Road signs recognition: state-of-the-art and perspectives

Btissam Bousarhane*

Smart Systems Laboratory (SSL),
National School of Computer Science and Systems Analysis ENSIAS,
Mohammed V University,
Rabat, Morocco
Email: bousarhane83@yahoo.fr
*Corresponding author

Saloua Bensiali

Department of Applied Statistics and Computer Science,
Hassan II Institute of Agronomy and Veterinary Medicine,
P.O. Box 6202, Morocco
Email: s.bensiali@iav.ac.ma

Driss Bouzidi

Smart Systems Laboratory (SSL),
National School of Computer Science and Systems Analysis ENSIAS,
Mohammed V University,
Rabat, Morocco
Email: driss.bouzidi@um5.ac.ma

Abstract: Making cars safer is a crucial element of saving lives on roads. In case of inattention or distraction, drivers need a performant system that is capable of assisting and alerting them when a road sign appears in their field of vision. To create such type of systems, we need to know first the major difficulties that still face traffic signs recognition, as presented in the first and second sections of this paper. We should also study the different methods proposed by researchers to overcome each of these challenges, as proposed in the third section. Evaluation metrics and criteria for proving the effectiveness of these approaches represents also an important element which section three of this article presents. Ameliorating the existing methods is crucial to ensure the effectiveness of the recognition process, especially by using deep learning algorithms and optimisation techniques, as discussed in the last section of this paper.

Keywords: road signs recognition; detection; classification; tracking; machine learning; deep learning; evaluation datasets; evaluation metrics; optimisation.

Biographical notes: Btissam Bousarhane is a PhD student at the National School of Computer Science and Systems Analysis ENSIAS, Smart Systems Laboratory (SSL), Mohammed V University, Rabat, Morocco.

Saloua Bensiali is a Professor at Hassan II Institute of Agronomy and Veterinary Medicine, Department of Applied Statistics and Computer Science, Morocco.

Driss Bouzidi is a Professor at the National School of Computer Science and Systems Analysis ENSIAS, Smart Systems Laboratory (SSL), Mohammed V University, Rabat, Morocco.

1 Introduction

Traffic accidents have a huge socio-economic impact on health and countries' development, and have also a considerable impact on quality of life. In addition to the significant economic costs of road accidents, traffic deaths exceed, globally, 1.2 million people each year, according to the Global Status Report on Road Safety [WHO, (2015), p.VII]. The report shows that 90% of road traffic fatalities were reported in low-and middle-income countries, knowing that these countries account for only 54% of the world's registered vehicles. The report highlights also that the risk of road traffic deaths is highest in the African region.

Morocco is one of the African countries that suffer enormously from this problem, and it pays each year, a heavy price in terms of human lives losses and economic costs. According to the latest statistics released, in 2017, by the Ministry of Equipment, Transport and Logistics (MAP, 2018b), the country has seen a rise of 9.99% in the number of traffic accidents (89,998), compared with 2016. However, the number of traffic fatalities has felled by 2.62% (943 deaths were recorded in urban areas, against 2,556 outside these areas). In fact, there are three main factors for traffic accidents in Morocco, which are: the quality of road infrastructure, the inadequate vehicles and the human factor. To face these huge problems, the country has adopted a new strategy (2017–2026 National Road Safety Strategy), that aims to halve the number of fatalities in traffic accidents by 2026 (MAP, 2018a).

In terms of infrastructure, Morocco makes considerable efforts to make roads safer, by implementing a certain number of effective measures to fix and improve their infrastructure [WHO, (2015), p.184]. Concerning the human factor, the National Committee of Traffic Accidents Prevention conducts each year many social campaigns to promote road safety, which especially targets changing road users' behaviour. Despite the progress that has been made by the country over the last decade, additional efforts should be made to address the challenge of rendering vehicles safer, knowing that there is an important increase in the number of registered vehicles in Morocco, and knowing also that drivers and cars' passengers represents 45% of traffic road deaths [WHO, (2015),

p.184]. Furthermore, a great number of vehicles fail to meet even the most basic international standards on vehicle safety.

Making cars safer, represents then, a crucial component of saving lives on the roads, and it is one of the most efficient ways to ensure road safety. Actually, vehicles' technologies have progressed enormously and can fulfil this role, by assisting drivers, specifically in case of inattention, distraction and adverse conditions (weather and light changes, etc.), and that by the integration of embedded systems that automate, especially, the detection and the recognition of road signs. These new technologies contribute also to make vehicles more autonomous, which means more secure for both drivers and pedestrians.

In this paper, our goal is to realise an overview on road signs recognition, to know the major difficulties and challenges that still face the progress of this field of research. In the second part of the paper, we will present a comparative study between, the classical approaches adopted, and those based on machine learning, for road signs recognition. For the third part, we will discuss the context conditions and limitations that should be taken into consideration, when evaluating the performance of these methods. Concerning the last part, we will present amelioration perspectives and challenges, for this field, in the context of deep learning.

2   Road signs recognition: difficulties and challenges

Traffic signs recognition is a process that contains two principal stages, which are detection and classification. The first stage consists on finding the precise location and size of potential road signs, while the classification stage consists on interpreting the meaning of information within each found sign, in order to identify its type. When addressing the recognition of road signs, many aspects should be taken into consideration to ensure the effectiveness of any system that targets to realise this complicated operation.

One of these important aspects is the diversity of road signs' types. Nevertheless, we notice that the majority of works that address this field of research (Akatsuka and Imai, 1987; Besserer et al., 1994; Safat et al., 2015b), etc., focus generally on one type of road signs, which are notably the static ones (horizontal and vertical), and tend to neglect the dynamic, electronic or variable message signs (United Nations, 1995), although they are integrated in Vienna Convention since 1995, and are actually used in roads by many countries.

For static road signs, their recognition still presents many challenges for researchers, due to their aspect that differs from country to country. In order to unify their appearance, many countries have adhered to the Vienna Convention on Road Signs and Signals, like Morocco for example. In contrast, we find that many other countries have not signed the convention, and have developed, instead, their own standards, like Canada, USA, etc. It is important to clarify also that even within the Vienna Convention itself, we find some differences in road signs' aspects, adopted by each country, thing which presents another problem that recognition systems should deal with.

Regardless of the context of use, many other external factors could affect, considerably, the reliability of the recognition process, precisely environmental, but also human factors. The environmental factors are relating, especially, to illumination and

weather changes (Lim et al., 2017; Safat et al., 2015a; Romdhane et al., 2016), which can affect the visibility, but also the appearance of road signs (fading colours, etc.). For the human factors, we can mention for example: damaged signs (by accidents, etc.), partial occlusions (Rehman et al., 2017), by pedestrians, cars or objects installed near to signs, like trees, etc.

The road scene presents also a problem for recognition systems (Surinwarangkoon et al., 2013), because it contains many objects that look like traffic signs (in terms of shape and colour), and contains also many cluttered objects that make the recognition more delicate. Furthermore, acquisition conditions influence, significantly, the process and could make it more complicated due to the angle of vision (Safat et al., 2015a), and the effect of vehicles' motion, which could make acquired images look blurry.

For dynamic road signs, although the fact that they are very different from the static ones, their recognition could face also a certain number of these challenges, especially, those relating to the complexity of the road scene, the acquisition conditions, the change of their aspect, and the external factors that could affect their visibility and appearance. In addition to these problems, the recognition of variable message signs faces other types of problems (Nienhuser et al., 2008; Wahyono et al., 2013, 2015), that are, principally, relating to the used technology (inadequate frequency between the camera and the variable signs), and also to the difficulties in recognising the displayed information, especially textual, because each character is represented by a certain number of dots, thing which complicates, enormously, the recognition process.

To face these challenges, that still face the recognition of static and dynamic road signs, considerable progress has been made by researchers over the last ten years. However, we find that, for static road signs, there is no approach that can handle the totality of these problems, and almost half of the faced challenges are treated by the existing approaches (Safat et al., 2015b). The literature review shows also that researchers focus more on some problems, rather than others, especially those relating to illumination changes, weather conditions, and occlusions in first place. In contrast, we notice that the effect of vehicles' motion, and damaged signs, for example, are not enough studied by researchers.

Concerning dynamic road signs, as mentioned before, few works have treated the recognition of this type of signs, and some researchers adopt, almost, the same approaches adopted for static ones. The adopted approaches are presented in the next section.

In fact, the impact of all these challenges is very obvious on cars' industry. We find that, many car brands have a system for traffic signs recognition (RACC and ADAC, 2011). However, almost all these brands realise the recognition of only one or two types of signs, which are exactly speed limit and no overtaking signs. Furthermore, the recognition process, in the majority of these brands, needs a navigation system to ensure the detection and the classification of road signs. Thus, they cannot be used as a standalone system, and their effectiveness depends, generally, on navigation system data, that should be actualised.

Despite all these challenges, the field of road signs recognition is in continuous progress, and new methods and techniques are proposed by researchers to overcome a certain number of these problems. These methods can be categorised in two main approaches, which are classical and machine learning-based approaches.

3   Classical vs. machine learning approaches

To ensure the recognition of traffic signs, researchers adopt multiple methods and techniques that can be divided in two main approaches, which are classical and machine learning approaches. Proving the effectiveness of any technique within these two approaches still raises a certain number of questions for researchers, especially concerning the appropriate metrics used, and evaluation criteria for methods benchmark.

3.1   Classical approaches

For road signs recognition, the first type of approaches (classical) relies, essentially, on road signs colour and geometric shape to realise the detection stage. Colour information is widely used by researchers to find the region of interest (ROI) that probably contains a road sign. To extract this type of information, many colour spaces are used: RGB, HSV, HIS, HSL, CIELAB, YCbCr, YUV, YIQ, OHTA, CMYK, CIECAM, etc. (Huang et al., 2017; Yang et al., 2016; Gao et al., 2008), etc.

We find that, the two spaces HSV and HSI are very used by researchers (Ardianto et al., 2017; Huang et al., 2017), etc. because they are based on human perception of colours, and encode colour information in one channel instead of three (contrary to RGB colour space). Certainly, the thresholding with HSV and MSER features (Sumi and Arun, 2017), for example, gives good results for the detection process. In contrast, we find that using OHTA space and HOG features ensures more effectiveness and accuracy for the detection (Yang et al., 2016). The comparison between a certain number of features confirms that the use of SIFT, HOG and TCH increases detection performances (Gokul et al., 2015).

Moreover, the results of the evaluation of segmentation algorithms, realised by Gomez-Moreno et al. (2010), show that RGB and OHTA normalised represent the best methods for traffic sign recognition. They confirm also that the use of some spaces like HIS does not improve, necessarily, the performance of the recognition process. Colour segmentation in RGB space is also used for the detection of dynamic road signs, like the approach adopted by Wahyono et al. (2015). The adopted method uses also density-based spatial clustering of applications with noise (DBSCAN) to group in regions the points that represent characters.

Unlike colour-based methods, geometric methods extract the ROIs by using information about traffic signs shape (circular, triangular, rectangular, etc.). This type of techniques is used, by researchers, to minimise the effects of illumination and weather changes. For road signs detection, these methods use a certain number of algorithms, like Hough transform, cross correlation, invariant moment, log-polar transform, fast radial transform, etc. As is the work of Nienhuser et al. (2008), which proposes a method for the detection of static and variable signs (speed limit signs), as well as their attribution to the concerned road, and that by using Hough transform.

For the recognition of Moroccan traffic signs, Romadi et al. (2017) adopt also an approach based on Hough Transform and Ramer-Douglas-Peucker filter to extract circular and triangular signs. For the classification, the method is based on a correlation technique using FLANN algorithm and SURF descriptor. Shape-based methods are very

sensible to occlusions and road signs deformation that affects considerably their performances. To overcome this problem, some researchers use Fuzzy shape recogniser and 3D reconstruction to ameliorate the detection results (Safat et al., 2015b).

In fact, traffic signs recognition is a more complicated question that needs a global vision to ensure its effectiveness, than just adopting one or more techniques within these classical approaches. From a realistic perspective, a road signs recognition system should face the totality of the challenges presented in the previous section. However, we find that, the methods used for illumination and weather changes, for example, adopt shape-based techniques to minimise the changes in road signs colours. In contrast, this same solution presents a problem for recognising occluded and damaged signs that need on the contrary a colour-based method. Another element, that is very important in this context, is that using a colour or a shape-based method will, certainly, mislead the system by detecting many objects in the road scene that are similar to traffic signs.

For that reason, researchers in recent papers choose, generally, to combine colour and shape information (Agrawal and Chaurasiya, 2017; Ellahyani and El Ansari, 2017b; Ellahyan et al., 2016; Yang et al., 2016) to improve the detection performances. Adopting this approach of combination can, certainly, help to minimise the rate of false positives, and that by eliminating the object that not have the same colours and shapes of road signs, but in contrast, that will decrease the rate of true positives. This problem is due to the fact that researchers that opt for this combination use, generally, colour method in the first place, and adopt then a shape-based technique. In this situation, the road signs affected by illumination changes and faded signs will not be correctly detected, Furthermore, this approach could be time consuming, especially when using vote algorithms based on edge detection, like Hough Transform, for example, thing which limits their use in real time systems.

To solve this problem, opting for a simultaneous (parallel) use of these two techniques (colour and shape) seems to be more effective. Surely, that will increase considerably the rate of true positives, but it will also increase the rate of false positives representing similar objects that look like traffic signs. To overcome some of these difficulties, a promising technique is proposed by some researchers, which is exactly the tracking. In fact, this method represents a key technique in many computer vision applications, including object and road signs recognition. The main objective of this technique is to track the position, the occlusion and the motion of segmented ROIs represented some objects, and that in every frame of a video scene. Integrating a tracking technique within these classical approaches, is very useful to handle some of the faced problems, and that by helping to predict the position of the road signs in the next frames, thing which will limit the zone of research to a very small region, and will decrease, considerably, the number of false positives, and ensure the effectiveness of true positives detections.

As a conclusion, we can say that, although the considerable improvement proposed by the classical approaches, this type of methods cannot ensure alone the effectiveness of traffic signs recognition. This situation lead us wonder if there is another type of approaches that can overcome all these challenges. To answer this question some researchers explore another type of approaches to solve these problems, which are precisely machine learning approaches.

3.2   Machine learning approaches

Machine learning approaches are very efficient, because they are, generally, based on a large dataset of annotated data that helps to train the algorithms to recognise efficiently road signs in different situations: lightening changes, adverse conditions, occlusions, damaged signs, etc.

For the detection stage, representing the ROIs with HOG features and a classifier like support vector machine (SVM) and RBF kernel (Yang et al., 2016) ameliorates, considerably, the obtained results. In the same context, using SVM and HOG with Liner Kernel, instead, shows also good results in terms of accuracy (Ma and Huang, 2015). Another performant method used in this stage, is convolutional neural networks, CNNs (Zhang et al., 2017), which are used, essentially, to deal with multiclass problems. Random forest (Ellahyani and El Ansari, 2017a) represents also one of the most popular methods used in this stage, and in some researches we find also the use of AdaBoost (Di et al., 2016). During this stage, we find that using a tracking technique like Kalman filter (Møgelmose et al., 2014), PWP 3D (Safat et al., 2015b), etc. helps to improve the recognition rate and optimises the response time.

CNNs are also used in classification stage, as shown in many researches in the literature (Yang et al., 2016; Stallkamp et al., 2011). SVM and random forest are also very used by researchers in this stage, especially with HOG features (Ardianto et al., 2017; Ellahyan et al., 2016). To further improve the obtained results, some researchers opt for the combination of these two techniques by using multiple features, like Dense_Sift, LBP, Gabor filter (Ma and Huang, 2015). SVM classifier with HOG represents also one of the most used techniques for the recognition of textual information in road signs, and that with the use of OCR, MSER, HSV and HSI (Sumi and Arun, 2017).

This classifier is used also by Nienhuser et al. (2008), for the classification of dynamic road signs. The proposed approach is based on a tracking technique that uses Kalman Filter, and a probabilistic approach for signs' attribution to the concerned road. Some researchers (Wahyono et al., 2015) use instead, Random Forest with local spatial pattern (LSP) for the classification of displayed information on dynamic traffic signs.

In the literature review, we find that neural networks are very efficient and accurate for traffic signs recognition, and are more suitable for dealing with multiclass classification problems. In contrast, we find that SVM is more efficient in real time applications, due to its quick response time, and we find also that SVM presents, in many cases, almost the same results.

But an important question still rises concerning these approaches, which is: how to know exactly the reel performances of these methods over the existing challenges, and how to measure and prove effectively their superiority in comparison to other approaches, especially classical ones. For that, we will present, in the next section, the evaluation criteria and metrics for methods' comparison and benchmark.

3.3   Performances evaluation

The progress of road signs recognition field relies on the efforts made by researchers around the world, and every new method or technique is, certainly, proposed with the aim

to solve one or more of the problems that face the process, in order to make it more efficient. However, proving the effectiveness of the proposed methods is not always an easy task, because that depends on a certain number of criteria, which are more precisely: true positives, true negatives, false positives, false negatives and execution time. To meet these criteria, researchers adopt different evaluation approaches: real time evaluation, evaluation with simulators, with synthetic images, with acquired images or videos, or with public datasets.

Real time evaluation is necessary to test the ability of the algorithm to achieve the response speed needed in real world situations, especially when increasing vehicles' speed. Despite its importance, it is not enough to prove algorithms real performances, because this type of evaluation is, generally, done within a limited period of time, which cannot reflects the totality of the problems that recognition process could face, like light and weather changes, occlusion, etc. To overcome these limitations using simulators could help to test the response speed of the algorithms by simulating real world situations and that with taking into consideration different and adverse conditions.

Synthetic images are also used by some researchers for evaluation purposes. However, they cannot show the real performances of recognition algorithms, because this type of images does not reflect the complexity of the real road scene. Other researchers use, instead, real world images collected from the web, or acquired in real roads. Generally, the number of test images, used by these approaches, is very small, and in some cases the used images contain only traffic signs, or in other cases, signs occupy the biggest part in these images, which not represents the case in real road scene. Another alternative to this type of approaches is doing an evaluation based on videos, rather than images. This method presents many advantages, because it combines the simulation of real world situations, and the exigencies of real time evaluation.

The diversity of the evaluation approaches, adopted by researchers, makes difficult to compare the performance of the different algorithms proposed. For that, many public datasets have been created, in order to facilitate the benchmark of detection and classification algorithms. Within these datasets, we find databases that respect Vienna Convention like, GTS Benchmark (Stallkamp et al., 2011; Houben et al., 2013), BTS Dataset (Mathias et al., 2013), Swedish TSD (Larsson and Felsberg, 2011), MASTIF (Šegvić et al., 2014), Korean TSD (Rehman et al., 2017), Italian TSD (RoCoCo, 2018), Russian TSD (Shakhuroa and Konushinba, 2016), etc. We find also datasets that respect only the standards of some countries, as is the case for LISA in the USA (Møgelmose et al., 2012, 2015), CTSD and CCTSDB (Zhang et al., 2017) and TT100K (Zhu et al., 2016b) in China, etc.

However, before choosing a dataset for testing the performance of any recognition algorithm, researchers should consider many factors that could affect, considerably, the reliability of obtained results. Some of these factors are: the type of input (images/video), the number of images, the type of signs, classes' balance, the variation of images aspects, sign size, etc. Tables 1 and 2 show the details of the most popular evaluation public datasets (for detection and classification), with respect to these factors.

Table 1   Public evaluation datasets for road signs detection

Dataset	Origin	Input		Image type		N. of images	Signs type	Balanced classes	Adverse conditions	Images size (px)	Standard	
		Videos	Images	Signs only	Road scene						Vienna	Other
GTSDB	Germany		X	X	x	900	Mandatory Danger Prohibitory	No	x	16–128	x	
BTS dataset	Belgium	X	X		x	9,006	Mandatory Danger Prohibitory	No		16–913	x	
Swidesh dataset	Sweden	X	X		x	+20,000	7 types	No	Occluded-Blurred	50–	x	
MASTIF	Croatia	X	X		x	1,000–6,000	5 main categories	No			x	
LISA	US	X	X		x	+12,000	Warning Speed limit No turn	No	Sunny weather only	6–222		US

Note: x – supported criterion.

Table 2   Public evaluation datasets for road signs classification

Dataset	Signs only	Road scene	N. of images	Signs type	Balanced classes	Adverse conditions	Images size (px)
GTSRB	X		+50,000	43 classes	No	x	15*15–222*193
BTSC dataset	X		7,125	63 classes	No		
rMASTIF	X		5,828	31 classes	No		

Note: x – supported criterion.

From the two tables, we notice that there is no balance between the classes in all the datasets, which makes difficult to test the real performances of algorithms in recognising different type of signs. The tables show also that only GTS Benchmark contains images acquired in adverse conditions, while the other datasets do not guarantee this diversity. As a result, they cannot be used to test algorithms that are designed to face problems like illumination changes, weather conditions, etc. Due the absence of datasets that meet all the criteria of effective evaluation, testing recognition algorithms should be made, then, using multiple datasets, in order to determine their real performances. For example, to test a tracking system, it is necessary to use a dataset that contains not only images but also videos, or sequence of images.

Table 3   Road signs detection methods

Reference	Method	AUC	Accuracy	Precision	Recall
Yang et al. (2016)	OHTA+SVM+HOG+ RBF Kernel	97.72%			
Zhang et al. (2017)	CNN	95.65%		95.31%	90.37%
Ellahyani and El Ansari (2017b)	RGB+Mean shift+Random Forest+Log-polar	94.22%			
Ellahyan et al. (2016)	HSI+Invariant Geometric Moments	93.69%			
Huang et al. (2017)	HSI+Selective Search+Hierarchical grouping methods		92.63%		
Ardianto et al. (2017)	HSV+color segmentation+Edge Detector		91.5%		
Madani and Yusof (2016)	HSV+LVQ+XOR			100%	97.67%
Ellahyani and El Ansari (2017a)	Normalized RGB+Polygonal approximation technique			95.72%	93.12%

Concerning the evaluation metrics used in these datasets we can mention: the recall, the precision, the accuracy and the area under curve (AUC). To compare the performances of some of the presented methods in Section 3, we have found many difficulties that are relating, especially, to the existing differences in evaluation approaches adopted by each of these methods (real time, small datasets, public datasets, etc.). In order to unify the criteria of comparison, we have chosen only the methods that use GTS Benchmark. But even within these methods, we find that researchers adopt different evaluation metrics, which makes the comparison very difficult, and even impossible in some cases. Tables 3 and 4 show the metrics used by some of these methods.

Table 4   Road signs classification methods

Reference	Method	AUC	Accuracy
Ma and Huang (2015)	SVM+ Random Forest+Dense Sift+LBP+Gabor Filter		98.76%
Ardianto et al. (2017)	SVM+HOG		98%
Yang et al. (2016)	CNN		97.75%
Ellahyan et al. (2016)	Random Forest+HOG+LSS		97.43%
Ellahyani and El Ansari (2017a)	SVM+White pixels features		96.53%
Huang et al. (2017)	CNN		80.5%

Between these metrics, we find that the AUC is the most reliable evaluation metric, for testing and comparing recognition algorithms, especially for machine learning algorithms, and that according to a study conducted by Ling et al. (2003). For that, we have chosen this criterion in order to classify the performances of the presented methods.

According to this evaluation metric, we can say that among the detection algorithms presented in Table 3, the method based on OHTA space, SVM, HOG and RBF Kernel presents the best performances (in terms of AUC) with 97.72%. We find also that the majority of the methods that perform very will are essentially based on machine and deep learning algorithms. Concerning the other detection methods, they are specifically based on shape detection (circles, triangles and rectangles).

For the classification stage, we adopt the accuracy as the evaluation metric (due to the lack of information provided about the AUC in the presented works). According to this, we can say that combining SVM + random forest and multiple features gives the best results, with 98.76%. We find also that CNNs are accurate and efficient for road signs classification and that with 98.24%.

When we further analyse the results presented in the two tables, we find that there are some methods that are not designed to detect and classify all traffic signs' types, but the recognition process concerns, instead, only few types of signs, especially red and speed limit signs, as is the case of the approach adopted by Huang et al. (2017) and Ardianto et al. (2017).

Another important point is that even the fact that the majority of the presented methods perform very well in GTSR dataset, we should underline that GTSR dataset contains only true positive road signs, and does not include in contrast, false elements (that are similar to traffic signs). Thing which not helps to evaluate exactly the ability of these methods to handle and eliminate false alarms (ROIs that are not rejected during the detection stage).

Where comes the necessity of testing the recognition system on datasets that include false positives, and also the importance of testing this type of systems on its totality, as realised by Ellahyan et al. (2016) and Yang et al. (2016). This evaluation of the system on its globality consists on extracting potential road signs in the detection stage, and using then the classification module to recognise the class of each detected ROI, and eliminate false alarms that look like traffic signs.

That let us also highlight the crucial role that plays the detection module in the recognition process. In fact, the detection stage is more important than the other stages (classification and eventually tracking). This importance is due to the fact that the performance of this stage affects considerably the performance of the whole system, and that even with the presence of a performant classification and tracking system. Because if the detection module is not very performant enough, especially by detecting a big number of false positives, that will surely induce the classification and tracking modules on error, and that by classifying and tracking elements that are not road signs. Thing which will decrease, enormously, the efficiency of the whole system. The next section discusses the importance of the detection stage, and its impact on the other stages.

3.4 Detection stage impact on the other stages (classification, tracking, etc.)

In fact, the detection process is more challenging than the other stages, because it is guided by three main questions, which are 'what', 'when' and 'where'.

When dealing with traffic signs recognition, the first question that researchers should answer is that: what is the best representative feature to look for, in a road scene image, in order to find traffic signs? In fact, it is very difficult to choose discriminative features that are able to locate perfectly the potential ROIs from a specific image. We can even say that, finding and extracting discriminative features, from real world images, that describe with a high precision the road signs, is the key element that considerably affects the performance of the whole recognition system. The detection of a huge number of false alarms will, surely, induce the classification module on error and that even if this stage includes a class for false elements, because a certain number of false positives could be classified within one of the signs' classes present in this inference stage. Thing which will lead to the perturbation of the drivers, and to the lost of their concentration. Furthermore, that will be time consuming, especially when dealing with a great number of false alarms. In addition to that, if the system includes a tracking stage, this stage will lose all of its usefulness, and that by tracking false elements, instead of tracking real traffic signs.

For the 'when' question, finding the best discriminative features is not enough when dealing with traffic signs recognition. When looking for these specific features? is another question to it, the detection system, should find a convenient answer. Looking for these features in every video frame (25 frames/s on average) is both computationally expensive, and also very time consuming. For that, some researchers choose to select, instead, an interval between these acquired frames to start the detection process. In fact, the 'when' question still rises many questions for researchers in order to find the best 'trigger' that could initiate the detection process every time a traffic sign appears in the road scene.

Table 5   Impact of detection approaches on classification and tracking stages

Detection stage		Classification stage		Tracking stage
Method	*Advantages and disadvantages*	*Method*	*Impact of detection on classification stage*	*Impact of detection on tracking stage*
Colour thresholding	• Fast • Low precision	Template matching	*Response time:* should deal with many detected objects that have the same colours as traffic signs and also with damaged, occluded and rotated signs detected from the previous stage. *Precision:* faded signs and signs affected by illumination changes could not be classified.	Could not track: faded signs and signs affected by illumination changes. Could track: occluded signs.
		Machine learning	*Response time:* fast, but it takes more time to extract features relating to detected ROIs that have the same colours as traffic signs. Precision: faded signs and signs affected by illumination changes could not be also classified. However, damaged, occluded and rotated signs could be classified (but that needs many classifiers).	Could not track: faded signs and signs affected by illumination changes. Could track: occluded, damaged signs and signs affected by rotations.
		Deep learning	*Response time:* time consuming (should deal also with many detected objects that have the same colours as traffic signs), but doesn't need manual extraction of features. *Precision:* could give the same or even outperforms machine learning methods, but that depends on the learning stage and the number of representative data used, especially relating to damaged, occluded and rotated signs. Contrary to machine learning methods, it needs instead few classifiers.	
Edge-based algorithms (Hough transform, etc.)	• Time consuming • Low precision.	Template matching	*Response time:* time consuming (should deal with detected objects that have the same shape as traffic signs) and it needs features extraction for these objects. *Precision:* the classification of faded signs and signs affected by illumination changes depends on the quality of the extracted edges.	Could not track: damaged and rotated signs. Could track: faded, occluded signs and signs affected by illumination changes (when a detected sign becomes partially occluded in the next frames).
		Machine learning	*Response time:* fast, but should deal with detected objects that have the same shape as traffic signs *Precision:* faded signs and signs affected by illumination changes could be classified. However, occluded, damaged and rotated signs could not be classified. It needs many classifiers and manual extraction of features.	Could not track: damaged signs and signs affected by rotations. Could track: faded signs, signs affected by illumination changes and occluded signs (when a detected sign becomes partially occluded in the next frames).
		Deep learning	*Response time:* time consuming and it takes more time to deal with detected objects that have the same shape as traffic signs. *Precision:* could outperform machine learning, especially if the learning dataset includes many samples of faded signs and signs affected by illumination changes. It needs few classifiers and extracts features directly.	

Table 5   Impact of detection approaches on classification and tracking stages (continued)

| Detection stage | | Classification stage | | Tracking stage |
Method	Advantages and disadvantages	Method	Impact of detection on classification stage	Impact of detection on tracking stage
Machine learning (SVM/random forest)	• Fast • High precision	Template matching	Response time: should deal with damaged, occluded and rotated signs detected and it needs features extraction for these object Precision: faded signs and signs affected by illumination changes could be classified	Could not track: damaged and rotated signs. Could track: faded signs, signs affected by illumination changes and Occluded signs (when a detected sign becomes partially occluded in the next frames).
		Machine learning	Response time: fast classification (most of signs are well detected and categorised by sub-classes). Precision: could reach high precision, but it needs many classifiers and manual extraction of features.	Could track: damaged signs, faded signs, signs affected by rotations, weather and illumination changes and occluded signs (even if it was for their first apparition).
		Deep learning	Response time: time consuming in comparison to machine learning. Precision: could outperform machine learning if the learning dataset includes more representative data.	
Deep learning (CNN)	• Time consuming • Very high precision	Template matching	Response time: should deal with damaged, occluded and rotated signs detected and it needs features extraction for these object Precision: faded signs and signs affected by illumination changes could be classified	Could not track: damaged and rotated signs. Could track: faded signs, signs affected by illumination changes and occluded signs (when a detected sign becomes partially occluded in the next frames).
		Machine learning	Response time: fast classification (detected signs are categorised by sub-classes and few number of false positives). Precision: could reach high precision by classifying signs with adverse conditions detected.	Could track: damaged signs, faded signs, signs affected by rotations, weather and illumination changes and occluded signs (even if it was for their first apparition).
		Deep learning (CNN)	Response time: time consuming in comparison to machine learning. Precision: could outperform machine learning if the learning dataset contains a lot of representative data for signs in adverse conditions and it needs few classifiers & extract directly features from images.	

The last important question to ask when dealing with a detection system is: Where looking to find these discriminative features?, because it is not logical to search for the traffic signs in the whole image. Firstly, because the road scene is very complex, and includes many objects that look, almost, like traffic signs, thing which will lead to many false detections. Secondly, from a driver perspective, the objective is not to find all the traffic signs present in a road scene, but more exactly the goal is to locate traffic signs that appear on the convenient road. So extracting traffic signs that belong to other roads is considered, in this specific case, as false alarms and not as true positive detections. Finding then the best zone of search, in road scene images, is another important question that researchers should explore more to find the convenient answer.

For all these reasons, the detection stage impacts considerably the classification and the tracking stages. Table 5 shows on details the impact of the different approaches adopted in the detection stage and that in terms of precision and response time.

Studying the impact of the detection stage on the classification and tracking stages, let us conclude that, combining machine and deep learning algorithms, especially SVM, Random Forest and CNN, helps to create performant recognition systems, which combine between the high performances and the response time needed for real time applications. The approach adopted by (Yang et al., 2016), that aims to create a real-time recognition system, confirms this conclusion. In fact, the proposed method consists on using machine learning algorithms for the detection stage (SVM with HOG and RBF Kernel). For the classification they used a Deep learning approach based on CNN. When testing their classification method on GTSR Dataset, the method have reached high performances with 97.75%. But when evaluating their system on its totality, the method have reached more than 98.24%. Thing which shows the high performances of CNNs for eliminating false positives extracted from the detection stage.

Furthermore, Deep learning approaches could bring some elements of answer to the key question that still guides the detection and also the classification stage, which concerns precisely the best features to look for in a road scene image. Especially, because deep learning methods do not need a manual extraction of features, and they deal directly with images to extract the best possible features.

The next section explores the progress made by researchers, in this context, and try to answer some important questions concerning the perspectives and challenges that still face deep learning approaches.

4   Deep learning: perspectives and challenges

Technologies are there to serve humanity, and to facilitate and improve the quality of people's lives. Road sign recognition systems are ones of these technologies that aim to make car drivers' experience easier and the most important thing, to make it safer for them and also for road users in general. Despite the important difficulties and challenges that still face this field, researchers, in this domain, have reached a significant progress, reflected by the diversity of the techniques and methods proposed for road signs recognition.

As for the other road sign recognition algorithms, deep learning methods aim to find the response of the most important question that inspires all the works done in this field, which is how to ensure a perfect recognition for road signs, under adverse conditions, that operates at the same time with a high response speed. Hence, there is three important key

elements that should be taken into consideration, which are: recognition rate, adverse conditions and response speed.

4.1 Recognition rate

Concerning the recognition rate, the state-of-the-art shows that almost the majority of deep learning-based approaches realise high performances, for road signs recognition, and that by using multiple public datasets (Zang et al., 2018; Hoang et al., 2018; Qian et al., 2016; Zhu et al., 2016a; Lau et al., 2015; Qian et al., 2015).

These approaches are, especially, based on CNNs, as is the work of Hoang et al. (2018). The proposed method includes ROIs extraction based on HSV space colour for road signs detection, and the use of a deep neural network for identifying the meaning of detected traffic signs. The method reaches the best accuracy on GTSRB with 99.99%.

For Zang et al. (2018), they propose a quaternion convolutional neural network (QCNN)-based method. To detect traffic signs, they use faster R-CNN. They utilise after that mean shift to track the detected signs. Then QCNNs are employed to extract the spatial and temporal features. For the final recognition, a third QCNN is used. To test the performance of the tracking system, they used a video-based evaluation dataset (MASTIF). The detection and classification rates are 99.31% and 99.15% respectively.

Another method that uses also CNN is proposed by Qian et al. (2016). This method utilises CNN for discriminative feature representation and max pooling positions for improving classification performance and speed. The accuracy obtained by this method is 98.86% on GTSRB. The same authors (Qian et al., 2015) propose a new method based on a multi-task CNN for road signs detection and classification (including digits, English letters and Chinese characters). The method is based on proposal regions to find traffic signs candidates, and that by using RGB space thresholding and edge detection. The second stage includes the use of a multi-task CNN for classification. The method achieves also a high accuracy on GTSRB with 95%.

In addition to the high accuracy achieved by these methods, deep learning approaches could even outperform human performances, for road signs recognition, and that according to a benchmark (man vs. computer) realised by Stallkamp et al. (2012).

4.2 Adverse conditions

When addressing the second challenge of adverse conditions, for deep learning approaches, we find that the existing public datasets do not reflect exactly the real world conditions and scenarios that really face recognition algorithms, thing which makes us wonder about the real performances of the methods that use these datasets for evaluation. To address the limitations of existing public datasets, and in order to test algorithms' performances under real world challenging conditions, a novel video dataset were introduced 'The Challenging Unreal and Real Environments for Traffic Sign Detection' (Temel et al., 2017). This dataset (CURE-TSD) was used to host the first edition of Video and Image Processing (VIP) CUP in 2017, denoted as 'Traffic Sign Detection under Challenging Conditions' (Temel and AlRegib, 2018). This competition included 19 teams from 10 countries.

For the video dataset used, it contains processed versions of captured and synthesised traffic videos with challenging conditions that vary from mild to severe: rain, snow, haze,

brightness, darkness, shadow, blur, decolourisation, dirty lens, noise, etc. The dataset includes 3,978 sequences for training, which represent 70%, and 1,755 sequences for testing algorithms' performances 30%.

Concerning the results of the final competition, held at the 2017 IEEE International Conference on Image Processing ICIP in Beijing, the winner team (Neurons) used separate CNNs to ensure the localisation and the recognition of traffic signs, and that after performing a pre-processing operation to eliminate the effect of adverse conditions. For the second team (Polytechnic), they utilised also CNNs, one for region proposal, and the second for road sign classification. The third team (Markovians) trained a recurrent CNN to identify the adverse conditions and detect traffic signs with a faster region-based CNN architecture. They also used a tracking system, based on Kalman approach to track signs over the frames.

We find that, the three finalist teams have used methods that rely on deep learning, especially on CNNs. These results let us conclude that CNNs perform well under real world adverse conditions, and that in comparison with other types of methods. We can say also that, this type of approaches is very performant for realising various tasks, including pre-processing, localisation and classification under challenging conditions. In contrast, we find that the best performing method at this competition achieves relatively low results in terms of precision (0.550) and recall (0.320).

The obtained results show then that the high performances that CNNs algorithms reach, in the most available public datasets, decrease considerably when testing the same algorithms in a dataset like CURE-TSD that includes more real world challenging scenarios. Furthermore, CNNs still face recognition problems relating, specifically, to rotation, image orientation and change in pose, due to the reduction of spatial information data when using max pooling layer. To overcome these difficulties, Tugirimana et al. (2018) take another path by developing a new type of neural networks, which is based on modules or capsules rather than pooling to recognise rotated and distorted images.

4.3   Response speed

Concerning the third key element (real time), deep learning systems should operate well, not only under challenging real world conditions, but they should also have a quick response. Considering this important factor, we find that deep learning approaches could not reach high performances, in comparison with other methods, especially machine learning algorithms. Certainly, deep learning approaches present many advantages, especially because they deal directly with images, instead of using features, and also because they need less classifiers. However, they require, on the contrary, a lot of GPU's, and also a lot of data to train the classifiers, which means a long time for algorithms training, and also an enormous effort to get balanced classes.

In addition to that, the real time recognition of road signs, using CNNs, still presents a challenge for researchers in order to improve the response time, because their processing time is quite long, in comparison with other methods like SVM for example. In fact, we find that SVM classifier combines between the high accuracy, and the speed response needed in real time recognition of road signs, and that by using small datasets, that take also a short time for the training stage.

As a conclusion, we can say that although the high performances of deep learning algorithms under adverse conditions (especially CNNs), they need, in contrast, high hardware requirements, thing which makes it difficult to use this kind of methods in

advanced driver-assistance systems (ADAS). For that reason, the real time recognition of road signs, using CNNs, should necessarily involve optimisation techniques (hardware and algorithmic optimisation) in order to improve the running time of this type of systems.

4.4 Optimisation techniques

For deep learning approaches, the optimisation involves modifying processing and networks architecture during the three following stages: training and both real time detection and classification.

4.4.1 Hardware optimisation

Concerning hardware optimisation, during the training stage, we find that the computational power of the hardware (GPU, APU, etc.) influences considerably the efficiency of deep learning systems. In fact, by taking into consideration these four important components: storage devices, main memory, CPU and GPU, the study conducted by Li et al. (2017) shows that hardware configuration have a significant impact on deep learning performances.

The study finds also that the memory does not affect dramatically the running time. In contrast, it demonstrates that using double and triple GPUs improve the execution time by 44% and 59% respectively. For storage devices, it highlights that HDD array and SSD array improve the performances of some models. The study mentions also that CPU frequency has a significant impact on running time. For the detection and classification of road signs, in real time, hardware optimisation is also needed, because such systems should process high quality images and dozens of frames per second (Alhamali et al., 2015).

Although the fact that hardware optimisation improves significantly the performances of deep learning systems, but it limits, in contrast, their use in low resources systems, such as portables, automatic cars, etc. It limits also their generalisation in ADAS systems, because this type of devices does not have the capabilities to perform deep inferences for real time applications (Li et al., 2017).

Hence, deploying deep learning techniques in such type of systems needs not only hardware optimisation, but it needs also the modification of networks' architecture (algorithmic optimisation) (Verhelst and Moons, 2017). Knowing that a promising progress is already done in this context of hardware optimisation by the introduction of new powerful and complex SoC for self-driving cars, as NVIDIA Xavier chipset released in 2018. This new chip has almost 9 billion transistors, and can perform over 30 trillion operations per second, and that by using only 30 watts of power.

4.4.2 Software optimisation

The algorithmic optimisation concerns also the three stages (training, detection and classification). For the training process, it is obvious that the performances of any deep learning system depend, especially, on the quality of this stage, which is in turn tightly related to the number of representative data used to train the system. However, labelling thousands of hours of videos to extract real world instances of road signs is a hard task to realise, and also it is very time consuming (takes many days or even weeks).

To overcome this problem, data augmentation is one of the classical techniques used by researchers to increase the amount of training data (Wong et al., 2016) by creating additional training samples, especially to improve performances in imbalanced class problems. There are two main approaches for generating new samples, which are applying transformations in feature or in data space. For the augmentation in data-space, it consists on creating transformations that maintain label information, with the validation of a human observer. Concerning the augmentation in feature space, it is used when it is not easy to validate these transformations.

Another performant method, used during this training stage, is the Iterative Search and Learn approach proposed by Overett and Wang (2017). In order to reduce the time needed for data annotation, they used a detector (cascade of boosted classifiers) to find positive road signs instances. This hand labelled data is further added to a growing dataset to retrain the detector from the start (until achieving a sufficient precision). In order to reduce the manual labelling required, they use auxiliary neural network classifiers to pre-sort true and false positive instances. They used then sorting and learning method in order to sort true positive road signs into sub-categories. The classification accuracy of this approach is 99%. However, the absence of ground truth data makes difficult to determine the overall recall of the detection stage.

For real time detection and classification of road signs, several approaches could be used by researchers to optimise deep learning algorithms, especially compression and acceleration approaches. The main objective of these approaches is to reduce storage and computational cost. There are four main compression approaches, which are parameter pruning and sharing, low-rank factorisation, transferred/compact convolutional filters and knowledge distillation (Cheng et al., 2017).

Concerning parameters pruning and sharing methods, they consist on reducing deep learning parameters that are redundant and not crucial (do not affect systems' performances). For low-rank factorisation, this method consists on estimating the informative parameters of deep learning systems. The third method aims to reduce the storage and the computational cost by designing special structural convolutional filters. Concerning the fourth method (knowledge distillation), its main objective is to train a small network that is able to achieve almost the same result obtained by a larger one (teachers/students networks).

These four methods are independent and complete each other. For deep object recognition (which includes road signs recognition) combining several compression and acceleration methods gives best results, especially, by using low-rank factorisation for convolutional layers, and parameter pruning and sharing techniques for fully connected layers. This combination helps to speed up considerably the running time, without dramatically lowering systems' accuracy.

Some researchers take a different path for algorithmic optimisation, by proposing a new approach (Lee and Lee, 2016). In this work, Instead of reducing deep learning parameters, as is the case of compression techniques, the new proposed approach consists on creating an intermediate classification output layer. The objective of the middle layer is to reduce execution time by classifying the simple images (that have a confident rank of 90% or more) as soon as possible without completing the whole process. For complex images, the inference does not stop until the final classification output layer.

From what we have presented in this section, we can say that optimisation methods, especially, algorithmic optimisation opens the way for developing performant systems,

based on deep learning architecture, that are able to realise a real time recognition of road signs, that ensures both high accuracy and precision, and also a very quick response time.

5 Conclusions

The field of road signs recognition knows a continuous progress in different domains, and a diversity of methods is proposed by researchers to improve the effectiveness of these systems. These techniques can be categorised in two main approaches (classical and machine learning). Machine learning approaches give more accurate and efficient results, especially deep learning-based methods. This type of methods gains in terms of recognition rate and challenging real world conditions, but it fails, considerably, to meet the third key element, which is real time response, because they are time consuming, in comparison to other methods, thing that makes their use in real time systems still a difficult challenge to handle.

To speed up deep learning systems, for training and real time recognition, hardware optimisation is no longer enough. For that, algorithmic optimisation becomes a necessity to ensure high accuracy, especially, when using low resources applications, such as autonomous cars, etc. and that will also help to generalise their use in ADAS systems. This optimisation involves necessarily the combination of several techniques to ensure getting satisfactory results.

Acknowledgements

References

Agrawal, S. and Chaurasiya, R.K. (2017) 'Ensemble of SVM for accurate traffic sign detection and recognition', *ICGSP '17 Proceedings of the International Conference on Graphics and Signal Processing*, Singapore, 24–27 June, pp.10–15.

Akatsuka, H. and Imai, S. (1987) 'Road signposts recognition system', *The International Conference on SAE Vehicle Highway Infrastructure: Safety Compatibility*, pp.189–196.

Alhamali, A., Salha, N., Morcel, R., Ezzeddine, M., Hamdan, O., Akkary, H. and Hajj, H. (2015) 'FPGA-accelerated Hadoop cluster for deep learning computations', *2015 IEEE International Conference on Data Mining Workshop (ICDMW)*, Atlantic City, NJ, USA.

Ardianto, S., Chen, C-J. and Hang, H-M. (2017) 'Real-time traffic sign recognition using color segmentation and SVM', *International Conference on Systems, Signals and Image Processing (IWSSIP)*.

Besserer, B., Estable, S., Ulmer, B. and Reichardt, D. (1994) 'Shape classification for traffic sign recognition', *Control Engineering Practice*, Vol. 2, No. 1, pp.483–488.

Cheng, Y., Wang, D., Zhou, P. and Zhang, T. (2017) 'A survey of model compression and acceleration for deep neural networks', *IEEE Signal Processing Magazine, Special Issue on Deep Learning for Image Understanding*, January 2018, pp.126–136.

Di, Z., Junqi, Z., Dongdong, Z., Bao, M., Cheng, J. and Tang, K. (2016) 'Traffic sign detection based on cascaded convolutional neural networks', *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), 2016 17th IEEE/ACIS International Conference*, Shanghai, China.

Ellahyan, A., El Ansari, M. and ElJaafari, I. (2016) 'Traffic sign detection and recognition based on random forests', *Applied Soft Computing*, September, Vol. 46, pp.805–815.

Ellahyani, A. and El Ansari, M. (2017a) 'A new designed descriptor for road sign recognition', *International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, Fez, Morocco.

Ellahyani, A. and El Ansari, M. (2017b) 'Mean shift and log-polar transform for road sign detection', *Multimedia Tools and Applications*, Vol. 76, No. 22, pp.24495–24513.

Gao, X., Hong, K., Passmore, P., Podladchikova, L. and Shaposhnikov, D. (2008) 'Colour vision model-based approach for segmentation of traffic signs', *EURASIP J. Image and Video Processing*, December 2007, 2008:386705.

Gokul, S., Suresh Kumar, S. and Giriprasad, S. (2015) 'Survey – an exploration of various techniques for sign detection in traffic panels', *ARPN Journal of Engineering and Applied Sciences*, Vol. 10, No. 9, pp.4018–4022.

Gomez-Moreno, H., Maldonado-Bascon, S., Gil-Jimenez, P. and Lafuente-Arroyo, S. (2010) 'Goal evaluation of segmentation algorithms for traffic sign recognition', *IEEE Transactions on Intelligent Transportation Systems*, Vol. 11, No. 4, pp.917–930.

Hoang, V., Le, M., Tran, T. and Pham, V. (2018) 'Improving traffic signs recognition based region proposal and deep neural networks', *ACIIDS 2018: Intelligent Information and Database Systems*, pp.604–613.

Houben, S., Stallkamp, J., Salmen, J., Schlipsing, M. and Igel, C. (2013) 'Detection of traffic signs in real-world images: the German traffic sign detection benchmark', *The 2013 International Joint Conference Neural Networks (IJCNN)*, Dallas, TX, USA.

Huang, S-C., Lin, H-Y. and Chang, C-C. (2017) 'An in-car camera system for traffic sign detection and recognition', *Fuzzy Systems Association and 9th International Conference on Soft Computing and Intelligent Systems (IFSA-SCIS)*, Otsu, Japan.

Lab RoCoCo (2018) *DITS – Data set of Italian Traffic Signs* [online] http://users.diag.uniroma1.it/bloisi/ds/dits.html (accessed 09/13/2019).

Larsson, F. and Felsberg, M. (2011) 'Using fourier descriptors and spatial models for traffic sign recognition', in Heyden, A. and Kahl, F. (Eds.): *Image Analysis. SCIA 2011. Lecture Notes in Computer Science*, Vol. 6688, pp.238–249, Springer, Berlin, Heidelberg.

Lau, M.M., Lim, K.H. and Gopalai, A.A. (2015) 'Malaysia traffic sign recognition with convolutional neural network', *2015 IEEE International Conference on Digital Signal Processing (DSP)*, Singapore.

Lee, H. and Lee, B-T. (2016) 'Selective inference for accelerating deep learning-based image classification', *International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju, South Korea.

Li, J., Zhang, C., Cao, Q., Qi, C., Huang, J. and Xie, C. (2017) 'An experimental study on deep learning based on different hardware configurations', *International Conference on Networking, Architecture, and Storage (NAS)*. Shenzhen, China.

Ling, C.X., Huang, J. and Zhang, H. (2003) 'AUC: a better measure than accuracy in comparing learning algorithms', in Xiang, Y. and Chaib-draa, B. (Eds.): *Advances in Artificial Intelligence. Canadian AI 2003. Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence)*, Vol. 2671, Springer, Berlin, Heidelberg.

Ma, Y. and Huang, L. (2015) 'Hierarchical traffic sign recognition based on multi-feature and multi-classifier fusion', *First International Conference on Information Science and Electronic Technology (ISET 2015)*.

Madani, A. and Yusof, R. (2016) 'Traffic sign detection based on simple XOR and discriminative features', *Jurnal Teknologi (Sciences & Engineering)*, Vol. 78, No. 6-2, pp.97–102.

MAP (2018b) *Road Traffic Fatalities Down 2.62% in 2017: State Secretary*, Portail du Maroc [online] http://www.maroc.ma/en/news/road-traffic-fatalities-down-262-2017-state-secretary (accessed 26 March 2018).

MAP, Moroccan News Agency (2018a) *Government Aims to Reduce by 50% Number of Fatalities in Traffic Accidents by 2026*, Portail du Maroc [online] http://www.maroc.ma/en/news/ government-aims-reduce-50-number-fatalities-traffic-accidents-2026 (accessed 26 March 2018).

Mathias, M., Timofte, R., Benenson, R. and Gool, L.V. (2013) 'Traffic sign recognition – how far are we from the solution?', *The 2013 International Joint Conference Neural Networks (IJCNN),* Dallas, TX, USA.

Møgelmose, A., Liu, D. and Trivedi, M.M. (2014) 'Traffic sign detection for U.S. roads: remaining challenges and a case for tracking', *2014 IEEE 17th International Conference Intelligent Transportation Systems (ITSC)*, Qingdao, China.

Møgelmose, A., Liu, D. and Trivedi, M.M. (2015) 'Detection of U.S. traffic signs', *IEEE Transactions on Intelligent Transportation Systems*, Vol. 16, No. 6, pp.3116–3125.

Møgelmose, A., Trivedi, M.M. and Moeslund, T.B. (2012) 'Vision-based traffic sign detection and analysis for intelligent driver assistance systems: perspectives and survey', *IEEE Transactions on Intelligent Transportation Systems*, Vol. 13, No. 4, pp.1484–1497.

Nienhuser, D., Gumpp, T. and Marius Zollner, J. (2008) 'Recognition and attribution of variable message signs and lanes', *2008 IEEE Intelligent Vehicles Symposium*, Eindhoven, Netherlands.

Overett, G. and Wang, W.W. (2017) 'Iterative search & learn for sign detection in large datasets', *IEEE Intelligent Vehicles Symposium (IV)*, Los Angeles, CA, USA.

Qian, R., Yue, Y., Coenen, F. and Zhang, B. (2016) 'Traffic sign recognition with convolutional neural network based on max pooling positions', *12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, Changsha, China.

Qian, R., Zhang, B., Yue, Y., Wang, Z. and Coenen, F. (2015) 'Robust Chinese traffic sign detection and recognition with deep convolutional neural network', *2015 11th International Conference on Natural Computation (ICNC)*, Zhangjiajie, China.

RACC and ADAC (2011) *Sistemas de Reconocimiento de Señales de Trafico en Turismos*, RACC, Cataluña.

Rehman, Y., Riaz, I., Fan, X. and Shin, H. (2017) 'D-patches: effective traffic sign detection with occlusion handling', *IET Computer Vision*, Vol. 11, No. 5, pp.368–377.

Romadi, M., Faizi, R. and Chiheb, R. (2017) 'A shape-based approach for detecting and recognizing traffic signs in a video stream', *4th International Conference on Control, Decision and Information Technologies (CoDIT)*, Barcelona, Spain.

Romdhane, N., Mliki, H. and Hammami, M. (2016) 'An improved traffic signs recognition and tracking method for driver assistance system', *15th International Conference on Computer and Information Science (ICIS), 2016 IEEE/ACIS*, Okayama, Japan.

Safat, B. et al. (2015a) 'An automatic traffic sign detection and recognition system based on colour segmentation, shape matching and SVM', *Mathematical Problems in Engineering*, November, Vol. 2015, No. 2015, pp.1–11.

Safat, B., Mohammad, A., Aini, H. and Salina, A. (2015b) 'Comparative survey on traffic sign detection and recognition: a review', *Przegląd Elektrotechniczny*, Vol. 91, NR. 12/2015.

Šegvić, S., Brkić, K., Kalafatić, Z. and Pinz, A. (2014) 'Exploiting temporal and spatial constraints in traffic sign detection from a moving vehicle', *Machine Vision and Applications*, Vol. 25, No. 3, pp.649–665.

Shakhuroa, V.I. and Konushinba, A.S. (2016) 'Russian traffic sign images dataset', *Computer Optics*, Vol. 40, No. 12, pp.294–300.

Stallkamp, J., Schlipsing, M., Salmen, J. and Igel, C. (2011) 'The German traffic sign recognition benchmark: a multi-class classification competition', *The 2011 International Joint Conference on Neural Networks (IJCNN)*, IEEE.

Stallkamp, J., Schlipsing, M., Salmen, J. and Igel, C. (2012) 'Man vs. computer: benchmarking machine learning algorithms for traffic sign recognition', *Neural Networks*, Vol. 32, pp.323–332, Elsevier.

Sumi, K. and Arun, K. (2017) 'Detection and recognition of road traffic signs: a survey', *International Journal of Computer Applications*, February, Vol. 160, No. 3, pp.1–5.

Surinwarangkoon, T., Nitsuwat, S. and Moore, E. (2013) 'Traffic sign recognition system for roadside images in poor condition', *International Journal of Machine Learning and Computing*, Vol. 3, pp.121–126, DOI: 10.7763/IJMLC.2013.V3.285.

Temel, D. and AlRegib, G. (2018) 'Traffic signs in the wild: highlights from the IEEE Video and Image Processing Cup 2017 Student Competition [SP Competitions]', *IEEE Signal Processing Magazine*, Vol. 35, No. 2, pp.154–161.

Temel, D., Kwon, G., Prabhushankar, M. and AlRegib, G. (2017) 'CURE-TSR: challenging unreal and real environments for traffic sign recognition', *31st Conference on Neural Information Processing Systems (NIPS), Machine Learning for Intelligent Transportation Systems Workshop*, Long Beach, CA, USA.

Tugirimana, J.d.D., Rulinda, J., Nzaramba, A. and Tresor, K.L. (2018) 'Road traffic sign detection and classification using capsule network', *International Journal of Scientific & Engineering Research*, Augsut, Vol. 9, No. 8.

United Nations Economic Commission for Europe (1995) *Convention on Road Signs and Signals: Done at Vienna on 8 November 1968, Amendment 1 (Incorporating the Amendments to the Convention which Entered into Force on 30 November 1995)*.

Verhelst, M. and Moons, B. (2017) 'Embedded deep neural network processing: algorithmic and processor techniques bring deep learning to IoT and edge devices', *IEEE Solid-State Circuits Magazine*, Vol. 9, No. 14.

Wahyono, Filonenko, A. and Jo, K-H. (2015) 'Automatic LED text recognition method on electronic road sign using local spatial pattern and random forest classifier', *2015 IEEE Intelligent Vehicles Symposium (IV)*, Seoul, Korea.

Wahyono, Park, H-S. and Kang-Hyun, J. (2013) 'Discontinuous character recognition method in variable message signs for intelligent transportation systems', *International Journal of Machine Learning and Computing*, Vol. 3, No. 6.

Wong, S.C., Gatt, A., Stamatescu, V. and McDonnel, M.D. (2016) 'Understanding data augmentation for classification: when to warp?', *International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, Gold Coast, QLD, Australia.

World Health Organization (WHO) (2015) *Global Status Report on Road Safety 2015*, WHO, Geneva.

Yang, Y., Luo, H., Xu, H. and Wu, F. (2016) 'Towards real-time traffic sign detection and classification', *IEEE Transactions on Intelligent Transportation Systems*, Vol. 17, No. 7.

Zang, D., Zhihua, W., Bao, M., Cheng, J., Zhang, D., Tang, K. and Li, X. (2018) 'Deep learning-based traffic sign recognition for unmanned autonomous vehicles', *Systems and Control Engineering*, IMechE.

Zhang, J., Huang, M., Jin, X. and Li, X. (2017) 'A real-time Chinese traffic sign detection algorithm based on modified YOLOv2', *Algorithms*, Vol. 10, No. 4, p.127.

Zhu, Y., Zhang, C., Zhou, D., Wang, X., Bai, X. and Liu, W. (2016a) 'Traffic sign detection and recognition using fully convolutional network guided proposals', *Neurocomputing*, 19 November, Vol. 214, pp.758–766.

Zhu, Z., Liang, D., Zhang, S., Huang, X., Li, B. and Hu, S. (2016b) 'Traffic-sign detection and classification in the wild', *Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference*, IEEE.