
<th>False alarm</th>
<th>Used dataset</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horak et al. (2016)</td>
<td>Detection</td>
<td>FRS and Harris Corner</td>
<td>RGB</td>
<td>93%</td>
<td>-</td>
<td>343 images</td>
<td>420 ms</td>
</tr>
<tr>
<td>Lai et al. (2018)</td>
<td>Classification</td>
<td>CNN-SVM</td>
<td>YCrCb</td>
<td>98.6%</td>
<td>-</td>
<td>1000 images</td>
<td>-</td>
</tr>
<tr>
<td>de la Escalera et al. (1997)</td>
<td>Classification</td>
<td>CNN</td>
<td>RGB</td>
<td>97%</td>
<td>-</td>
<td>1620 images</td>
<td>30–40 ms</td>
</tr>
<tr>
<td>Yakimov (2015)</td>
<td>Detection</td>
<td>GHT</td>
<td>HSV</td>
<td>97.3%</td>
<td>2.7%</td>
<td>GTSDB</td>
<td>-</td>
</tr>
<tr>
<td>Cao et al. (2019)</td>
<td>Recognition</td>
<td>LeNet-5 CNN</td>
<td>HSV</td>
<td>99.75%</td>
<td>-</td>
<td>GTSRB</td>
<td>5.4 ms</td>
</tr>
</tbody>
</table>
the existing road infrastructure to improve traffic safety required in the future (Ben-Bassat et al., 2019; Gámez Serna and Ruichek, 2018). Traffic signs are one type of road traffic engineering equipment; their design change is one of those challenges that will affect the road infrastructure. Ben-Bassat et al. (2019) studied traffic signs to assess the compatibility of traffic signs with similar meanings by presenting a small dataset of traffic signs to road engineering experts in several different countries. They found that many Vienna Convention signs could be improved using ergonomics assessment techniques. It is also possible to design signs that meet the three essential ergonomics criteria: familiarity, compatibility, and standardization. When improving the signs, the focus should be on the global compatibility of signs.

Sayin et al. (2020) have contended that smart road signs, which have intelligent codes (such as those visible in infrared) on their surface to provide innovative vehicles with more accurate and detailed information, are a potential trend in future intelligent transportation systems. Since humans cannot realize or understand these intelligent road signs, they do not participate in their identification. These smart codes make the road sign classification problem more compatible with communication settings than the traditional classification.

In their study, Lengyel et al. (2021) presented a set of influential printable and publishable labels that can consistently mislead a traffic sign classification algorithm based on deep learning. Natural traffic signs and printed physical labels were used in the laboratory, and a 99.00% NN self-generated mobile application was used to assess hostile detection and attack. Suitable test models for vehicles have yet to be obtained. Thus, cognitive modules based on deep learning that are also vulnerable to adversarial attacks are only sometimes clear from the information provided by the manufacturers.

7 Conclusion

This paper has argued that traffic signs, are imperative for road driving, including autonomous driving systems. This study has briefly reviewed the impact of traffic sign diversity and recognition by autonomous vehicles and focused on classifying traffic sign datasets using different methods; these methods are divided into three classes:
color-based categorization according to color space, shape-based detection and deep learning methods. This theme improves road safety by transmitting information about the vehicle's external environment, such as speed limits, drivable lanes, and temporary closures to the vehicle's driving system, which is the decision control center. The investigation with CNNs showed that despite the various types of deep learning networks in this field, many have outstanding advantages for classification accuracy and algorithm time consumption; the detection rate in modern methods ranges between 90 to 100%. However, it is difficult to determine the best approach among them. An ADAS application that can detect and classify road traffic signs in real-time is still needed to be developed.

Accordingly, the question is: do the current methods of detecting traffic signs have the ability to prove the same performance in actual applications with different datasets? Can these methods detect, distinguish, and make the correct decision at the right time in cases of combining different types of signs on a single pole, as humans do? Therefore, this paper recommends that these issues be resolved before AVs start moving and mixing with traffic in larger amount. It also urges that all relevant governments, organizations, road designers, and planners consider changes to the road safety infrastructure and attempt to compile and unify a comprehensive global database of traffic signs and find appropriate algorithms to identify them before proceeding with larger number of AVs on the road networks. Future research should investigate:

1. the Impact of the physical properties of current traffic signs upon AVs equipped with LiDAR's, sensors and cameras,
2. the effects of the traffic sign retroreflectivity on the accuracy of traffic sign recognition,
3. the possibility of using an intelligent barcode traffic sign, and
4. the possibility of combining more than one sign on one pole.

References


European Road Assessment Programme (EuroRAP) (2021) "SLAIN Project (Saving Lives Assessing and Improving TEN-T Road Network Safety)", [online] Available at: https://eurorap.org/slain-project [Accessed: 04 November 2021]


European Road Assessment Programme (EuroRAP) (2021) "SLAIN Project (Saving Lives Assessing and Improving TEN-T Road Network Safety)", [online] Available at: https://eurorap.org/slain-project [Accessed: 04 November 2021]


https://doi.org/10.1007/978-3-319-30222-5_21

https://doi.org/10.1177/0959651818758865

https://doi.org/10.1109/MWSCAS.2010.5548904

https://doi.org/10.1109/CVPR.2016.232