

Article

Understanding Traffic Signs by an Intelligent Advanced Driving Assistance System for Smart Vehicles

Riadh Ayachi ^{1,*}, Mouna Afif ¹, Yahia Said ^{1,2}, Abdessalem Ben Abdelali ¹¹ Laboratory of Electronics and Microelectronics (E μ E), Faculty of Sciences of Monastir, University of Monastir, Monastir, Tunisia² Electrical Engineering Department, College of Engineering, Northern Border University, Arar, Saudi Arabia

*Correspondence: riadhayachi@fsm.rnu.tn

Abstract: Recent technologies have made life smarter. vehicles are vital components in daily life that are getting smarter for a safer environment. Advanced Driving Assistance Systems (ADAS) are widely used in today's vehicles. It has been a revolutionary approach to make roads safer by assisting the driver in difficult situations like collusion, or assistance in respecting road rules. ADAS is composed of a huge number of sensors and processing units to provide a complete overview of the surrounding objects to the driver. In this paper, we introduce a road signs classifier for an ADAS to recognize and understand traffic signs. This classifier is based on a deep learning technique, and, in particular, it uses Convolutional Neural Networks (CNN). The proposed approach is composed of two stages. The first stage is a data preprocessing technique to filter and enhance the quality of the input images to reduce the processing time and improve the recognition accuracy. The second stage is a convolutional CNN model with a skip connection that allows passing semantic features to the top of the network in order to allow for better recognition of traffic signs. Experiments have proved the performance of the CNN model for traffic sign classification with a correct recognition rate of 99.75% on the German traffic sign recognition benchmark GTSRB dataset.

Keywords: Intelligent Advanced Driving Assistance System, Deep Learning, Convolutional Neural Networks, Traffic Signs Classification

How to cite this paper : Ayachi, R., Afif, M., Said, Y., & Ben Abdelali, A. (2021). Understanding Traffic Signs by an Intelligent Advanced Driving Assistance System for Smart Vehicles. *Journal of Artificial Intelligence and Big Data*, 1(1). Retrieved from <https://www.scipublications.com/journal/index.php/jaibd/article/view/148>

Received: September 15, 2021

Accepted: October 16, 2021

Published: October 17, 2021



Copyright: © 2021 by the authors. Submitted for possible open access publication under the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Intelligent ADAS are used to assist drivers in difficult driving situations. indeed, the ADAS can assist them in the continuous control of the vehicle. In addition, the ADAS system can access the vehicle functionalities like engine and brakes, and apply numerous actions to avoid a dangerous situation like collisions with other vehicles or with pedestrians. Generally, an ADAS system is composed of several sensors (camera included). In the research presented here, the visual information provided by the cameras will be used to help the ADAS system to recognize and interpret correctly the road signs.

To process visual information, a well-known deep learning model was used as convolutional neural network [1], widely exploited in image processing tasks such as object recognition, image classification [2], and object localization. The success of CNN models has been achieved in various computer vision tasks [3,4]. Indeed, the CNN is characterized by the power of processing visual data that mimics the biological system. In a CNN every neuron of the network is applied in a restricted region of the receptive field [5]. The power of the CNN to learn directly from the image [6], unlike other classification algorithms that need a hand-crafted feature to learn from, makes it at the top of used algorithms for image processing tasks.

Enhancing deep learning algorithm performance is based on the amount of training data. For an artificial system, extensive computation effort are required for recognizing

and classifying a traffic sign. From one country to another, the shape and the color of the same road sign are different. Figure 1 presents the school sign in different countries (intra-class variability). Each country has different classes of signs (inter-class variability). Moreover, the same road sign may have a different appearance because of environmental factors such as dust, rain, sun, its observation point, etc. Therefore, the mentioned above challenges need to be solved successfully in order to obtain a robust road sign classifier, i.e., leading to the minimum rate of errors.



Figure 1. School signs in different countries

This paper introduces a two steps approach to traffic sign classification which solves all the above-mentioned problems and offers a high processing rate (up to 40 image/s).

The remainder of the paper is organized as follows. Section 2 introduces some related works on traffic sign classification. Section 3 outlines the novel two-step approach to traffic sign recognition. Section 4 provides the results of the evaluation of the proposed approach. Section 5 summarizes the paper and to proposes its extensions.

2. Related works

The need for a robust traffic sign classifier became an important part of ADAS. Fleyeh et al. [19] propose the use of the principal component analysis algorithm to classify traffic signs. Their method was evaluated on 2 different datasets and the highest accuracy achieved was 97.6%.

In [20], a traffic sign recognition method combines two functional modules. The first module is a feature extraction descriptor (histogram of oriented gradient (HOG)). the second module was the classifier based on the extreme learning machine. The proposed method achieves high performance in both recognition accuracy and computational efficiency.

Berkaya et al. [21] propose 2 traffic sign recognition methods for circular signs. The first method is based on a combination of feature extraction techniques (histogram of oriented gradients, local binary patterns, and Gabor features), and a classifier based on the support vector machines. The second method is based on a circle detection algorithm and an RGB-based color thresholding technique. Both techniques achieve acceptable accuracy, but the second method was suitable for real-time processing.

Gecer et al. [18] use a Combination of Shifted Filter Responses for a color-blob method to select circular regions from images (as a features extraction module) and use

the support vector machines as a classifier. This technique achieves an accuracy of 98.94% on the GTSRB dataset and 89.20% on the butterfly dataset.

Many works based on deep learning techniques were presented.

Rachmadi *et al.* [9] introduced a cascade Convolutional Neural Network to perform the classification of the Japanese road signs. The proposed method achieved an inference speed of less than 20 ms per image.

A Multi-Column deep neural network was proposed in [17]. It is a combination of many convolutional neural networks. Each convolutional neural network was trained on different preprocessed data. The proposed method results system is invariant to contrast and illumination changes.

Recently, vehicles are integrating new techniques for traffic signs classification; e.g. BMW has integrated a traffic sign recognition system in the BMW 5 Series [22]. Moreover, other vehicle manufacturers started to implement those technologies [10]. Volkswagen has integrated a traffic sign classifier into the Audi A8 [11].

The literature review on the traffic signs recognition proved that this technology is very important for ADAS systems or even for self-driving cars. The mentioned works are improved separately, either for accuracy or inference speed. The proposed method provides a good balance between accuracy and speed by achieving the highest classification accuracy and a good inference speed.

3. Proposed approach

The proposed method focuses on the balance between speed and recognition accuracy by deploying a data preparing technique to enhance the image quality and the use of a Convolutional Neural Network.

A two-steps approach to traffic sign classification encompasses: (1) data preparation pipeline to enhance data quality and to improve performance. (2) deep learning model to recognize and classify traffic signs. The data preprocessing method is composed of 4 steps:

1. First, loading the data and apply data augmentation techniques
2. Second, images resizing to 32*32 with 3 channel RGB color space and shuffled [12].
3. Third, a local histogram equalization [7,13] was applied;
4. Finally, the images were normalized to feed them to the proposed convolutional neural network.

For traffic sign recognition, a custom CNN model composed of 13 layers, 6 convolution layers, 3 max-pooling layers, 2 fully connected layers, and a softmax layer was proposed. The data preparation workflow is presented in Figure 2.

First, the data was loaded and new examples were generated using a data augmentation technique. The data augmentation technique was applied to maximize the amount of the training data. Second, input images were resized 32*32 in the RGB color space. Then, to avoid obtaining highly correlated mini-batches, the images were shuffled. So, the training algorithm obtains different mini-batch for each forward pass. Third, contrast of the images was enhanced using local histogram equalization [14].

Data Preparation



Figure 2. Data preparation process

Globally speaking, the proposed preparation allows to increase the contrast of areas with lower local contrast and makes global contrast higher. Finally, the data was normalized to make all examples at the same scale ensuring balanced representation of all features. Data preparation process enhances data quality for both training and testing processes.

Then a CNN model was proposed for traffic sign recognition. CNN is a deep learning model used for solving computer vision tasks. It has achieved a wide success in images classification task and it has the flexibility of reuse of existing models for new applications.

Inspired by the ResNet [8] that proposed implementing residual blocks instead of classic plain blocks, we proposed a custom CNN model for traffic sign recognition. A residual block consists of connecting previous layers with more than just the next layer. By using those residual blocks, ResNet achieved a top-5 error of 3.57% [8] in the ILSVRC 2015 [15].

For the proposed traffic sign recognition, a CNN model based on residual connection was proposed. Figure 3 presents the proposed architecture. The network is composed of 7 convolution layers. For the first convolution layer, 3 filters with 1x1 kernel size with a stride of 1 were used to learn features from the whole image. Those 3 filters are used to produce feature maps from each of the 3 channels RGB space color. The 1x1 kernel size was used to make a significant improvement to the network performance. For the other convolution layers, a 3x3 kernel is used starting with 32 filters and the number of filters was doubled after each pooling layer. 3x3 kernels were used because smaller kernel sizes result in better performance. For the pooling, the max-pooling with 3x3 kernel size and a stride of 2 was used. The flatten layer converts tensors to vectors to ensure the connection between the max-pooling layer and the fully connected layer. Two fully connected layers

summarized the features and learned global features. For the classification, the soft max layer was deployed to provide predictions. As a residual connection, all the max-pooling layers are connected to the fully connected layer to get features from all the stages of the network. For performance enhancement, the dropout technique was deployed after each max pooling layer and after the fully connected layers.

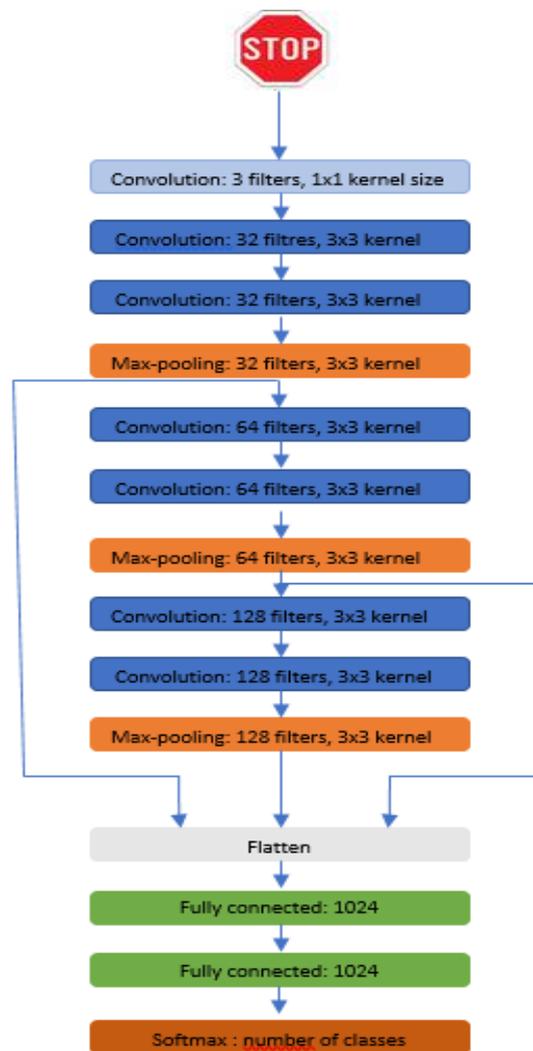


Figure 3. Proposed architecture

4. Results

Experiments were performed on a desktop with an Intel i7 CPU and a Nvidia GTX960 GPU. The network was built using the TensorFlow framework. For model training and testing, the German traffic signs dataset GTSRB [16] was used. The GTSRB dataset is traffic signs classification benchmark. The dataset was divided into training set and a testing set. It contains more than 50000 natural scenes images from 43 traffic signs classes.

To test in detail the system performance, various configurations were tested such as the image sizes, different batches, dropout rate, and different learning algorithms (optimizers).

At the start, the parameters of the experimental environment were set as follows: all images were resizing to 32*32, a batch of 1024 samples was used, a dropout rate of 0.25 was assumed and the stochastic gradient descent was selected as learning algorithm.

Once the training is finished and several tests were performed, the best configuration was fixed by an images size of 96*96, using a mini-batch of 256 samples, a dropout rate of 0.5, and the Adam optimizer for learning.

For model training and testing, we started by preprocessing the data and use at as input to the CNN model. First, the data was loaded, that the augmentation technique was applied. Next, the images were resized, shuffled to generate mixed mini-batches and the local histogram equalization was applied. At the final step of the preparing pipeline, the generated data was used for the training of the proposed CNN.

Table 1 summarize the achieved results compared to existing works repoted on the GTSRB dataset. as repoted, an accuracy of 99.75% was achieved. Compared to human accuracy, the achieved results is very acceptable. this result outperforms the state-of-the-art models in the traffic signs classification task.

Most of the false-negative predictions are due to totally or partially damaged images after the data pre-processing.

Table 1. The model’s sizes before and after applying the proposed approach

Network	Accuracy (%)
Committee of CNN [17]	99.4
Color-blob-based COSFIRE filters [18]	98.97
Proposed	99.75

For real-time implementation, a balancing speed and accuracy is important. The proposed CNN is suitable for implementation in real-time applications. Usually, an average of 25 images per second are required for real-time constraints. The traffic sign recognition system has achieved up to 143 FPS using high-performance GPU. So, switching to on low-performance hardware such as raspberry and achieving high performance in term of speed and accuracy is possible.

After the train and tests, the CNN model was implemented in a traffic sign classification application where the images were labeled with human-understandable labels. The application was tested using new images. Figure 4 presents an example of the application output of top five predictions.

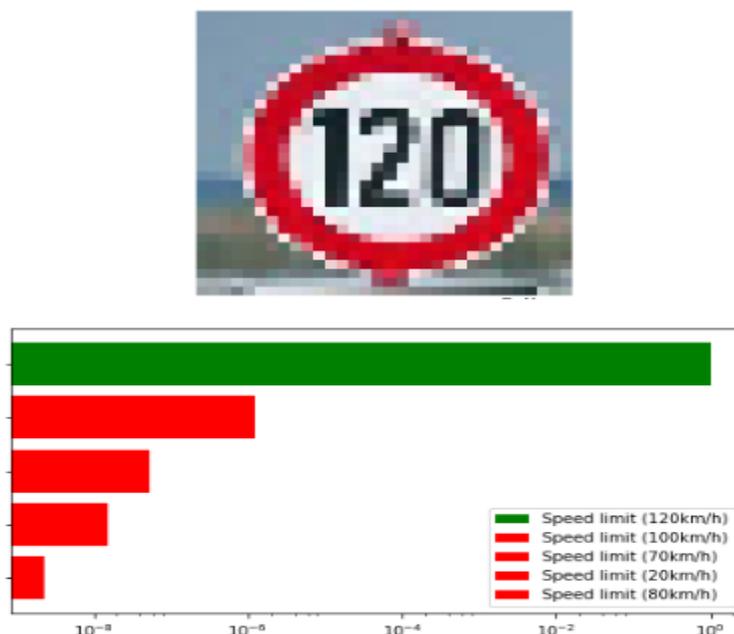


Figure 4. The output of the traffic sign application

5. Conclusions

This paper proposed a 2 stages method based on a data preparation and a CNN model for the traffic signs recognition. The data preparation stage enhances the image quality and improves the recognition step. The obtained results prove the efficiency the proposed approach by achieving real-time speed and high accuracy which outperforms human performance and state-of-the-art methods.

Author Contributions: Conceptualization, Riadh Ayachi.; methodology, Mouna Afif; software, Riadh Ayachi; validation, Riadh Ayachi; formal analysis, Mouna Afif; investigation, Mouna Afif; resources, Riadh Ayachi; data curation, Riadh Ayachi; writing—original draft preparation, Riadh Ayachi; writing—review and editing, Riadh Ayachi; visualization, Yahia said.; supervision, Abdesalem Ben Abdelali. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Data Availability Statement: In this section, please provide details regarding where data supporting reported results can be found, including links to publicly archived datasets analyzed or generated during the study.

Conflicts of Interest: The authors declare no conflict of interest.

References

- [1] O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458.
- [2] Ciresan, D. C., Meier, U., Masci, J., Maria Gambardella, L., & Schmidhuber, J. (2011, July). Flexible, high-performance convolutional neural networks for image classification, IJCAI Proc.-Int. Joint Conf. on Artificial Intelligence (Vol. 22, No. 1, p. 1237).
- [3] Convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).
- [4] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
- [5] Convolutional Neural Network (CNN). Available: <https://medium.com/@raycad.seedotech/convolutional-neural-network-cnn-8d1908c010ab>
- [6] LeCun, Y., Jackel, L. D., Bottou, L., Cortes, C., Denker, J. S., Drucker, H., ... & Vapnik, V. (1995). Learning algorithms for classification: A comparison on handwritten digit recognition. *Neural networks: the statistical mechanics perspective*, 261, 276
- [7] Zhu, H., Chan, F. H., & Lam, F. K. (1999). Image contrast enhancement by constrained local histogram equalization. *Computer vision and image understanding*, 73(2), 281-290.
- [8] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- [9] Rachmadi^{1/2}, R. F., Komokata^{1/2}, Y., Íchimura^{1/2}, K., & Koutaki^{1/2}, G. (2017). Road sign classification system using cascade convolutional neural network.
- [10] Continental. Traffic Sign Recognition. Available online: http://www.contionline.com/generator/www/de/en/continental/automotive/general/chassis/safety/hidden/verkehrszei_chenerkennung_en.html (accessed on 20 April 2017).
- [11] Choi, Y.; Han, S.I.; Kong, S.-H.; Ko, H. Driver status monitoring systems for smart vehicles using physiological sensors: A safety enhancement system from automobile manufacturers. *IEEE Signal Process. Mag.* 2016, 33, 22–34.
- [12] Dean, J., & Ghemawat, S. (2008). MapReduce: simplified data processing on large clusters. *Communications of the ACM*, 51(1), 107-113.
- [13] Kim, J. Y., Kim, L. S., & Hwang, S. H. (2001). An advanced contrast enhancement using partially overlapped sub-block histogram equalization. *IEEE transactions on circuits and systems for video technology*, 11(4), 475-484.
- [14] Abdullah-Al-Wadud, M., Kabir, M. H., Dewan, M. A. A., & Chae, O. (2007). A dynamic histogram equalization for image contrast enhancement. *IEEE Transactions on Consumer Electronics*, 53(2).
- [15] Deng, J., Berg, A., Satheesh, S., Su, H., Khosla, A., & Fei-Fei, L. (2012). ILSVRC-2012, 2012. URL [http://www. image-net.org/challenges/LSVRC](http://www.image-net.org/challenges/LSVRC).
- [16] J. Stallkamp, M. Schlipsing, J. Salmen, C. Igel, Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition, *Neural Networks*, Available online 20 February 2012, ISSN 0893-6080, 10.1016/j.neunet.2012.02.016. (<http://www.sciencedirect.com/science/article/pii/S0893608012000457>) Keywords: Traffic sign recognition; Machine learning; Convolutional neural networks; Benchmarking.
- [17] CireşAn, Dan, et al. "Multi-column deep neural network for traffic sign classification." *Neural networks* 32 (2012): 333-338.
- [18] Gecer, B., Azzopardi, G., & Petkov, N. (2017). Color-blob-based COSFIRE filters for object recognition. *Image and Vision Computing*, 57, 165-174.

-
- [19] Fleyeh, Hasan, and ErfanDavami. "Eigen-based traffic sign recognition." *IET Intelligent Transport Systems* 5, no. 3 (2011): 190-196.
 - [20] Huang, Zhiyong, Yuanlong Yu, Jason Gu, and Huaping Liu. "An efficient method for traffic sign recognition based on extreme learning machine." *IEEE transactions on cybernetics* 47, no. 4 (2017): 920-933.
 - [21] Berkaya, Selcan Kaplan, HuseyinGunduz, OzgurOzsen, CuneytAkinlar, and Serkan Gunal. "On circular traffic sign detection and recognition." *Expert Systems with Applications* 48 (2016): 67-75.
 - [22] Bui-Minh, Thanh, Ovidiu Ghita, Paul F. Whelan, Trang Hoang, and Vinh Quang Truong. "Two algorithms for detection of mutually occluding traffic signs." In *2012 International Conference on Control, Automation and Information Sciences (ICCAIS)*, pp. 120-125. IEEE, 2012.