TRAFFIC SIGN DETECTION AND RECOGNITION USING ADAPTIVE THRESHOLD SEGMENTATION WITH FUZZY NEURAL NETWORK CLASSIFICATION

Abdulrahman S Alturki
Department of Electrical Engineering
College of Engineering, Qassim University
Saudi Arabia
abdulrahman@qec.edu.sa

Abstract: Traffic Sign Recognition (TSR) system is a significant component of Intelligent Transport System (ITS) as traffic signs assist the drivers to drive more safely and efficiently. In this paper, a new traffic sign detection and recognition approach is presented by using Fuzzy Neural Network (FNN) and it is including three stages. The first stage segments the images to extract ROIs. The segmentation is usually performed based on Adaptive thresholding to overcome the color segmentation problems. The second one detects traffic shapes. Given that the geometric form of traffic signs is limited to triangular, circular, rectangular and octagonal forms, the geometric information is used to identify traffic shapes from ROIs provided by the first stage. The third stage recognizes the traffic signs based on the information including included in their pictograms. Moreover, in this work, six types of features are extracted. These features were provided to the FNN classifier to perform the recognition. As a classifier, FNN, Artificial Neural Network (ANN) and Support Vector Machine (SVM) classifiers have been tested together with the new descriptor. The proposed method has been tested on both the German Traffic Sign Detection and Recognition Benchmark dataset. The results obtained are satisfactory when compared to the state-of-the-art methods.

Keywords: Traffic sign, Intelligent Transport System, Fuzzy Neural Network, Adaptive thresholding, ANN, FNN.

1. INTRODUCTION

Traffic sign and classification is part of the Traffic Sign Recognition (TSR) system and it is one of the developed systems under Advance Driver Assistance (ADAS) which help to improve the safety issue on the road. ADAS play an important role in enhancing car safety and driving comfort. One of the most important difficulties that ADAS face is the understanding of the environment and guidance of the vehicles in real outdoor scenes [1]. Humans driving are a task based almost entirely on visual information, and one of the tasks in successful driving involves the identification of traffic signs. Traffic signs provide information about the current state of the road, restrictions, prohibitions, warnings, and other helpful information for navigation. The information provided by the road signs is encoded in their visual traits: shape, color and pictogram. Road sign recognition has been a challenge problem for many years and is an important task not only for ADAS, but also for other real-world applications including urban scene understanding, automated driving, or even sign monitoring for maintenance.

To improve the ADAS system performance, many different approaches to traffic sign recognition have been proposed and it is difficult to compare between those approaches since they are based on different data. Moreover, some articles concentrate on subclasses of signs, for example on speed limit signs and digit recognition. Regarding the detection problem, different approaches have been proposed. In the older studies, e.g. [1, 2], as well as in many recent ones, e.g. [3, 4], it was common to employ color segmentation [5].

Ruta et al. [6] used the Color Distance Transform, where a DT is computed for every color channel separately. Larsson et al. [7] used locally segmented contours combined with an implicit star-shaped object model as prototypes for the different sign classes. The contours are described by Fourier descriptors. Hough trans-form is another technique employed to detect shapes. In [8] a proprietary and undisclosed algorithm is used to detect rectangles, and Hough Transform for the detection of circles. Loy and Zelinsky[9] proposed a technique similar to Hough transform called fast radial transform, which was successfully used for sign detection in[10,11]. Many recent approaches use gradient orientation information in the detection phase, for example, in [12], Edge Orientation Histograms are computed over shape-specific sub-regions of the image.

After the localization of region of interests ROIs, classification techniques employed to determine the content of the detected traffic signs. Learning approaches are the most used techniques. Maldonado et al. in [13] utilized different one-vs-all Support

978-1-5386-3779-1/18/$31.00 ©2018 IEEE

Vector Machines (SVMs) with Gaussian kernel for each color and shape classification to recognize signs. In [14] SVMs are used with HOG features to carry out classification on candidate regions provided by the interest region detectors. It with stand great appearance variations thanks to the robustness of local features, which typically occur in outdoor data, especially dramatic illumination and scale changes. In [15], the authors suggest a hinge loss stochastic gradient descent method to train convolutional neural networks (CNNs).The method yields to high accuracy rates. However, a high computing cost is paid to train the data. Lim et al. in 16] used also Neural Networks (NNs), and improved their results by preselecting the color-shape features using Principal Components Analysis (PCA) and Fisher Linear Discriminant. Many other researchers use Nearest Neighbour approaches to classify traffic signs. For example, Kuo et al. in [17] used K-d tree to identify the content of the sign and yields to high accuracy rates. In [18], the identification of signs is carried out by a normalized correlation-based pattern matching using a traffic-sign database.

In general, the quality of the results obtained by any study on TSR varies from one research group to another. It is very difficult to decide which approach gives better overall results, mainly due to the lack of a standard database of road images. So, in this paper, a FNN classifier presented to recognize the traffic sign. The performance results show that the presented scheme attained better results compared than existing classifier schemes like ANN and SVM.

2. PROPOSED METHODOLOGY

In this section, each step of the proposed approach has been described.

As depicted in Fig. 1, the proposed method is achieved in three main steps. The first one segments the images to extract ROIs. The second one detects the shapes from the ROIs. The last step recognizes the information included in the detected traffic signs. In this process, initially, the input traffic image is acquired from different high way roads of Bangladesh. Then, it has been processed by three steps to recognize the signs. Finally, the performance results have been measured in term of detection accuracy.

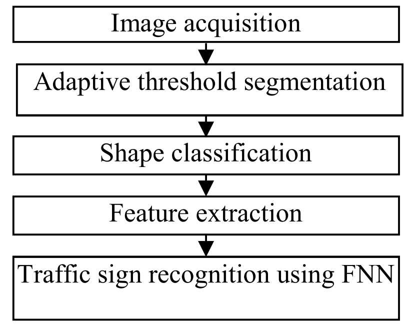

Figure 1: Overall process of presented scheme

Segmentation

Color segmentation algorithms are influenced by weather condition, day time, shadows, orientation of objects in relation to the sun and many other parameters [19]. These parameters change frequently in dense urban area scenes. In addition, there are many other objects in the street of the same color as traffic signs (red and blue). Therefore, the color information is only used to generate ROIs without performing classification.

To overcome the difficulties related to illumination changes and possible deterioration of the signs, the HSI color space is used in our system. Each image pixel is classified according to its hue, saturation, and intensity using adaptive thresholds for red and blue colors. To determine the adaptive threshold, we have used a dynamic threshold selection process (T_1 *and* T_2) by Eq. 1 and Eq. 2 based on Gray-threshold function.

$$T_1 = G_t(N) \quad (1)$$

Where G_t is defined as gray threshold

$$T_2 = G_t\big(N(N > T_1)\big) \quad (2)$$

Where N is defined as the normalization process

$$N = \frac{I - \min(I)}{\max(I) - \min(I)} \quad (3)$$

Where, the extracted three color channels represented by I. The segmentation step provides binary images where the ROIs are represented with white pixels. Insignificant ROIs are discarded based on size and aspect ratio constraints.

Shape classification

The approach used to classify shapes from extracted ROIs is described in this section. Most of the methods available in the literature call some classifiers, e.g. SVM, to detect the sign shape. Here, a simple invariant geometric moment's based method is used to achieve shape classification. Invariant moments are introduced in[24] where the famous Hu's seven invariant geometric moments were derived. Hu described his h1to h6moments as absolute orthogonal invariants (independent of position, size, and orientation) and h7as a skew orthogonal invariant (useful in distinguishing mirror images). These features are capable of recognizing simple objects.

The shapes need to be recognized are circles, triangles, and rectangles. They are all simple objects and we believe the invariant moments can help to recognize them perfectly. The ROIs are binary patches to provide to our shape classification system. The invariant moments are computed for each ROI and compared with those of the target patches of the possible three shapes. Binary patches of the different shapes are created. Note that the octagonal shapes are considered belonging to the same shape class as the circular ones.

Among the detected ROIs in the segmentation step, only those having their moments close to those of the target shapes will be considered as valid shape classes. Different metrics have been tested to match the ROIs with the appropriate shape classes. The metric used is defined as follows:

$$I(X,Y) = \sum_{i=1}^{7} |m_i^X, m_i^Y| \qquad (4)$$

Where m_i^X and m_i^Y are defined as follows

$$m_i^X = sign(h_i^X) \log |h_i^X| \qquad (5)$$
$$m_i^Y = sing(h_i^Y) \log |h_i^Y| \qquad (6)$$

where h$h_i^X$and $h_i^Y$are the values of the Hu moments of the ROI andthe patch respectively.

The metric I will be used to find correspondences between thedetected ROIs and the patches in Fig. 3. After computing metricsof the ROI over the three patches, the metric with the minimum value indicate the class of the shape. Note that a ROI is rejected if its corresponding metric value is above a threshold which was empirically derived based on collected images from the dataset.

Feature extraction

Once the candidate ROIs (blobs) are classified into a shape class they are provided to the recognition module in charge of identifying the sign. Most of the road signs contain a pictogram, a string of characters, or both. The recognition module is a classifier which should be fed with features describing the signs to identify. We have found out many features but chosen only six of them, as large number of input to FNN makes the computer system complex to train the FNN reading large sets of rules.

Total black pixel. As size of characters are different, so the number of black pixels(number of 0s) are also different.

$$\text{Total black pixel} = \text{row} * \text{column nnz(Image)} \qquad (7)$$

Where nnz is the number of nonzero elements

Entropy-Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image.

$$\text{Entropy} = \text{entropy(Image)} \qquad (8)$$

GLCM. A statistical method of examining texture that considers the spatial relationship of pixels is the gray-level cooccurrence matrix (GLCM). The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix. GLCM contains four measures of an image.

Contrast. Returns a measure of the intensity contrast between a pixel and its neighbor over the whole image.

Correlation. Returns a measure of how correlated a pixel is to its neighbor over the whole image.

Energy. Returns the sum of squared elements in the GLCM.

Homogeneity. Returns a value that measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal.

Recognition using FNN

A fuzzy Neural Network (FNN) consists of 4 layers [20-22] of functions for train the nodes to detect the malicious node and the functions are:

- Initial layer is input, based on the radial basis function NNs (RBFNNs) and is simply the nodes inputs out to the next layer;
- The second layer is a hidden layer and is fuzzifies the inputs, for example, fuzzy rule antecedents with a fuzzy truth for each node, and is obtained through passing each input value through a Fuzzy Set

Membership Function (FSMF) for a linguistic variable [22].

- Third is a rule layer, it implies a consequent fuzzy variable in this layer from certain fuzzifying nodes.
- Final layer is defuzzification [21] and is processed.

In FNN, use the simplified case of only two classes of correctly detected sign and non-corrected signs . FNN training process has been considered the above two classes. The proposed system has N number nodes in the input samples. Here, the two classes are there in the training example data $\{(x^{(q)}, t^{(q)}): q = 1, \ldots, Q\}$, i.e., the $t^{(q)}$ has two lables, as well as the proposed system use K=2 like k_1, k_2 class groups of hidden nodes, and these nodes represents a Gaussian function midpoint with a correlated label. In a class grouped, each Gaussian has a different center but the same label. The first group nodes are considered as class 1.

In this process, which node feature or behaviour is close to another node means that nodes are grouped with same label. It is based on the number of centers and thus Gaussians or behaviours of nodes is given by two classes. After that, fuzzy truth of input node x is in the same class as $x^{(q)}$ is given by the Gaussian FSMF centered on $x^{(q)}$ and is defined as

$$x \rightarrow g(x, x^{(q)}) = \exp\{-\|x - x^{(q)}\|^2 / (2\sigma^2)\} \qquad (9)$$

Where σ can be taken to be one-half of the average distance between all pairs.

Gaussian centers in class 1 are feed from their Gaussian input nodes to the maximize node of the class 1 fuzzy truths. These are performs as a fuzzy or node in choosing the agent or CH center and fuzzy truth. The x represents to some $x^{(k)}$ for Class 1. The maximum fuzzy truth of class 1 for x is send to the final output maximize node of agent. As well as the maximum fuzzy truth of class 2 for x is send to the final output maximize node of classification. As a result the input x belongs to the output class determined through the label of the output Gaussian center vector. Based on the above procedures the signs are correctly recognized and it measured in term of accuracy.

3. RESULTS AND DISCUSSION

In this segment, the performance of presented scheme has been evaluated and compared with existing algorithms like ANN [23] and SVM.

Dataset

The public available data sets called German Traffic Sign Recognition Benchmark (GTSRB), German Traffic Sign Detection Benchmark (GTSDB), and Swedish Traffic Signs (STS) are adopted for the performance evaluation. The GTSRB data set contains 51,839German traffic signs in 43 classes (39,209 training images and12,630 test images). These 43 classes of traffic signs have been divided into six subsets: speed limit, other prohibitory, derestriction, mandatory, danger, and unique signs Figure 2. The GTSDB dataset provides 900 full images (600 for training, 300 for testing).and the Swedish Traffic Signs Data sets.

Figure 2: Subsets of traffic signs in the GTSRB data set. (a) Speed limit signs. (b) Other prohibitory signs. (c) Derestriction signs. (d) Mandatory signs. (e) Danger signs. And (f)unique signs.

Segmentation and shape classification

In the following of the paper, we refer to shape detection stage as the segmentation step followed by the shape classification one of the proposed method for traffic sign detection and recognition. To evaluate the detection stage, we used two different data. The first one is formed from 300 images of the GTSDB data set. All the images were normalized to 640 × 480 pixels using bilinear interpolation.

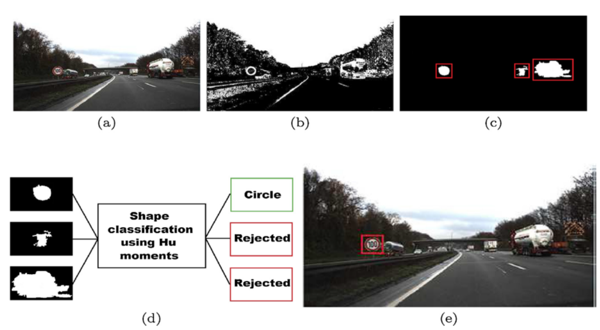

Figure 3: Step by step results of sign detection

Figure 3(a) shows an example among images used to test the proposed detection approach. The corresponding segmentation results with and without using size and aspect ratio constraints are illustrated in Figure 3(b) and (c), respectively. Figure 3(d) shows the shape classification results. The appropriate shape has been assigned to the first ROI, which is classified as circular. No shapes have been assigned to the two other ROIs. Figure 3(e) shows the final detection results by proposed detection method. Red bounding box represents detected region of traffic sign.

Figs. 4, 5 and 6 illustrate examples of recognition results when the proposed approach applied to images of various traffic environments. In Fig. 4, the traffic signs contained in the images have been successfully detected and recognized. In Fig. 5(a) and (b), the road signs were too far to be detected. After the color segmentation, they were discarded because they were not meet-ing the size constraint. In Fig. 5(c) and (d), the signs color in the images was changed due to the shadows. Consequently, the ROIs corresponding to the signs were not extracted by the segmentation method. In Fig. 6, the traffic signs contained in the images have been successfully detected. However, the system could not recognize them due to the motion blur in the signs.

Figure 4: Sign Detection and recognition results

Figure 5: Sing recognition with misdetection

Figure 6: Recognition with confused classification

Performance measures

A correct detected sign is considered true positive if the corresponding bounding box overlaps with at least 50% of the area covered by the right traffic sign present in the image. The evaluation of the detection accuracy is performed based on precision–recall curve, where the recall and precision values are computed as follows:

$$Recall = \frac{number\ of\ corretly\ detected\ signs}{number\ of\ true\ signs} \times 100 \qquad (10)$$

$$Precision = \frac{number\ of\ corretly\ detected\ signs}{number\ of\ detected\ signs} \times 100 \qquad (11)$$

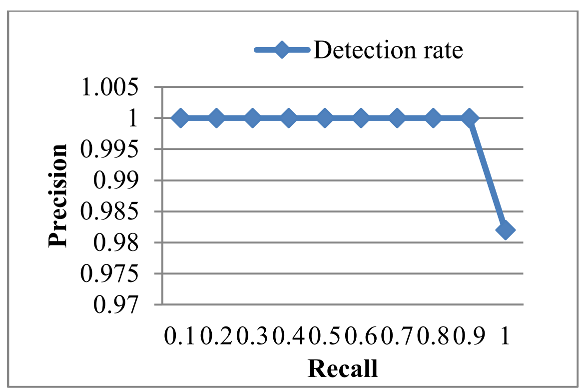

Figure 7: Precision–recall curves of the proposed FNN detection and recognition method

Table 1: Overall perfromance analysis among all sign recognition methods

Methods	FNN	ANN	SVM
Accuracy	98.2	97	95.6
Precision	100	98.2	96.5
Recall	97.45	96.3	94.2

From table 1, the proposed FNN attained better accuracy compared than other algorithms. Due to the less error rate, high positive detection rate, and effectual segmentation, the proposed FNN attained good results.

4. CONCLUSION

In this paper, FNN classified with a three stages system for real-time Traffic Sign Detection and Recognition has been presented. The first stage segments the images into ROIs based on color information. Only significant ROIs will be considered referred to their size and aspect ratio constraints. In the second stage, the circular, rectangular and triangular shapes are detected using invariant geometric moments. In the recognition stage, six features are extracted to classify the detection. The experimental results show that the proposed FNN has been attained 98.2% of accuracy compared than existing ANN and SVM algorithms due to the effectual segmentation and feature extraction.

References

1. A. De La Escalera, L.E. Moreno, M.A. Salichs, J.M. Armingol, Road traffic signdetection and classification, IEEE Trans. Ind. Electron. 44 (6) (1997) 848–859.
2. D.M. Gavrila, Multi-feature hierarchical template matching using distancetransforms, in: Fourteenth International Conference on Pattern Recognition,Proceedings, vol. 1, IEEE, 1998, pp. 439–444.
3. J. Lillo-Castellano, I. Mora-Jiménez, C. Figuera-Pozuelo, J. Rojo-Álvarez, Trafficsign segmentation and classification using statistical learning methods, Neu-rocomputing 153 (2015) 286–299.
4. C. Yao, F. Wu, H.-j. Chen, X.-l. Hao, Y. Shen, Traffic sign recognition using hog-svm and grid search, in: 12th International Conference on Signal Processing(ICSP), IEEE, 2014, pp. 962–965.
5. M. Bénallal, J. Meunier, Real-time color segmentation of road signs, in: CanadianConference on Electrical and Computer Engineering, IEEE CCECE 2003, Vol. 3,IEEE, 2003, pp. 1823–1826.
6. Ruta, Y. Li, X. Liu, Real-time traffic sign recognition from video by class-specific discriminative features, Pattern Recogn. 43 (1) (2010) 416–430.
7. F. Larsson, M. Felsberg, Using Fourier descriptors and spatial models for trafficsign recognition, in: Image Analysis, Springer, 2011, pp. 238–249.
8. F. Moutarde, A. Bargeton, A. Herbin, L. Chanussot, Robust on-vehicle real-timevisual detection of American and European speed limit signs, with a modulartraffic signs recognition system, in: Intelligent Vehicles Symposium, IEEE, IEEE,2007, pp. 1122–1126.
9. G. Loy, A. Zelinsky, Fast radial symmetry for detecting points of interest,IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (8) (2003)959–973.
10. G. Loy, N. Barnes, Fast shape-based road sign detection for a driver assistancesystem, in: IEEE/RSJ International Conference on Intelligent Robots and Sys-tems, (IROS 2004), Proceedings, vol. 1, IEEE, 2004, pp. 70–75.
11. C.F. Paulo, P.L. Correia, Automatic detection and classification of traffic signs, in:Eighth International Workshop on Image Analysis for Multimedia InteractiveServices, WIAMIS’07, IEEE, 2007, p. 11.
12. B. Alefs, G. Eschemann, H. Ramoser, C. Beleznai, Road sign detection from edgeorientation histograms, in: Intelligent Vehicles Symposium, IEEE, IEEE, 2007,pp. 993–998.
13. S. Maldonado-Bascón, S. Lafuente-Arroyo, P. Gil-Jimenez, H. Gómez-Moreno, F.López-Ferreras, Road-sign detection and recognition based on support vectormachines, IEEE Trans. Intell. Transp. Syst. 8 (2) (2007) 264–278.
14. S. Salti, A. Petrelli, F. Tombari, N. Fioraio, L. Di Stefano, Traffic sign detection viainterest region extraction, Pattern Recogn. 48 (4) (2015) 1039–1049.
15. J. Jin, K. Fu, C. Zhang, Traffic sign recognition with hinge loss trained con-volutional neural networks, IEEE Transactions on Intelligent TransportationSystems 15 (5) (2014) 1991–2000.
16. K.H. Lim, K.P. Seng, L.M. Ang, Intra color-shape classification for traffic signrecognition, in: International Computer Symposium (ICS), IEEE, 2010, pp.642–647.
17. W.-J. Kuo, C.-C. Lin, Two-stage road sign detection and recognition, in: IEEEInternational Conference on Multimedia and Expo, IEEE, 2007, pp. 1427–1430.
18. J. Miura, T. Kanda, Y. Shirai, An active vision system for real-time traffic signrecognition, in: Intelligent Transportation Systems, Proceedings, IEEE, IEEE,2000, pp. 52–57.
19. B. Soheilian, N. Paparoditis, B. Vallet, Detection and 3d reconstruction of trafficsigns from multiple view color images, ISPRS J. Photogr. Remote Sens. 77 (2013)1–20.

20. B. Gabrys and A. Burgiela, “General fuzzy min-max neural network for clustering and classification,” IEEE Trans. Neural Networks 11(3), 769-783, 2000.
21. H. M. Lee, C. M. Chen and Y. L. Jou, “An efficient fuzzy classifier with feature selection based on fuzzy entropy,” IEEE Trans. SMC-B 31(3), 426-432, 2001.
22. Langari, R., Wang, L., & Yen, J. (1997). Radial basis function networks, regression weights, and the expectation-maximization algorithm. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 27(5), 613-623.
23. Ellahyani, A., El Ansari, M., El Jaafari, I., & CHARFI, S. (2016). Traffic sign detection and recognition using features combination and random forests. International Journal of Advanced Computer Science and Applications, 7(1), 683-693.