Automatic road sign detection and recognition based on neural network

Redouan Lahmyed (lahmyed.redouan@gmail.com)
Ibn Zohr University Faculty of Science Agadir: Universite Ibn Zohr Faculte des Sciences Agadir
https://orcid.org/0000-0003-1023-2502

Mohamed El Ansari
Ibn Zohr University Faculty of Science Agadir: Universite Ibn Zohr Faculte des Sciences Agadir

Zakaria Kerkaou
Ibn Zohr University Faculty of Science Agadir: Universite Ibn Zohr Faculte des Sciences Agadir

Research Article

Keywords: Traffic sign detection, Traffic sign recognition, Color segmentation, Artificial neural networks (ANN), Support vector machines (SVMs), Histogram of dominant silhouette orientation, Gradient local binary patterns (GLBP), Local self-similarity (LSS)

Posted Date: June 4th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-408446/v1

Version of Record: A version of this preprint was published at Soft Computing on January 12th, 2022. See the published version at https://doi.org/10.1007/s00500-021-06726-w.

Automatic road sign detection and recognition based on neural network

Redouan Lahmyed* · **Mohamed El Ansari** · **Zakaria Kerkaou**

Received: date / Accepted: date

Abstract Road sign detection and recognition is an integral part of intelligent transportation systems (ITS). It increases protection by reminding the driver of the current condition of the route, such as notices, bans, limitations and other valuable driving information. This paper describes a novel system for automatic detection and recognition of road signs, which is achieved in two main steps. First, the initial image is pre-processed using DBSCAN clustering algorithm. The clustering is performed based on color information, and the generated clusters are segmented using Artificial neural networks (ANN) classifier. The resulting ROIs are then carried out based on their aspect ratio and size to retain only significant ones. Then, a shape-based classification is performed using ANN as classifier and HDSO as feature to detect the circular, rectangular and triangular shapes. Second, a hybrid feature is defined to recognize the ROIs detected from the first step. It involves a combination of the so-called GLBP-Color which is an extension of the classical gradient local binary patterns (GLPB) feature to the RGB color space and the local self-similarity (LSS) feature. ANN, Adaboost and support vector machine (SVM) have been tested with the introduced hybrid feature and the first one is selected as it outperforms the other two. The proposed method has been tested in outdoor scenes, using a collection of common databasets, well known in the traffic sign community (GTSRB, GTSDB and STS). The results demonstrate the effectiveness of our method when compared to recent state-of-the-art methods.

Keywords Traffic sign detection · Traffic sign recognition · Color segmentation · Artificial neural networks (ANN) · Support vector machines (SVMs) · Histogram of dominant silhouette orientation · Gradient local binary patterns (GLBP) · Local self-similarity (LSS).

1 Introduction

Advanced Driver Assistance (ADAS) systems are designed to improve vehicle safety and driving comfort. One of the most significant difficulties facing ADAS is the perception of the landscape and guidance of the vehicles in actual outdoor scenes including pedestrian detection [42, 31, 11, 30], vehicle environment perception [29, 50, 28, 13, 12], traffic sign detection [15, 14, 16, 36], and so on.

This research work was supported by the National Center for Scientific and technical Research (CNRST), research grant No: 20UIZ2015.

Redouan Lahmyed (Corresponding author)
LabSIV, Department of Computer Science, Faculty of Science, Ibn Zohr University, BP 8106, 80000 Agadir, Morocco.
Tel.: +212-622483298
E-mail: lahmyed.redouan@yahoo.fr or lahmyed.redouan@gmail.com

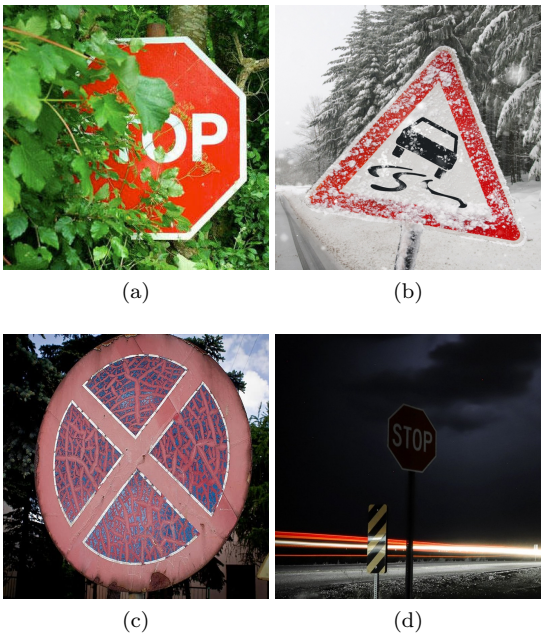
Mohamed El Ansari
Informatics and Applications Laboratory, Department of Computer Science, Faculty of Science, My Ismail University Meknès, Morocco.
LabSIV, Department of Computer Science, Faculty of Science, Ibn Zohr University, BP 8106, 80000 Agadir, Morocco.
E-mail: m.elansari@uiz.ac.ma or melansari@gmail.com

Zakaria Kerkaou
LabSIV, Department of Computer Science, Faculty of Science, Ibn Zohr University, BP 8106, 80000 Agadir, Morocco.
E-mail: kerkaou.zakaria@gmail.com

Human driving is an activity that is almost exclusively dependent on visual knowledge, and one of the tasks involved in good driving is to recognize road signs. Otherwise, it can pose a threat of people's lives due to lack of concentration or ignorance. Road signs offer updates on the existing status of the route, limits, bans, alarms, and other important navigation information.

Over the last two decades, the area of road sign detection and recognition systems has attracted substantial research interest. A number of systems have been proposed and implemented not only for ADAS, but also for other real-world applications. We mention here automated driving, urban scene understanding, and sign monitoring for maintenance. For such applications, accuracy and fast response time are highly significant metrics. However, in certain realistic situations, the identification of traffic signs is challenging, if not difficult. Some of those situations are illustrated in Fig. 1 and listed below:

- Obstacles e.g. trees, cars, and people may affect the identification of traffic signs (Fig. 1 (a)).
- Weather conditions such as snow, rain, and fog and air pollution, make the detection and recognition phases very complex (Fig. 1 (b)).
- Color fading: The color of the sign fades with time as a result of long exposure to sunlight, and the reaction of paint to air (see Fig. 1 (c)).
- Changes in lighting conditions at various periods (day and night)(Fig. 1 (d)).



(a) (b) (c) (d)

Fig. 1: Examples of challenges facing road sign recognition systems.

In this paper, a two-stage traffic sign detection and recognition approach is presented. The first stage consists of detecting the traffic signs from the input images, which is achieved in two sub-steps. The first sub-step segments the images to extract ROIs based on DBSCAN clustering and ANN. The DBSCAN clustering algorithm used to partition the initial image into a set of connected components based on color information. Then, the segmentation stage is carried out by the ANN classifier. The second sub-step verifies if the ROIs provided by the previous sub-step represent traffic signs or not by performing classification on the basis of the HDSO descriptor, which is inspired by the silhouette pattern of road sign, together with ANN classifier. The second stage is performed by developing an extension of the GLBP feature to RGB color images that we name GLBP-Color. It is combined with the LSS to define the hybrid feature that we propose to adopt for the traffic recognition method.

The structure of this paper is organized as follows. Section 2 presents state-of-the-art road sign detection and recognition. The proposed traffic sign detection and recognition is described precisely in Section 3. Experimental results to assess the performance of the proposed approach are shown in Sec. 4. Section 5 concludes the paper and present an outlook on further possible improvements.

2 Related work

Several different approaches to traffic sign detection and recognition have been suggested in the literature. Thus, the result of the research obtained differs from one group to another. Comparing and determining the best among these systems require more effort since they are based on non-availability of standard dataset, which makes the result less reliable. Among the dataset used in the field, there is German Traffic Signs Dataset (GTS) [49], Swedish Traffic Signs Dataset (STSD) [33], Stereopolis Dataset [3], LISA Dataset [40] and so on. Each dataset is characterized by some properties such as the number of classes, purpose, and the number of images.

Regarding the purpose property of the dataset, the traffic sign algorithms could be divided into three major categories of methods: (1) methods aimed to detect the road sign in the image [15, 34, 43, 6], (2) methods intended to recognize the sign class [45, 59, 39], and (3) methods designed to both detect and recognize the traffic sign simultaneously [16, 60, 56].

In the first category, different road sign detection approaches have been introduced where two directions are defined. Here, we refer to color-based and shape-based methods. Some relevant works in the first direction propose to carry out the color segmentation in various color spaces including RGB (Red Green Blue) [43], HSV (Hue, Saturation, Value) [34], YUV (Luma, Blue projection and Red projection components) [39], and L-a-b (Lightness axis, a-axis "green to red", b-axis "blue to yellow") [34]. Li et al. [43] adopt the color enhancement method to segment red, yellow and blue colors to detect road signs using the RGB color space. The same space has been also used in the work [4]. The authors have employed the difference between red and blue, and the difference between red and green channels to build two stable features in road sign detection. Besides RGB space, other approaches choose to employ other color spaces. For example, Both L-a-b and HSI systems are considered in [34]. They are utilized to extract candidate blobs for chromatic signs. Miura et al. [39] performed their system in YUV space to detect blue rectangular signs.

In the second direction, some works suggest utilizing geometric information to identify and detect traffic signs. Typically, these shape-based techniques are used either directly on road scene images or as a second phase after color segmentation. Bascón and Rodríguez [2] proposed a function of the angle defined as the distance from the blob center to its edge to classify the blobs as squares, triangles, or circles. In [15], a method for road sign detection based on mean shift clustering algorithm, random forests classifier, and log-polar transform technique is described. Dariu M. Gavrilla et al. [22] employed the techniques of Distance Transform (DT) and Template Matching (TM) to localize triangular and circular signs.

In the second category, numerous published studies have taken road sign recognition as the main problem, where different features and classifiers have been exploited including HOG, LBP, Haar-like wavelet, local self-similarity (LSS), Gabor filters, SVM and ANN classifiers. For instance, the authors in [59] utilized features combination of various sized HOG together with Fisher Linear Discriminant feature space reduction algorithm to label the road sign according to the information included in its pictogram. For the same purpose, Salti et al. [45] suggested using the HOG features with SVMs classifier in the classification phase. A normalized correlation-based pattern matching using a road sign dataset to determine the content of the traffic signs is proposed by Miura in [39]. However, these methods tend to perform poorly in real scenarios, since they are often applied after detecting and localizing the traffic signs.

For the sake of alleviating this problem, some relevant works suggested combining the two categories into one (The third category) to design systems for detecting and recognizing traffic sign in urban scenarios. On this basic concept, various approaches have been proposed such as the one in [16] that applies invariant geometric moments to classify shapes and HSI-HOG combined with LSS to recognize the traffic signs. The same authors proposed another system [14], in which the Distance to borders (DtBs), HOG, LSS and random forests classifier are used. In the same context, a method for road sign images is proposed using CNN in [56], However, a very high cost of computation (time and hardware) is required. In [60], the authors proposed two convolutional neural networks (CNNs) system in this field. The first one is used for the detection alone, while the second one is utilized for simultaneous detection and classification purposes. The proposed system shares the same difficulties that face the one published in [56].

In spite of the achievements obtained, it is an unfortunate duty to report that most of the above-motioned systems are inherently restricted by non-accurately predicted responses under certain circumstances such as weather conditions, disorientation of the sign, and different illumination levels. We alternatively present in this paper a novel method for road traffic detection and recognition.

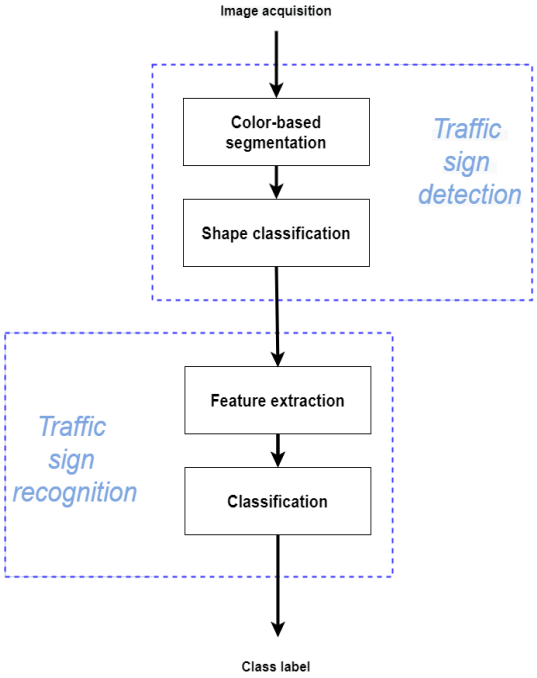


Fig. 2: Algorithm scheme.

The proposed method is an amelioration of the methods presented recently in Refs [16, 15, 18]. It is developed based on machine learning techniques to carry out the segmentation, and a silhouette pattern of road sign descriptor for shape classification. It gives better results compared to the ones proposed in [16, 15]. Moreover, the recognition phase is performed on the basis of integrating the color information into the GLBP features by using the RGB components to compute the descriptor instead of gray-scale images. The computed GLBP-Color is combined with LSS features to create a novel descriptor. These features were provided to the ANN classifier to recognize the traffic sign.

3 Proposed method

In this section, a novel approach for detecting and recognizing traffic signs is presented. As schematized in the Figure. 2, the proposed approach is achieved in two main stages. In the first stage, we aim to detect traffic signs by employing color, shape cues and machine learning techniques. The second one also uses the color information along with texture, gradient and internal geometric layout of local self-similarities information for features computation to identify the information included in the provided traffic signs by the first step. Here, we detail each of the aforementioned steps of the proposed method.

3.1 Detection

The first stage of the proposed system consists of two sub-steps: (1) Segmentation where the locations of possible traffic signs (ROIs) in natural scene images are determined and (2) shape classification, where tests are performed to verify the presence of road signs in the generated ROIs. The details of each of these sub-steps are given in this section.

3.1.1 Segmentation

Despite the fact that the road signs are distinct from each other, there are several similar properties for signs under one target category. For instance, triangular red borders characterize danger signs, mandatory signs are known by white arrows and blue backgrounds, derestriction signs have backgrounds with white color, prohibitory signs have circular borders colorized by red. Therefore, color

segmentation procedure is adopted since it aims at detecting white, red and blue colors. It is only used for extracting the ROIs (candidate traffic signs) rather than performing road sign detection due to many factors such as the presence of some objects with the same color as traffic signs in the road scene and the change of weather that may affect the color segmentation process. The color-based segmentation is done by performing machine learning techniques on the input image.

First, the image pixels are clustered into a set of groups. The clustering procedure is performed using the density-based spatial clustering of applications with noise (DBSCAN) algorithm [19] that groups together pixels with many nearby neighbors and considers pixels whose nearest neighbors are too far away as outliers. It is one of the most efficient density-based clustering techniques, which can withstand noise, densities and shapes as well. Its functionality entirely relies on two parameters: ϵ (epsilon) the radius from a corresponding pixel (pix_{corre}), which includes neighbouring pixels, and Min_{pxls} is the minimum number of pixels needed to form a cluster.

The main concept of the DBSCAN algorithm is based on the idea of creating a cluster if and only if a pixel (pix) has nearly enough neighbors within the radius ϵ. Otherwise, pix is labeled as noise (outlier).

The procedure begins by choosing an arbitrary unvisited pixel in the image. Taking the pixel and the value of input ϵ, Min_{pxls} within the region formed will be verified. If the total number of pixels in ϵ-neighborhood of pix is equal to or exceeds the input Min_{pxls} value, then it will build a cluster. Pixels lying outside the cluster form noise pixels. If a new pixel is added into a cluster, implies that its neighbors within ϵ-distance are all added to the same cluster as well. We continue to build further clusters by the same manner.

Once the clusters are obtained, the ANN classifier carries out the segmentation step. Here, all pixels of each cluster are fed to the ANN classifier to determine the color components (CC) that they represent (See Fig. 3). Once the process is done, the pixels will participate in the vote to identify the color of the cluster by choosing the CC with the highest value of vote. The above process is repeated until all clusters colors are identified.

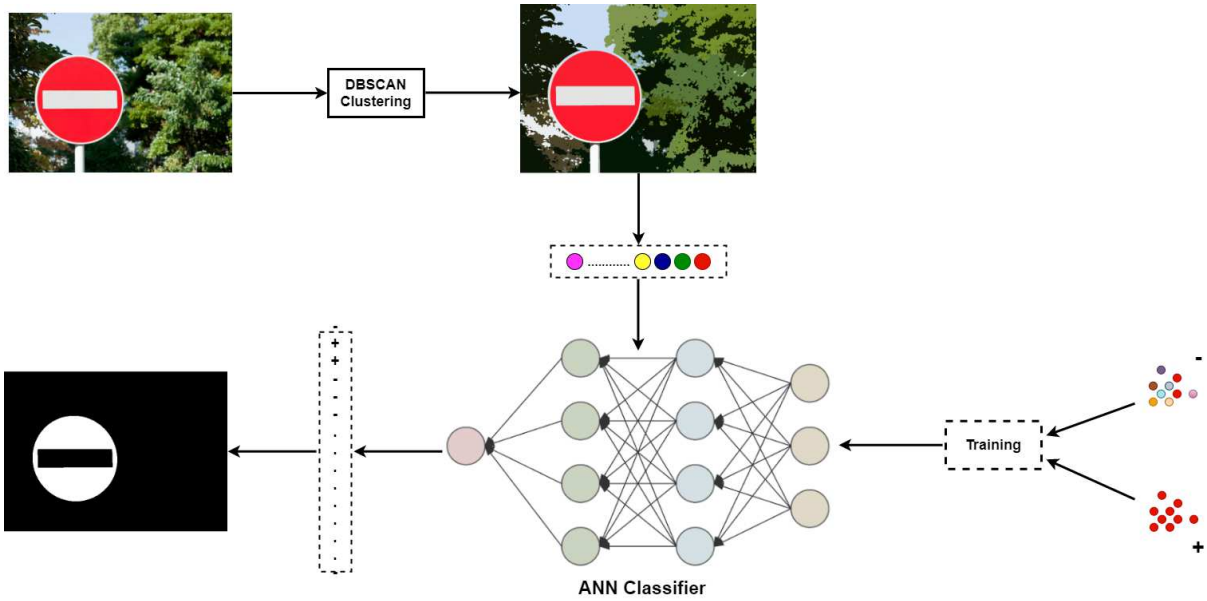


Fig. 3: Color segmentation procedure.

Figure 4 depicts an example of the segmentation results obtained when the proposed approach is applied on a road scene image. The segmentation is illustrated with a binary image where the ROIs are represented with white pixels.

The ROIs we get from the segmentation phase are mapped into their corresponding regions in the initial image. Each ROI in the binary image generates a ROI in the corresponding visible image (Fig. 5).

To ameliorate the detection approach performance, we discard the insignificant ROIs based on their sizes and aspect ratio.

(a) (b)

Fig. 4: (a) Input image. (b) Segmentation results.

(a) (b)

Fig. 5: Segmentation results mapped into the input image.

A ROI is taken in consideration if:

- The aspect ratio is between 1/2.03 and 2.03.
- The size is between $w \times h/27$ and $w \times h/2.6$, where h and w are the height and the width of the image sample, respectively.

The aspect ratio and the size are selected empirically regarding the collected data from German Traffic Sign Detection Benchmark (GTSDB) and Swedish Traffic Signs (STS). Figure 6 represents an example of the ROIs obtained in the image before and after discarding the insignificant ROIs. We can remark clearly form Fig. 6 (b) how the number of maintained ROIs is decreased.

(a) (b)

Fig. 6: Extracted ROIs in the original image (a) before and (b) after taking into account the size and aspect ratio constraints.

3.1.2 Shape classification

Once the ROIs are generated (candidate road signs), they are provided to the classification module to classify them as road sings or non-road signs. In this section, the approach utilized to classify

the provided ROIs according to their shapes is detailed. Generally, three shapes are assumed to be detected as traffic signs. The shapes are rectangle, triangle, and circle. In this work, we refer to histogram of dominant silhouette orientation (HDSO) descriptor together with ANN classifier to perform this classification.

Motivated by its success in the field of pedestrian detection [32] and inspired by the silhouette pattern of road sign as well, HDSO features is adopted to recognize road signs shape. The descriptor procedure requires three primary phases: The first phase, silhouette edge extraction, defines the silhouette of the road sign from the input ROI. The second phase of the proposed algorithm, polar transform, uses the polar representation of the coordinates instead of the Cartesian one. The last step, histogram computation, handles the procedure for HDOS feature vector extraction. The main three subsequent phases of HDSO descriptor are illustrated in Fig. 7. More details are listed below.

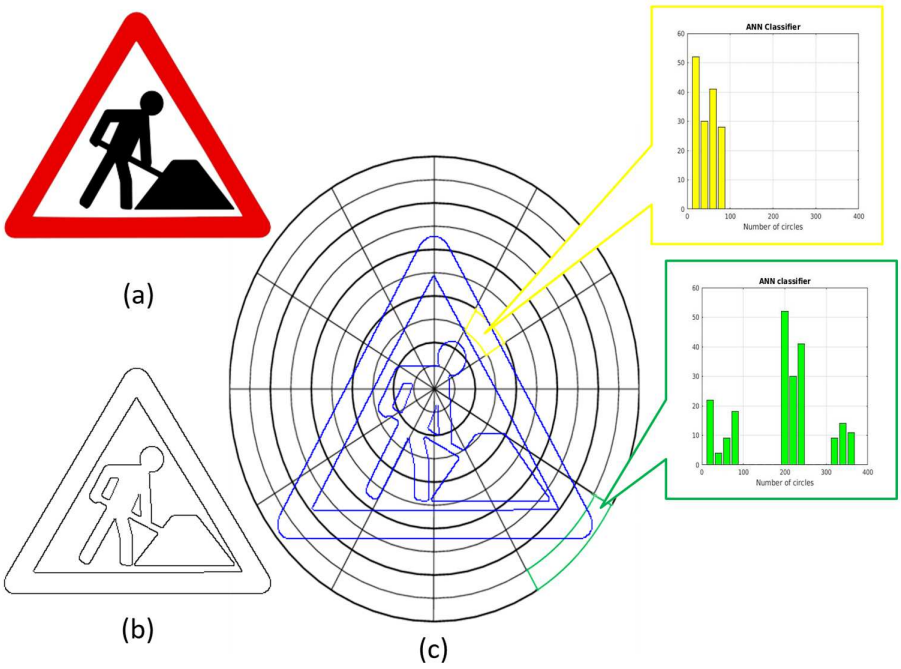


Fig. 7: HDSO descriptor illustration: (a) human silhouette, (b) human silhouette edge extraction, and (c) HDSO descriptor computation.

a. Silhouette edge extraction

To computed HDSO features and as mentioned above, we first start by silhouette edge extraction step using Canny operator [7]. Once it is done, the silhouette edge center (x_c, y_c) is determined by $y_c = \frac{1}{N}\sum_{i=0}^{N} y_i$ and $x_c = \frac{1}{N}\sum_{i=0}^{N} x_i$, where N is the total number of silhouette pixels. Then, we calculate the orientation ω_p at every edge point $p(x, y)$ according the the following equation:

$$\omega_p = \arctan(\frac{G_{y_p}}{G_{x_p}}) \tag{1}$$

where the gradient of the sample image along the directions x and y are utilized.

b. Polar transform

To explain efficiently the road sign form, polar representation is utilized. The cartesian coordinate system is converted into the corresponding polar coordinate system using Eqs. 2 and 3

$$r_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \tag{2}$$

and

$$\theta_i = \tan^{-1}(\frac{y_i - y_c}{x_i - x_c}) \tag{3}$$

where, θ_i represents the orientation relative to the silhouette centroid $c(x_c, y_c)$, and r_i signifies the euclidean distance between p_i and c. Once the computation is done, three information must be assigned for each p_i and they are recorded as follow: $[\omega_p, p_{r_i}, p_{\theta_i}]$.

c. Histogram computation

Descriptor computation begins with the creation of a log-polar histogram that represents dominant silhouette orientations distribution under consistent partitions around the silhouette's center $c(x_c, y_c)$ (Fig. 7 (c)). We partition the silhouette edge area into K cells by uniformly partitioning r_{max} (maximum radius) into n components that indicate the circles number, and angles into d directions such that $K = n \times d$. Then, for each cell $K_i = e \times f$ (where $e = 0, 1, 2, ..., n-1$ and $f = 0, 1, 2, ..., d-1$), we build a histogram $H^{e \times f}$ with 12 orientations spaced over 0° to 360° as follows:

$$\underset{\alpha_b = b \times \frac{\pi}{6}}{H^{e \times f}}(b) = \#\{p_{r_i, \theta_j} \mid \alpha_b \leq \omega_p < \alpha_b + \frac{\pi}{6}\} \quad \forall\, b = 0...11 \tag{4}$$

where the ranges of θ_j and r_i are defined by $\frac{2\pi v}{m} \leq \theta_j \leq \frac{2\pi(v+1)}{m}$ and $\frac{u}{n} r_{max} \leq r_i \leq \frac{u+1}{n} r_{max}$ respectively.

In order to globalize the description of HDSO, we select the dominant orientation $\beta^{e \times f}$ of the histogram from each cell according to the following equation:

$$\beta^{e \times f} = argmax_{b=0...11}(H^{e \times f}(b)) \tag{5}$$

The maximum value $\beta^{e \times f}$ would be taken as the cell $e \times f$ feature value. The above process is straightforwardly repeated until the dominant orientation is computed for each cell K_i of the built histogram. The resulting feature descriptor HDSO is constructed from a vector containing all cells values with a size of $n \times d$.

After computing the descriptor, it is fed to the ANN classifier to classify the provided ROIs into appropriate shapes. An overview of ANN classifier is presented in sub-section 3.2.2.

3.2 Recognition

Once the candidate blobs are classified into a shape class, the recognition procedure is taken place in order to identify the sign. In this section, we describe how the proposed hybrid feature is created. It is a combination of two features: The first one is proposed based on GLBP feature by involving RGB color information (named GLBP-Color) and the second one is picked among the features that have good performance in the road sign representation (e.g. HOG, LSS and Gabor). Among the Gabor, LSS and HOG, we look for the one that succeeds in improving the performance of the recognition results when it is combined with the GLBP-Color.

3.2.1 Features extraction

a. GLBP feature

The first feature involved in our experiments is the GLBP [27].It is a version of the LBP [41] descriptor based on the idea of combining two types of information (texture and gradient) to construct powerful and even more discriminative attributes. Its main concept consists to exploit the uniform LBP patterns to compute the histogram of oriented gradients in order to reduce the effect of the noise on the recognition results. The GLBP histogram dimension is defined by all possible width and angle values. Precisely, eight potential angle values or Freeman directions are available; meanwhile, the value of "1" in the uniform patterns varies from one to seven. This yields a GLBP histogram of 7×8 in which gradient attributes are accumulated. After computation, the normalization (L2-norm) is taken place to derive the GLBP histogram within the image.

b. GLBP-Color feature

Computation of the GLBP descriptor in the work of Ning Jiang. [27] begins by converting the input color image into a grayscale one for simplifying the computations. However, this preprocessing step affects badly the descriptor quality. It discards all color information and leaves only the luminance of each pixel. Hence, we are persuaded that computing the GLBP from color images instead of the grayscale ones could improve the quality of the recognition and classification as well.

In this paper, we propose a new approach to compute the color-based GLBP feature. The computed GLBP from the blue (B), red (R), green (G) components using our approach, so-called GLBP-Color, has the same size as the classical GLBP [27], and has also better performance of the classification and recognition compared to the one introduced in [27].

At each given image pixel p with the coordinates (i, j), we refer to its B, G and R value components by $\mathrm{B}(i, j)$, $\mathrm{G}(i, j)$ and $\mathrm{R}(i, j)$ respectively. We begin first by computing the bit binary code (BBC) of p by comparing its value with those of its 8 neighbor pixels one by one at each blue, green and red components. We denote the computed bit binary code of p from B, G and R components by $BBC_B(p)$, $BBC_G(p)$ and $BBC_R(p)$ respectively. The three computed values are used to compute the final bit binary code of the pixel p according to the following formula:

$$BBC(p) = XOR(BBC_B(p), BBC_G(p), BBC_R(p)) \tag{6}$$

XOR in Eq. 6 signifies the exclusive disjunction logical operation.

Once $BBC(p)$ is checked as a uniform pattern, we compute two parameters. The first parameter is the width (ω). It represents the number of occurrences of "1" in the binary code. The second parameter is angle (θ). It is the direction code of the middle pixel in "1" area of its binary code. Both θ and ω parameters are employed for mapping the position of bin in GLBP-Color histogram.

We then directly compute the gradient components (G_x, G_y) using the following equations:

$$G_x(i, j) = max\left(G_x^B(i, j), G_x^G(i, j), G_x^R(i, j)\right) \tag{7}$$

and

$$G_y(i, j) = max\left(G_y^B(i, j), G_y^G(i, j), G_y^R(i, j)\right) \tag{8}$$

where

$$G_x^{CC}(i, j) = CC(i + 1, j) - CC(i - 1, j), \;\; CC = B,\ G\ or\ R \tag{9}$$

and

$$G_y^{CC}(i, j) = CC(i, j + 1) - CC(i, j - 1), \;\; CC = B,\ G\ or\ R \tag{10}$$

CC signifies the color component which could be either B, G or R. Once the G_x and G_y are identified using Eqs. 9 and 10, respectively, the magnitude values at the pixel $p(i, j)$ can be computed as:

$$magnitude = \sqrt{G_x^2 + G_y^2} \tag{11}$$

The magnitude value is used as weights for voting by following the same steps as the classical GLBP.

c. Gabor feature

The second one is the Gabor feature [10] which is a set of band-pass filters that have been employed in different computer vision, pattern recognition, and signal processing problems, including texture analysis, due to their optimal properties in both spatial and frequency domains. Gabor filters have been used in numerous traffic sign recognition applications [17] in recent years, inspired by their success in extracting the essential activations to build a sparse object representation and their capability to multi-orientation and multi-scale image analysis and subsequently, one of the most effective contour detection and texture tools.

d. HOG feature

The third feature adopted in this work is the HOG feature. It was first introduced by Triggs and Dalal [9] for pedestrian detection. It is extensively adopted in various image processing and computer vision problems for detecting objects. The main concept of the descriptor is that the shape and the appearance of an object can overwhelmingly be described rather well by the distribution of the local intensity gradients or edge directions, even without having a precise knowledge of the corresponding edge or gradient positions. Its computation is done on an intensive grid of uniformly spaced cells and the accuracy is improved using overlapping local contract normalization.

e. LSS feature

LSS is the last feature involved in this work. It presents a local self-similarity description operator for object detection, first proposed by Irani et al. in [47]. The basic idea behind the LSS is to capture

the internal geometric layout of local regions and compare it throughout the images. This indicates that the input image is partitioned into small cells of the same size. Then, they are compared to a patch located at the center of the sample. The resulted distance surface is normalized and projected into a log-polar representation divided by radial and angle intervals. The feature value is regarded as the extreme value in the interval space.

f. Hybrid-based feature model

Newly, feature combination has been adopted exceedingly in different object detection and pattern recognition fields. It has become one of the most powerful alternatives in different complex systems including road sign recognition [18, 16]. Thus, we believe that our road sign recognition system can be enhanced by taking advantage of different discriminant information such as local information (self-similarity of color, edges, repetitive patterns) and entire information (texture, edge direction, color and shape) of images. In our paper, two combinations to build novel hybrid-based feature models are studied. These combinations are: GLBP-Color+Gabor, and GLBP-Color+LSS.

The proposed hybrid descriptors are formed according to the following formula:

$$H_f = [F_1||F_2] = (x_1, x_2, x_3, x_4, ..., x_n, y_1, y_2, y_3, y_4, ..., y_m) \tag{12}$$

Where (——) indicates the concatenation between the two features vectors $F_1 = (x_1, x_2, x_3, x_4, ..., x_n)$ and $F_2 = (y_1, y_2, y_3, y_4, ..., y_m)$. By verifying the validity and the effectiveness of the proposed hybrid features in Sec. 4, we are convinced that the last one (GLBP-Color+LSS) outperforms the former ones. Consequently, GLBP-Color+LSS will be the feature adopted by the proposed recognition method.

3.2.2 Classifiers

a. SVM

SVM is a supervised learning model designed by Vladimir N. Vapnik and Alexey Ya. Chervonenkis [53]. The fundamental principle of this classifier is to convert the input vectors by a non-linear transformation into a higher-dimensional space, and then find a hyperplane that separates the results. The found hyperplane (H) should have the greatest potential for generalization and isolate the negative samples from the positive ones. As depicted in the Fig. 8, the training dataset which belongs to two classes are presented by black and white circles. The hyperplane (H) that separates the negative samples from the positive ones is formed, in which the margin between the closest negatives and positives is maximal. The data located on the boundaries ($H1$ and $H2$) of the two classes are named support vectors. SVM was first used to solve problems of binary classification. It is, moreover, often used to solve multi-class problems, such that it is done by combinations of the binary classification problems.

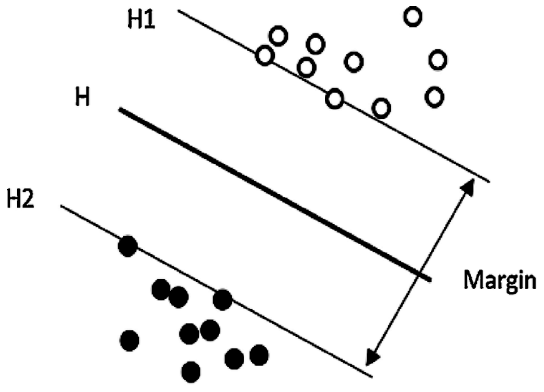


Fig. 8: The SVM binary classification.

b. Adaboost

Adaboost "Adaptive Boosting" was introduced by Freund Yoav et al. in [21]. It has been applied to different scientific fields by many researchers and has achieved good results. It consists of producing a strong classifier from a set of weak ones. The basic concept of boosting is selecting the best simple

and weak classifier after each iteration. The selected classifier is weighted based on the accuracy of classifying the training samples. Likewise, the samples incorrectly classified are weighted to choose the best weak classifier in the following iteration. This classifier uses exponential error loss as criterion. Ultimately, the chosen weak classifiers are linked together with various weights to construct a powerful and complex classifier.

c. ANN

ANNs have attracted great attention in machine learning due to their efficiency in difficult, complicated, multivariate nonlinear fields, such as road sign detection and recognition [26, 44]. It is a powerful and versatile tool that is capable of capturing and representing complex input/output relationships, and there is no need to assume an essential data allocation such as the ones usually done in statistical modeling. Broadly, the ANNs classifier consists of a series of simulated neurons operating in parallel, as one or multi-layer that are often composed of three layers: input, hidden and output layer (see Fig. 9).

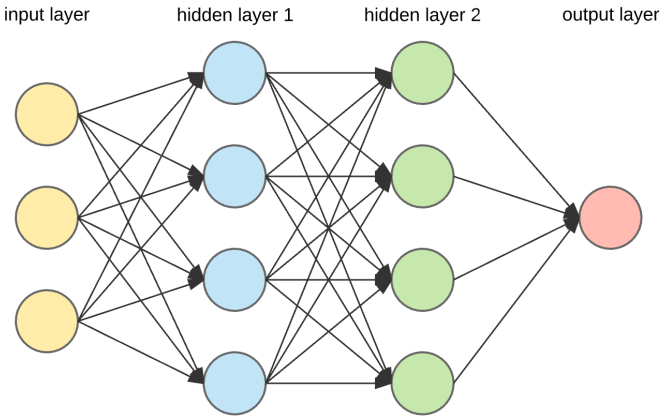


Fig. 9: Artificial neural network (ANN) architecture.

The function of the second layer (hidden) is to interact in any suitable way between the external input (vector samples) and the network output. For ROIs classification in both detection and recognition phases, we used Feed forward multilayer neural network ($FFNN$) model which is used mainly to classify inputs into a set of target categories based on feature selection parameters [38]. Furthermore, two types of signals are defined in this network (function and error signals). The first one, which is also acknowledged as the input signals, that are fed to the input of the $FFNN$ network, propagate forward through the network (neuron by neuron), and reach the network output end as output signals. The second signal originates at the output neuron of the network and propagates backward (layer by layer) via the network. The neural network's output can be expressed according to the following equation:

$$y = F_o(\sum_{j=0}^{M} W_{0j}(F_h(\sum_{i=0}^{N} W_{ji}X_i))) \tag{13}$$

where W_{0j} is the synaptic weights from neurony in the second layer to the single output neuron. F_h (resp. F_o) represents the activation function of the neurons from the hidden (resp. output) layer. X_j signifies the i^{th} element of the input vector, and W_{ji} is the connection weights between the neurons of the second layer and the inputs.

4 Experimental results

To evaluate the efficiency of the proposed traffic sign detection and recognition method, we carry out a series of comparative experiments using the three public datasets (GTSDB, STS, GTSRB) presented in Section 4.1. The obtained results clarify the contribution of each component of the proposed approach along with the entire approach using a 2.40 GHz Intel i5 processor.

4.1 Datasets

As mentioned above, three publicly available datasets have been used to assess system performance. The datasets are German Traffic Sign Detection Benchmark (GTSDB), Swedish Traffic Signs (STS), and German Traffic Sign Recognition Benchmark (GTSRB). These datasets were collected in urban areas with different weather and outdoor lighting conditions using the visible cameras.

The GTSDB dataset contains 900 full images. Those images are divided into two sets. The first one with 600 images is used for the training phase while the second one (300 images) is utilized for the testing phase.

The STS dataset provides 20000 images with 20% labeled. The images are captured from Swedish highways and cities, and they contain more than 3400 traffic signs.

The GTSRB dataset offers more than 50000 German traffic signs images in total, divided into 43 classes. The images format is 24-bit color PPM and their size is varying from 15×15 to 250×250 pixels. Fig. 10 illustrates the GTSRB dataset classes, which have been partitioned into six subsets (Speed limit, derestriction, mandatory, danger, other prohibitory and unique signs).

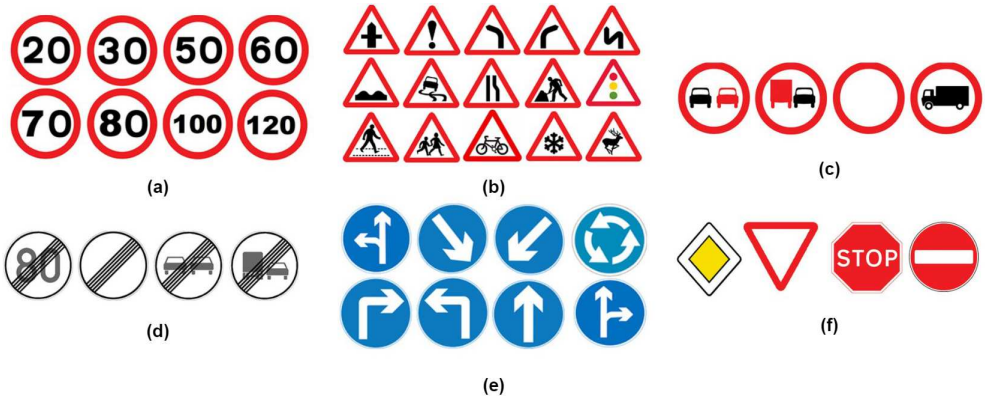


Fig. 10: Subsets of road signs in the GTSRB dataset: (a) Speed limit. (b) Derestriction. (c) Mandatory. (d) Danger. (e) Other prohibitory. (f) Unique.

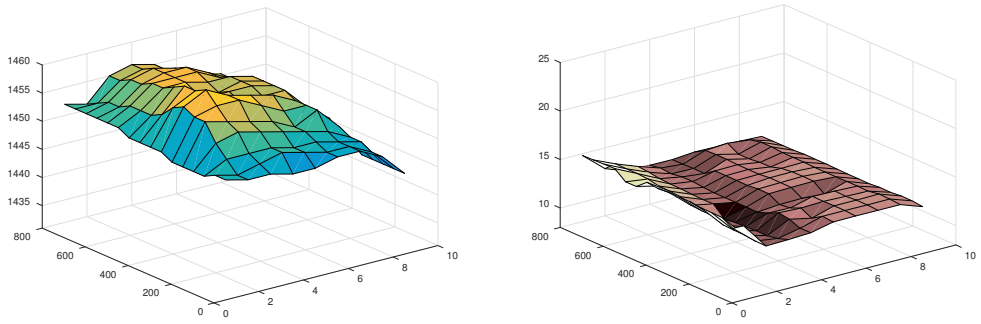
Both first and second datasets are used to evaluate the performance of the detection phase once the images are normalized to 640×480 pixels using bilinear interpolation. In the recognition phase, we adopt the last dataset for evaluation, which will permit us to easily compare our approach with other state-of-the-art methods.

4.2 Parameters Setting

In this section, the effect of the parameters involved in the different steps of the proposed system is investigated. The parameters are chosen empirically using some sample images from GTSDB, STS and GTSRB datasets.

In the color segmentation process, two parameters of DBSCAN clustering technique were used: Min_{pxls} and ϵ (epsilon) the radius from a corresponding pixel, which includes neighboring pixels. Figure 11 depicts the number of true positives (TPs) obtained and the corresponding computational time while changing the value of these two parameters over more than 250 images chosen from the GTSDB and STS datasets. Note that a correct detected road sign is counted TP if its corresponding bounding box overlaps with at least 50% of the area covered by the road sign present in the image. As shown in Fig. 11, both parameters have a noteworthy influence on the obtained results. $Min_{pxls} = 300$ and $\epsilon = 3$ are chosen since they guarantee high accuracy (more than 1455 TPs) while the consumed time is at its lowest value (less than 12 ms).

The ANN parameters are also obtained empirically using data from GTSRB, GTSDB and STS datasets on both detection and recognition phase. We remark from Fig. 12 that the classification accuracy enhances with the number of nodes (N_{Nodes}) and becomes constant once this number reaches a specific value. Here, detection and recognition stages hit the highest accuracy score when



(a) (b)

Fig. 11: (a) Number of TPs and (b) the computational time while varying the parameters ϵ and Min_{pxls}.

N_{Nodes}=18 and N_{Nodes}=9, respectively. Thus, the value N_{Nodes}=18 has been adopted to be the number of nodes in ANN classifier. The ANN parameters used for the training in both detection and recognition stages are summarized in Table 1.

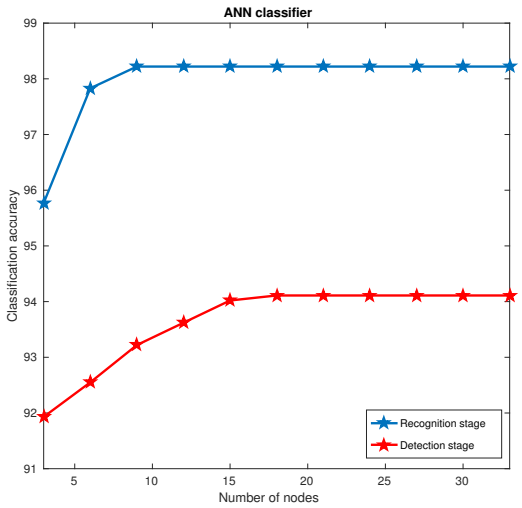


Fig. 12: The average classification accuracy of the ANN classifier in different method stages.

Table 1: The ANN network parameters used for training phase.

Parameters	ANN
Number of input layer units	13
Number of hidden layer	02
Number of first hidden layer unit	10
Maximum number of epochs to train	2300
Learning rate	0.62
Minimum performance gradient	1e - 10
Error after learning	0.000042

To achieve the optimum parameters of the features used in the proposed method, a procedure based on cross-validation experiments has been used. The datasets are partitioned into two subsets

(training and validation). The classifiers on the two sub-sets are trained and evaluated using various features and the parameters that optimize the accuracy of validation are chosen. Using the selected features parameters, the classifiers were retrained one more time on the training dataset. Here, the GLBP-Color, HOG, LBP, LSS and HDSO features are computed as follows:

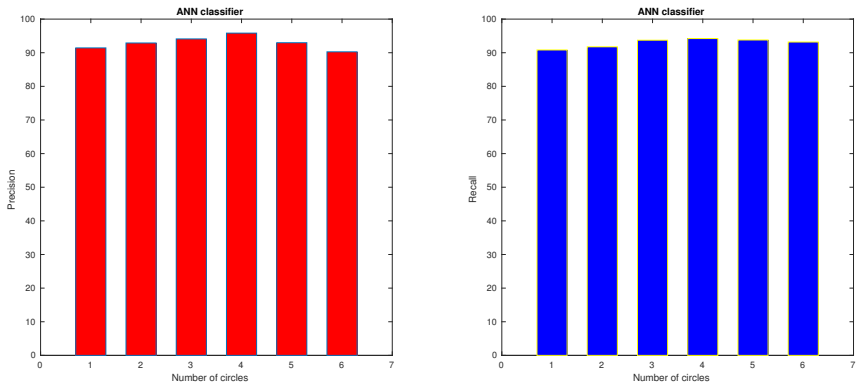
- For GLBP-Color, once the detected road sign ROI is normalized to 64×128, it is partitioned into $7 \times 15 = 105$ blocks with a size of 16×16. For each one of these blocks, a histogram has been constructed using 56 (8×7) bins. To identify the most efficient version of GLBP-Color for road sign detection, we successively experimented with each version under the GTSDB and STS datasets. Table 2 summarizes the performance evaluation of GLBP-Color$_{P,R}$ descriptor provided for different parameters (where P is the sampling points and R is the radius of circle) values. The values $P = 8$ and $R = 3$ are obtained from cross-validation analyses performed on the training dataset.
- For HOG features, the detected ROI is normalized to 64×128 and partitioned into $7 \times 15 = 105$ overlapping blocks. We split each block into 2×2 cells with 8×8 pixels. For each cell, we compute a gradient histogram using 9 bins. We form a 3780 HOG vector.
- For the LBP features, we normalize the sample to 64×128 as well and divided into 8×8 blocks. We build a vector of 59 for each block using the uniform patterns approach. This results in a 3776 LBP vector.
- For LSS features, 5×5 patches in a larger surrounding image region equal to 40×40 pixels have been adopted and the log-polar coordinates are divided into 4 radial intervals and 20 angles.
- The number of angles (m) and the number of circles (n) are the parameters for HDSO features. The values of both parameters affect detection efficiency. As depicted in Fig. 13, the highest results scores are obtained when $m = 36$ and $n = 4$. Therefore, the values 36 and 4 are adopted for m and n respectively.

Table 2: Performance evaluation of GLBP-Color descriptor provided for different parameters values in terms of CCR (%) and running time (ms).

Version	CCR(%)	Computing time (ms)
GLBP-Color$_{8,1}$	97.54	19.36
GLBP-Color$_{8,2}$	97.60	19.42
GLBP-Color$_{8,3}$	**97.97**	**19.57**
GLBP-Color$_{16,1}$	97.61	19.82
GLBP-Color$_{16,2}$	97.59	19.98
GLBP-Color$_{16,3}$	97.53	20.21
GLBP-Color$_{32,1}$	97.46	20.37
GLBP-Color$_{32,2}$	97.43	20.42
GLBP-Color$_{32,3}$	97.38	20.49

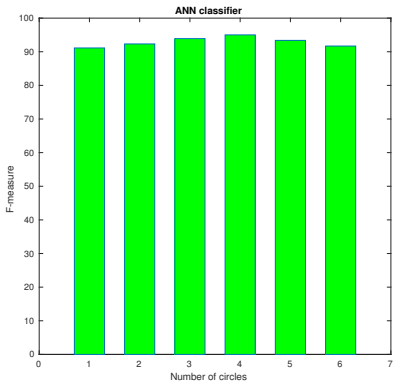
4.3 Results

Fig. 14. illustrates an example of the results provided by the proposed module at its main stages applied to a sample image captured by the visible camera. First, the input image (Fig. 14 (a)) is segmented using the DBSCAN algorithm together with the color information (Fig. 14 (b)). The resulting clusters are projected on the image to obtain their corresponding ROIs (Fig. 14 (c)). Some of these resulting ROIs are eliminated according to their aspect ratio and size. This procedure speeds up the detection since the number of ROIs to be treated is reduced (Fig. 14 (d)). The segmentation approach retrieves the road sign present in the Fig. 14 (a) along with some other undesirable ROIs that they do not represent any traffic signs. To validate the detected ROIs, a shape classification procedure should be applied. Figure 14 (e) depicts the traffic sign detection results when the combination between HDSO and ANN are used as a feature and classifier, respectively on the obtained ROIs. To identify the detected traffic sign, a green bounding box is used. Once the road sign is detected, it is provided to the recognition module to identify which class it belongs to. Fig. 14 (f)) represents the recognition results using ANN as a classifier and a color-based GLBP together with LSS as features. The computational time of each of the stages involved in the detection and recognition process is depicted in Table 3.



(a)

(b)



(c)

Fig. 13: HDSO descriptor performance provided for various parameters values. (a) Precision. (b) Recall. (c) F-measure.

Table 3: Running time of each step of the proposed approach in ms/f.

	Detection	Recognition
Consuming time	37.43	19.62

To demonstrate the effectiveness of the proposed approach, we analyze and compare the obtained results from each step.

We first evaluate the proposed detection module and compare the obtained results with other reported works. Second, we assess the performance of the proposed recognition approach and we perform a comparison with the state-of-the-art approaches on the GTSRB dataset. The evaluation of the detection and recognition stages is carried out on the basis of recall, precision, F-measure, and accuracy, where they are computed as follows:

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \times 100 \tag{14}$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \times 100 \tag{15}$$

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ samples} \tag{16}$$

$$F - measure\ (F - score) = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{17}$$

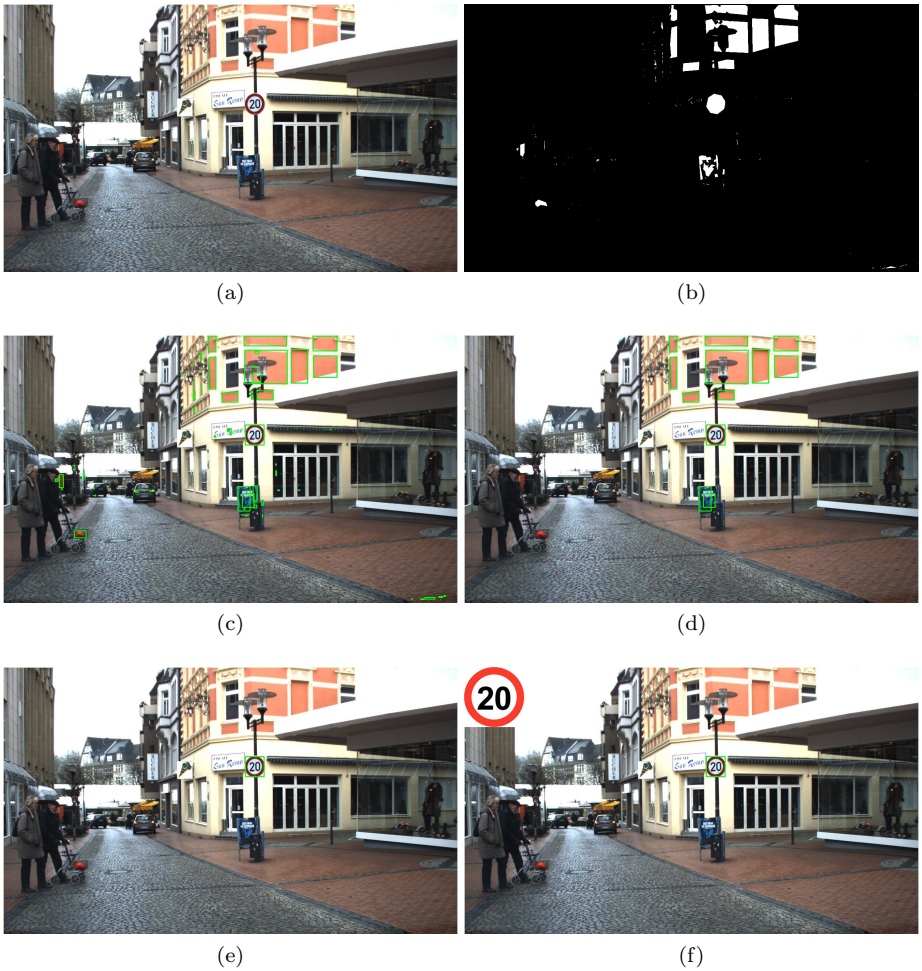


Fig. 14: Example of the results of each step of the proposed methodology. (a) Original image. (b) Segmentation results. (c) Segmentation results mapped into the original image. (d) Segmentation results after taking into account the size and aspect ratio constraints. (e) Road sign detection results. (f) Road sign recognition results.

As depicted in Table 4 and Table 5, the proposed detection method yields the scores with a precision of 95.83% at a recall of 94.22%, and a precision of 96.07% at a recall of 94.89% in GTSDB and STS datasets respectively. Figures 15 (a) and Figure 15 (b) show the precision-recall curves of the proposed approach when applied to GTSDB and STS datasets respectively. The AUC of the two ROC curves is 95.76% and 96.67%, respectively.

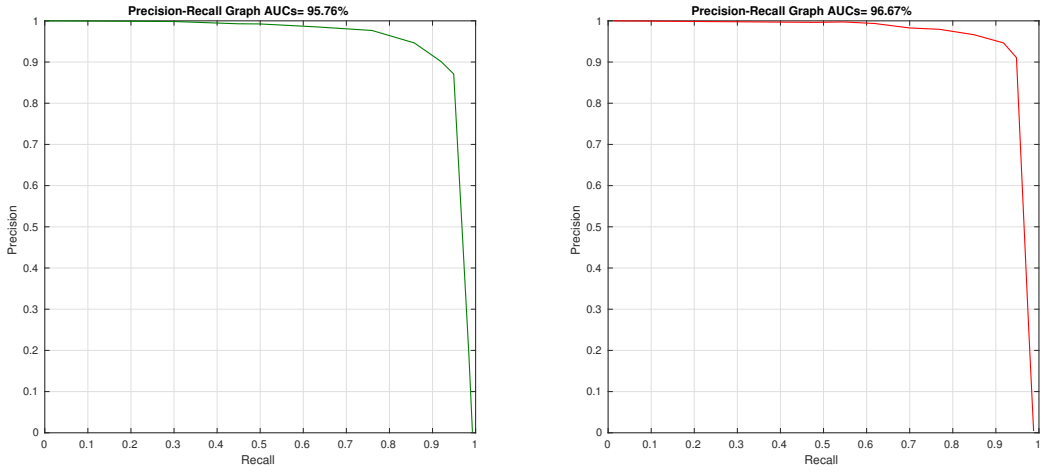
Table 4: The best trade-off between the precision, recall values and the AUC obtained by the proposed method and the ones reported in [16] and [15] on the GTSDB dataset.

	The proposed method	The method in [16]	The method in [15]
Precision	**95.83%**	90.13 %	94.03 %
Recall	**94.22%**	91.07 %	92.98 %
AUC	**95.76%**	93.69 %	94.22 %

The proposed detection approach has been compared to other reported methods in order to assess its performance using the GTSDB dataset. The methods as well as their results in terms of precision, recall, and F-measure are listed in Table 6. We notice from this table that our method outperforms

Table 5: The best trade-off between the precision, recall values and the AUC obtained by the proposed method and the ones reported in [16] and [15] on the STS dataset.

	The proposed method	The method in [16]	The method in [15]
Precision	**96.07%**	90.27 %	94.15 %
Recall	**94.89%**	93.27 %	93.87 %
AUC	**96.67%**	94.05 %	95.17 %



(a) (b)

Fig. 15: Precision-recall curves of the proposed method when applied to: (a) GTSDB and (b) STS dataset.

the ones introduced in [15, 20, 16, 57, 25, 58, 24, 23, 2] and [54] for which the F-measure scores are 93.50, 90.97, 90.59, 88.73, 70.70, 65.28, 65.07, 62.66, 54.57, and 46.42, respectively.

Table 6: Quantitative GTSDB traffic detection comparison between the proposed method and other published approaches using F-measure (in %).

Reference	F-measure (%)	Method description
Yang C [54]	46.42	SDA[a]
Bascón [2]	54.57	HST[b]
Gómez-Moreno [23]	62.66	RGBNT[c]
Greenhalgh [24]	65.07	MSERs[d]
Zaklouta [58]	65.28	Win-HOG[e]
Houben [25]	70.70	CVS[f]
Yuan X [57]	88.73	GBR[g]
Ellahyani [16]	90.59	HSI-Hu[h]
Fan Y [20]	90.97	NN-HOG[i]
Ellahyani et al. [15]	93.50	Mean shift + Log-polar transform + Random Forest
Our proposed Method	**95.02**	DBSCAN clustering + HDSO feature + ANN

[a] SDA: graph-based saliency detection algorithm.
[b] HST: SVM hyper-parameters optimization strategy.
[c] RGBNT: RGB space Normalized Threshold method.
[d] MSERs: Maximally Stable Extremal Regions.
[e] Win-HOG: sliding window algorithm with HOG features.
[f] CVS: learned color gradient with the constant vote system.
[g] GBR: Graph-Based Ranking and segmentation algorithm.
[h] HSI-Hu: HSI based segmentation and Hu moments algorithm.
[i] NN-HOG: Neural Networks with Random Weights combined with HOG features algorithm.

More detection results are depicted in Fig. 16. The clustering results we obtain when we apply the DBSCAN algorithm to the test image are illustrated in Figure 16 (b). The generated ROIs after mapping the resulting clusters on the image (Fig. 16 (a)) with and without taking into account the aspect ratio and size are shown in Figure 16 (d) and Figure 16 (c) respectively. To validate the detected ROIs, a classification procedure was performed on the basis of the HDSO features and ANN classifier. The classified ROI is illustrated in the test image by green bounding boxes in Fig. 16 (e).

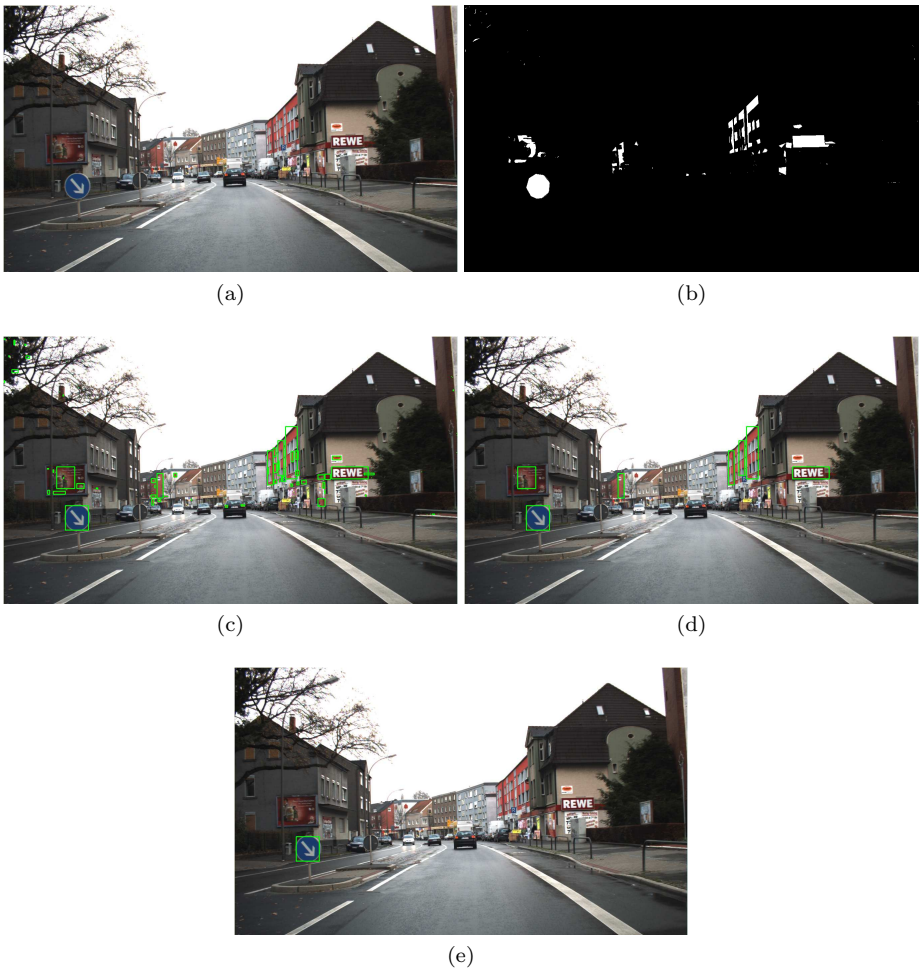


Fig. 16: (a) Input image. (b) Segmentation results. (c) Segmentation results projected into the original image. (d) ROIs obtained after taking into account the size and aspect ratio constraints. (e) Detection results.

To evaluate the recognition module, GTSRB dataset has been used. Here, five evaluations have been included to prove the relevance of the proposed GLBP-Color. We first evaluate the performance of GLBP-Color and compare it with the classical GLBP [27] and some other LBP versions (RGB-LBP [1] , ALBP [37] and tLBP [52]). In the second evaluation, we compare the performance of GLBP-Color with some single feature descriptors widely utilized in the field (HOG [9] , LSS [47] and Gabor [10] features). The third one assesses results of GLBP-Color with some color descriptors including Hue SIFT [5] , Rg SIFT [5] , HSV SIFT [5] , RGB SIFT [5] , HSV-HOG [16] , and RGB-HOG [55]. In the fourth evaluation, we try to combine the proposed GLBP-Color together with Gabor and LSS features to look for possible improvements. The last evaluation compares the results provided by ANN and those obtained by the SVM with radial basis function (RBF) kernel and Adaboost classifiers.

Table 7 represents the accuracy values obtained while using the proposed GLBP-Color, grayscale, GLBP, LBP, RGB-LBP, ALBP and tLBP to the GTSRB dataset which is composed of six subsets

(See Fig. 10). It is clear from Table 7 that the GLBP-Color outperforms all the other descriptors in terms of accuracy in all subsets. Its corresponding accuracy of all traffic signs is 97.97%, which has been improved by 2.07%, 2.01%, 1.97%, 2.08% and 0.74% when the new GLBP-Color is compared to tLBP, ALBP, RGB-LBP, LBP, grayscale GLBP, respectively.

Table 7: Performance of LBP, ALBP, tLBP, RGB-LBP, GLBP, GLBP-Color features when tested on the GTSRB dataset.

Feature	**CCRs(%) of all traffic signs**	**CCRs(%) of each subset**					
		(a)	(b)	(c)	(d)	(e)	(f)
LBP [41]	95.89	96.17	88.50	98.17	95.10	98.85	98.54
tLBP [52]	95.90	96.11	88.53	98.21	95.10	98.85	98.60
ALBP [37]	95.96	96.23	88.56	98.21	95.17	98.91	98.66
RGB-LBP [1]	96.00	96.28	88.58	98.25	95.17	99.03	98.73
GLBP [27]	97.23	97.23	91.55	99.24	96.29	99.40	99.68
GLBP-COLOR	**97.97**	**98.00**	**93.59**	**99.32**	**97.55**	**99.52**	**99.81**

Table 8 lists the results obtained while using the proposed GLBP-Color, HOG, LSS, and Gabor features. We remark from the table that among the four features, the GLBP-Color feature still performs better than the HOG, LSS and Gabor features.

Table 8: Performance of the single features when tested on the GTSRB dataset.

Feature	**CCRs(%) of all traffic signs**	**CCRs(%) of each subset**					
		(a)	(b)	(c)	(d)	(e)	(f)
Gabor [10]	95.97	96.28	88.59	98.29	95.17	99.09	98.40
LSS [47]	95.97	96.28	88.61	98.33	95.17	98.97	98.47
HOG [9]	96.39	96.55	89.26	98.86	95.32	99.09	99.24
GLBP-COLOR	**97.97**	**98.00**	**93.59**	**99.32**	**97.55**	**99.52**	**99.81**

The same remark could be seen on the basis of the results listed in Table 9, when we compare GLBP-Color with some other color descriptors using ANN classifier.

Table 9: Performance of GLBP-Color and some color-based descriptors when tested on the GTSRB dataset.

Feature	**CCRs(%) of all traffic signs**	**CCRs(%) of each subset**					
		(a)	(b)	(c)	(d)	(e)	(f)
Hue SIFT [5]	96.55	96.78	89.59	98.93	95.47	99.15	99.36
Rg SIFT [5]	96.72	96.84	90.14	98.97	95.69	99.27	99.43
HSV SIFT [5]	96.91	96.95	90.63	99.05	95.92	99.34	99.55
RGB SIFT [5]	97.00	97.00	90.97	99.16	95.99	99.34	99.55
HSV-HOG [16]	97.41	97.61	91.88	99.24	96.51	99.46	99.75
RGB-HOG [55]	97.60	97.78	92.31	99.28	97.10	99.40	99.75
GLBP-COLOR	**97.97**	**98.00**	**93.59**	**99.32**	**97.55**	**99.52**	**99.81**

With the aim to enhance the performance of the proposed GLBP-Color, it has been combined with other features, i.e., Gabor and LSS. Table 10 indicates the results obtained by the GLBP-Color alone, the combination of GLBP-Color and Gabor and the combination of GLBP-Color and LSS. We remark that the combination of GLBP-Color and LSS succeeds in improving the accuracy in all GTSRB datasets subsets. It provides 98.22% in terms of accuracy of all subsets. Therefore, the combination GLBP-Color and LSS adopted as the main feature vector in the recognition stage of the proposed method.

In order to justify the use of the ANN as a classifier, a comparison with the SVM with basis function (RBF) kernel and Adaboost classifiers is made. Table 11 indicates the results given by the

Table 10: Performance of GLBP-Color, GLBP-Color + Gabor and GLBP-Color + LSS features when tested on the GTSRB dataset.

Feature	**CCRs(%) of all traffic signs**	**CCRs(%) of each subset**					
		(a)	(b)	(c)	(d)	(e)	(f)
GLBP-COLOR	97.97	98.00	93.59	99.32	97.55	99.52	99.81
GLBP-COLOR+Gabor	98.07	98.22	93.65	99.47	97.70	99.58	99.81
GLBP-COLOR+LSS	**98.22**	**98.39**	**93.84**	**99.62**	**97.92**	**99.70**	**99.87**

three classifiers when they are feed with all features already used. The SVM parameters used in this comparison are: C=5 and G=0.07. It is obvious from the results that the ANN outperforms the other classifiers. Thus, the ANN is chosen in the proposed recognition method as a classifier.

Table 11: The accuracy and the average running time of the classifiers used in this work.

Feature	**Accuracy(%) of all dataset**			**Run time (ms/frame)**		
	ANN	SVM	Adaboost	ANN	SVM	Adaboost
Gabor	95.97	95.80	95.63	23.21	42.09	41.79
LBP	95.89	95.77	95.54	13.72	34.07	33.91
tLBP	95.90	95.78	95.55	13.58	33.82	33.76
ALBP	95.96	95.79	95.59	13.67	34.00	33.87
RGB-LBP	96.00	95.80	95.61	13.80	34.15	34.08
LSS	95.97	95.81	95.64	13.92	34.23	34.13
HOG	96.39	96.24	96.10	19.83	40.19	40.11
Hue SIFT	96.55	96.44	96.32	19.88	40.26	40.18
Rg SIFT	96.72	96.96	96.83	19.94	40.35	40.20
HSV SIFT	96.91	97.01	96.89	19.96	40.39	40.26
RGB SIFT	97.00	97.04	96.91	19.95	40.33	40.21
GLBP	97.23	97.15	97.00	19.87	40.03	39.90
HOG-HSV	97.41	97.30	97.15	19.67	39.89	39.83
HOG-RGB	97.60	97.52	97.36	19.63	39.81	39.75
GLBP-COLOR	97.97	97.83	97.71	19.57	40.11	38.72
GLBP-COLOR+Gabor	98.07	97.87	97.75	19.99	40.38	40.25
GLBP-COLOR+LSS	**98.22**	98.14	98.04	**19.62**	40.29	40.22

A comparison versus some state-of-the-art works is given to evaluate the performance of the proposed approach. The works used for the comparison, their descriptions and their results are presented in Table 12. We notice from this table that Committee of CNNs [8] and Multi-scale CNNs [46] achieve high accuracy compared to our method. However, the adoption of CNNs technology is expensive in terms of resources including computational time and the hardware used in the experiments. Furthermore, both approaches extend the training dataset by encoding, scaling, and rotating the samples using random values in the training dataset. The proposed method surpasses the ones reported in [48, 59, 51, 16, 35] and [18], with gains of 2.54%, 2.08%, 1.03%, 0.79%, 0.39% and 0.26% in the accuracy compared with the state-of-the-art methods [48, 59, 51, 16, 35, 18] respectively, and the computational time is very low when we compared to CNNs-based approaches

To assess the performance of the entire traffic sign detection and recognition system, the GTSDB dataset will be adopted. Fig. 17 shows the precision-recall curve obtained by the whole proposed method. As illustrated in this figure, the value of AUC, which is mathematically acknowledged as the integral of the curve over the precision and the recall labels, obtained is 95.17% on an average run time of 8-10 frames/second.

Figures 18, 19 and 20 present some examples of the detection and recognition results when the proposed approach is applied to sample images. As can be seen in Fig. 18, the system successfully detects recognizes the traffic signs included in the two images. The traffic signs presented in Fig. 19 have been successfully detected. However, they could not be recognized well by the reason of the motion blur in the signs. In Figure 20, the road sings could not be detected due to different reasons. The ROIs corresponding to road sings presented in the images were not extracted by the segmentation approach

Table 12: Quantitative GTSRB traffic recognition comparison between the proposed method and other published approaches using CCR (in %).

Reference	Accuracy (%)	Method description
Ciresan et al. [8]	99.46	Committee of CNNs
Sermanet et al. [46]	98.31	Multi-scale CNNs
Our proposed Method	**98.22**	GLBP-Color + LSS + ANN
Ellahyani et al. [18]	97.96	Log-polar transform + HOG + LBP + LSS + Random Forests
Liu et al. [35]	97.83	Log and sparse coding
Ellahyani et al. [16]	97.43	HSI-HOG + LSS + Random forests
Sun et al. [51]	97.19	BW-ELM[a]
Zaklouta et al. [59]	96.14	Random forests
Stallkamp et al. [48]	95.68	LDA on HOG2[b]

[a] BW-ELM: Between-category to within-category sums of squares - extreme learning machine.
[b] LDA on HOG2: Linear Discriminant Analysis with HOG features.

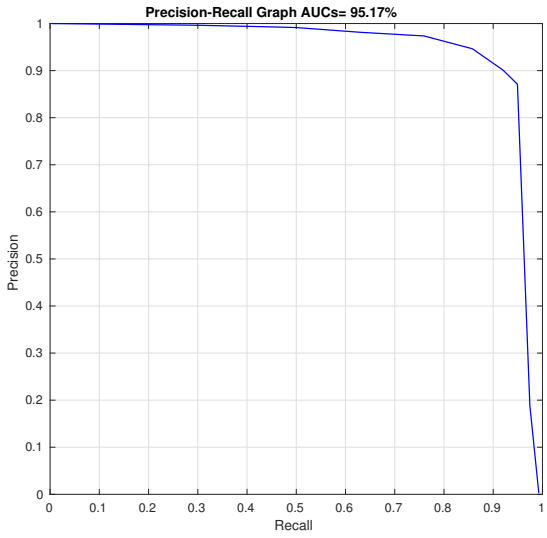


Fig. 17: Precision–recall curve of the proposed detection and recognition approach.

(a) (b)

Fig. 18: Examples of detection and recognition results.

5 Conclusion and perspectives

This paper presents a two stages system for real-time road sign detection and recognition. The first step performs the detection on the basis of color and shape cues. A clustering technique is carried out on the initial image to form a set of connected components. The resulting clusters are provided to the ANN classifier for segmentation according to their color. The obtained ROIs (possible traffic signs) are then processed based on their size and aspect ratio to keep only significant ones. Then, we

(a) (b)

Fig. 19: Examples of detection with miss-recognition.

(a) (b)

Fig. 20: Examples of misdetections.

refer to HDSO feature and ANN classifier to detect circular, triangular and rectangular shapes on the resulting ROIs. In the recognition step, we combine the so-called GLBP-Color with LSS feature to form a new descriptor. This descriptor is then used with the ANN classifier to identify the signs from the detected ROIs. Results obtained on the public GTSDB, STS, and GTSRB datasets justify the effectiveness and robustness of the proposed method.

In future work, we are intending to enhance the quality of the results obtained by the proposed method in both detection and recognition phases. We aim also to use other machine learning techniques to accelerate the classification procedure, and improve the robustness of the system.

Compliance with ethical standards

– **Funding** : This research work was supported by the National Center for Scientific and technical Research (CNRST), research grant No: 20UIZ2015.
– **Conflict of Interest** : All the authors have no conflict of interest to declare.
– **Ethical approval** : This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Sugata Banerji, Abhishek Verma, and Chengjun Liu. Lbp and color descriptors for image classification. In *Cross Disciplinary Biometric Systems*, pages 205–225. Springer, 2012.
2. S Maldonado Bascón, J Acevedo Rodríguez, S Lafuente Arroyo, A Fernndez Caballero, and Francisco López-Ferreras. An optimization on pictogram identification for the road-sign recognition task using svms. *Computer Vision and Image Understanding*, 114(3):373–383, 2010.
3. Rachid Belaroussi, Philippe Foucher, Jean-Philippe Tarel, Bahman Soheilian, Pierre Charbonnier, and Nicolas Paparoditis. Road sign detection in images: A case study. In *2010 20th International Conference on Pattern Recognition*, pages 484–488. IEEE, 2010.

4. Mohamed Benallal and Jean Meunier. Real-time color segmentation of road signs. In *CCECE 2003-Canadian Conference on Electrical and Computer Engineering. Toward a Caring and Humane Technology (Cat. No. 03CH37436)*, volume 3, pages 1823–1826. IEEE, 2003.
5. Simone Bianco, Davide Mazzini, DP Pau, and Raimondo Schettini. Local detectors and compact descriptors for visual search: a quantitative comparison. *Digital Signal Processing*, 44:1–13, 2015.
6. Amal Bouti, Med Adnane Mahraz, Jamal Riffi, and Hamid Tairi. A robust system for road sign detection and classification using lenet architecture based on convolutional neural network. *Soft Computing*, pages 1–13, 2019.
7. John Canny. A computational approach to edge detection. In *Readings in Computer Vision*, pages 184–203. Elsevier, 1987.
8. Dan Cireşan, Ueli Meier, Jonathan Masci, and Jürgen Schmidhuber. A committee of neural networks for traffic sign classification. In *The 2011 international joint conference on neural networks*, pages 1918–1921. IEEE, 2011.
9. Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893. IEEE, 2005.
10. John G Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *JOSA A*, 2(7):1160–1169, 1985.
11. Mohamed El Ansari, Redouan Lahmyed, and Alain Trémeau. A hybrid pedestrian detection system based on visible images and lidar data. In *VISIGRAPP (5: VISAPP)*, pages 325–334, 2018.
12. Mohamed El Ansari, Lhoussaine Masmoudi, and Abdelaziz Bensrhair. A new regions matching for color stereo images. *Pattern recognition letters*, 28(13):1679–1687, 2007.
13. Mohamed El Ansari, Stéphane Mousset, and Abdelaziz Bensrhair. Temporal consistent real-time stereo for intelligent vehicles. *Pattern Recognition Letters*, 31(11):1226–1238, 2010.
14. Ayoub Ellahyani and Mohamed El Ansari. Complementary features for traffic sign detection and recognition. In *2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA)*, pages 1–6. IEEE, 2016.
15. Ayoub Ellahyani and Mohamed El Ansari. Mean shift and log-polar transform for road sign detection. *Multimedia Tools and Applications*, 76(22):24495–24513, 2017.
16. Ayoub Ellahyani, Mohamed El Ansari, and Ilyas El Jaafari. Traffic sign detection and recognition based on random forests. *Applied Soft Computing*, 46:805–815, 2016.
17. Ayoub Ellahyani, Mohamed El Ansari, Ilyas El Jaafari, and Said Charfi. Traffic sign detection and recognition using features combination and random forests. *International Journal of Advanced Computer Science and Applications*, 7(1):686–693, 2016.
18. Ayoub Ellahyani, Mohamed El Ansari, Redouan Lahmyed, and Alain Trémeau. Traffic sign recognition method for intelligent vehicles. *JOSA A*, 35(11):1907–1914, 2018.
19. Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231, 1996.
20. Yanjun Fan and Weigong Zhang. Traffic sign detection and classification for advanced driver assistant systems. In *2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, pages 1335–1339. IEEE, 2015.
21. Yoav Freund, Robert Schapire, and N Abe. A short introduction to boosting. *Journal-Japanese Society For Artificial Intelligence*, 14(771-780):1612, 1999.
22. Dariu M Gavrila. Multi-feature hierarchical template matching using distance transforms. In *Proceedings. Fourteenth international conference on pattern recognition (Cat. No. 98EX170)*, volume 1, pages 439–444. IEEE, 1998.
23. Hilario Gómez-Moreno, Saturnino Maldonado-Bascón, Pedro Gil-Jiménez, and Sergio Lafuente-Arroyo. Goal evaluation of segmentation algorithms for traffic sign recognition. *IEEE Transactions on Intelligent Transportation Systems*, 11(4):917–930, 2010.
24. Jack Greenhalgh and Majid Mirmehdi. Real-time detection and recognition of road traffic signs. *IEEE transactions on intelligent transportation systems*, 13(4):1498–1506, 2012.
25. Sebastian Houben. A single target voting scheme for traffic sign detection. In *2011 IEEE Intelligent Vehicles Symposium (IV)*, pages 124–129. IEEE, 2011.
26. Kh Tohidul Islam and Ram Gopal Raj. Real-time (vision-based) road sign recognition using an artificial neural network. *Sensors*, 17(4):853, 2017.

27. Ning Jiang, Jiu Xu, and Satoshi Goto. Pedestrian detection using gradient local binary patterns. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 95(8):1280–1287, 2012.
28. Zakaria Kerkaou, Nawal Alioua, Mohamed El Ansari, and Lhoussaine Masmoudi. Edge points-based stereo matching approach for omnidirectional images. *Journal of Electronic Imaging*, 27(5):053015, 2018.
29. Zakaria Kerkaou and Mohamed El Ansari. Support vector machines based stereo matching method for advanced driver assistance systems. *Multimedia Tools and Applications*, 79(37):27039–27055, 2020.
30. Redouan Lahmyed and Mohamed El Ansari. Multisensors-based pedestrian detection system. In *2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA)*, pages 1–4. IEEE, 2016.
31. Redouan Lahmyed, Mohamed El Ansari, and Ayoub Ellahyani. A new thermal infrared and visible spectrum images-based pedestrian detection system. *Multimedia Tools and Applications*, 78(12):15861–15885, 2019.
32. Redouan Lahmyed, Mohamed El Ansari, Alain Tremeau, and Zakaria Kerkaou. Camera-light detection and ranging data fusion-based system for pedestrian detection. *Journal of Electronic Imaging*, 27(6):063011, 2018.
33. Fredrik Larsson and Michael Felsberg. Using fourier descriptors and spatial models for traffic sign recognition. In *Scandinavian conference on image analysis*, pages 238–249. Springer, 2011.
34. JM Lillo-Castellano, I Mora-Jiménez, Carlos Figuera-Pozuelo, and José Luis Rojo-Álvarez. Traffic sign segmentation and classification using statistical learning methods. *Neurocomputing*, 153:286–299, 2015.
35. Huaping Liu, Yulong Liu, and Fuchun Sun. Traffic sign recognition using group sparse coding. *Information Sciences*, 266:75–89, 2014.
36. Yi Liu, Jie Ling, Qianhong Wu, and Bo Qin. Scalable privacy-enhanced traffic monitoring in vehicular ad hoc networks. *Soft Computing*, 20(8):3335–3346, 2016.
37. Yifeng Liu, Lin Zeng, and Yan Huang. An efficient hog–albp feature for pedestrian detection. *Signal, Image and Video Processing*, 8(1):125–134, 2014.
38. Ashkan Madadlou, Zahra Emam-Djomeh, Mohamad Ebrahimzadeh Mousavi, Mohamadreza Ehsani, Majid Javanmard, and David Sheehan. Response surface optimization of an artificial neural network for predicting the size of re-assembled casein micelles. *Computers and Electronics in Agriculture*, 68(2):216–221, 2009.
39. Jun Miura, Tsuyoshi Kanda, and Yoshiaki Shirai. An active vision system for real-time traffic sign recognition. In *ITSC2000. 2000 IEEE Intelligent Transportation Systems. Proceedings (Cat. No. 00TH8493)*, pages 52–57. IEEE, 2000.
40. Andreas Mogelmose, Mohan Manubhai Trivedi, and Thomas B Moeslund. Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey. *IEEE Transactions on Intelligent Transportation Systems*, 13(4):1484–1497, 2012.
41. Timo Ojala, Matti Pietikäinen, and David Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29(1):51–59, 1996.
42. KR Sri Preethaa and A Sabari. Intelligent video analysis for enhanced pedestrian detection by hybrid metaheuristic approach. *Soft Computing*, pages 1–9, 2020.
43. Andrzej Ruta, Yongmin Li, and Xiaohui Liu. Real-time traffic sign recognition from video by class-specific discriminative features. *Pattern Recognition*, 43(1):416–430, 2010.
44. Sanjit Kumar Saha, Dulal Chakraborty, and Md Al-Amin Bhuiyan. Neural network based road sign recognition. *International Journal of Computer Applications*, 50(10), 2012.
45. Samuele Salti, Alioscia Petrelli, Federico Tombari, Nicola Fioraio, and Luigi Di Stefano. Traffic sign detection via interest region extraction. *Pattern Recognition*, 48(4):1039–1049, 2015.
46. Pierre Sermanet and Yann LeCun. Traffic sign recognition with multi-scale convolutional networks. In *The 2011 International Joint Conference on Neural Networks*, pages 2809–2813. IEEE, 2011.
47. Eli Shechtman and Michal Irani. Matching local self-similarities across images and videos. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007.
48. Johannes Stallkamp, Marc Schlipsing, Jan Salmen, and Christian Igel. The german traffic sign recognition benchmark: a multi-class classification competition. In *The 2011 international joint conference on neural networks*, pages 1453–1460. IEEE, 2011.

49. Johannes Stallkamp, Marc Schlipsing, Jan Salmen, and Christian Igel. Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition. *Neural networks*, 32:323–332, 2012.
50. D Sudha and J Priyadarshini. An intelligent multiple vehicle detection and tracking using modified vibe algorithm and deep learning algorithm. *Soft Computing*, 24:17417–17429, 2020.
51. Zhan-Li Sun, Han Wang, Wai-Shing Lau, Gerald Seet, and Danwei Wang. Application of bw-elm model on traffic sign recognition. *Neurocomputing*, 128:153–159, 2014.
52. Jirí Trefnỳ and Jirí Matas. Extended set of local binary patterns for rapid object detection. In *Computer vision winter workshop*, pages 1–7, 2010.
53. Vladimir Naumovich Vapnik and Vlamimir Vapnik. *Statistical learning theory*, volume 1. Wiley New York, 1998.
54. Chuan Yang, Lihe Zhang, Huchuan Lu, Xiang Ruan, and Ming-Hsuan Yang. Saliency detection via graph-based manifold ranking. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3166–3173, 2013.
55. Jiaolong Yang, Wei Liang, and Yunde Jia. Face pose estimation with combined 2d and 3d hog features. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pages 2492–2495. IEEE, 2012.
56. Xing Yang. Enhancement for road sign images and its performance evaluation. *Optik*, 124(14):1957–1960, 2013.
57. Xue Yuan, Jiaqi Guo, Xiaoli Hao, and Houjin Chen. Traffic sign detection via graph-based ranking and segmentation algorithms. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 45(12):1509–1521, 2015.
58. Fatin Zaklouta and Bogdan Stanciulescu. Warning traffic sign recognition using a hog-based kd tree. In *2011 IEEE intelligent vehicles symposium (IV)*, pages 1019–1024. IEEE, 2011.
59. Fatin Zaklouta and Bogdan Stanciulescu. Real-time traffic-sign recognition using tree classifiers. *IEEE Transactions on Intelligent Transportation Systems*, 13(4):1507–1514, 2012.
60. Zhe Zhu, Dun Liang, Songhai Zhang, Xiaolei Huang, Baoli Li, and Shimin Hu. Traffic-sign detection and classification in the wild. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2110–2118, 2016.