Low-cost Deep Learning-based Prototype for Automatic Identification of Traffic Signs in Vehicles

Enol García González, José R. Villar, and Enrique de la Cal

Computer Science Department, University of Oviedo, Oviedo, Spain
enolgargon@gmail.es, {villarjose,delacal}@uniovi.es

Abstract. The automotive industry has evolved in recent years with new driving assistance systems, which make it evolve towards autonomous driving, in which the figure of the driver becomes less relevant over time until it becomes unnecessary to have a person driving the vehicle. As a contribution to this evolution towards autonomous driving, a low-cost prototype is presented, which can be installed in any type of car, capable of capturing images of driving by means of a camera and processing information from traffic signs on the road. This information can be used as input for another system, in which, thanks to this already processed signal information, it can make other types of driving decisions.

Keywords: Computer Based Driving Assistance, Traffic Signal Identification, Deep Learning.

1 Introduction

The automotive industry, like all industries, is in continuous evolution, seeking to improve vehicles by incorporating new technologies that are being discovered, researched, and improved. As part of this quest for continuous improvement, the automotive industry has been incorporating new safety and driving assistance systems in vehicles, such as:

- **Speed control**, which allows limiting or maintaining a constant speed in the vehicle without the driver needing to interact with the pedals. Some research on how to approach speed control can be found in [18, 9, 25, 3, 2, 26]
- **Line tracking**, which allows recognition of the lines on the road that delimit the lanes. This improvement is implemented in vehicles in two different ways:
  - *Unintentional lane change warning.* These systems aim at driver safety, alerting the driver when changing lanes if the system considers that the change is not intentional, but rather that it is a matter of absent-mindedness [8, 16, 1].
  - *Steering control.* This type of system has been under investigation for less time and is based on an intelligent system controlling the lateral movement of a vehicle, for example, to keep it in a lane. Some research in this field can be found in [13, 21].

– **Emergency braking**, This system is responsible for stopping the vehicle in the event of a dangerous or hazardous situation that requires it to stop. On this system, there are solutions with classic conditional logic and with artificial intelligence systems [23, 10, 19].

The combination and integration of these systems are moving the automotive industry towards what is known as autonomous driving. To properly define the term autonomous driving, the Society of Automotive Engineers (SAE) has published the J3016 standard[20], which refers to the 6 different levels into which a vehicle can be classified according to its ability to drive autonomously (see Figure 1):

– **Level 0**. This is the level of driving that has been occurring since the invention of the automobile. At this level, there is no system that intervenes in the driving and the driver is in charge of performing all the necessary actions on board the vehicle.
– **Level 1**. This level incorporates a driver assistance system, speed control. At this level, the driver can disengage from controlling the speed of the vehicle and focus on controlling the steering.
– **Level 2**. At this level, the driver can momentarily "disengage" from speed and direction control, and the vehicle can take over these two tasks. However, the driver is necessary to control the environment in which the vehicle is driving, as well as to give orders to the speed and steering systems so that the vehicle knows what to do and takes over in case of emergency.
– **Level 3**. At this level, the vehicle can control speed and direction and is also able to observe the environment in which it is driving to make decisions such as which exit to take at a traffic circle, what to do at a junction, overtaking, etc. The driver is still needed to take control in case of emergency, as he/she must be partially in control of what is happening on the road.
– **Level 4**. At this level, the driver's work becomes optional, since the vehicle is capable of performing all the necessary actions while driving, including making decisions in case of emergency. However, the driver can take control if he/she wishes.
– **Level 5**. At this level, there is no driver, as the vehicle will have neither pedals nor steering wheels. In no way can a person drive.

The present project is going to focus on the automatic identification of traffic signs. There are works based on different techniques both in the industry and in the literature: positioning-based systems, sensor-based systems. Positioning-based systems are the most common to be found nowadays and propose to use data labeled on digital maps[5, 12, 17] in order to obtain additional information about the road on which they run. This method has the disadvantage of labeling work and the risk of having an obsolescent map due to new legislation, signage or construction work.

Sensor-based systems obtain information from the road by means of sensors installed in the vehicle, which allows real-time information to be available. The

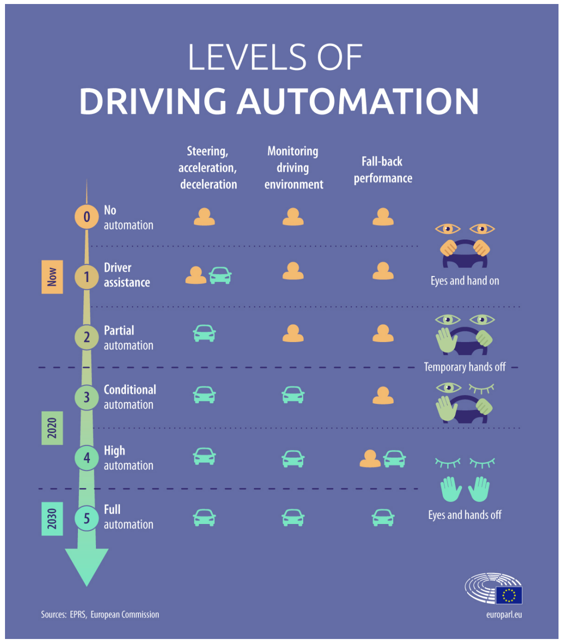

Fig. 1: Classification of autonomous driving levels. Currently, vehicles that are on the road are in levels 0, 1 and 2.

sensors most commonly used in the literature and which offer the most complete information are cameras [14, 6, 11, 15] and LIDAR sensors[27, 24, 7].

This project aims to generate a low-cost prototype to be implemented in a vehicle that, using a camera, captures the traffic signs present on the road, showing them to the user and allowing other driving assistance systems to consult and make use of this information. It is important that the system is fast enough to capture and process the information from a traffic sign before the vehicle reaches it. Deep Learning has been used as the modeling tool for signal identification.

The structure of this study is as follows. The next Section includes both the deployed hardware and the modeling issues. This section also deals with the data set that has been gathered for this research and with the experimental setup. Section 3 includes the results and the discussion that follows. The study ends with the conclusions drawn.

2 Materials and Methods

2.1 Deployed infrastructure

For the construction of the system, there are different alternatives to choose for its construction. It is necessary to select which product will be used for the following components of the system:

- **Hardware**. Among all the hardware alternatives that could have been used, Intel NUC was chosen because of its power and versatility. In addition, a USB camera and a small HDMI screen are also required to display the information.
- **Operating System**. GNU/Linux Ubuntu has been chosen as the operating system because it offers ease for multiprocess development and communication between several processes, besides being a free and open system that will not increase the cost of the prototype.
- **Development software**. For the implementation, the Python programming language will be used together with Keras, since it is a solution widely used in the literature.

From the design point of view, it is composed of 3 modules:

- **Camera control module**. This module will be in charge of taking snapshots of the camera periodically. These images will constitute the data input to the system.
- **System logic module**. This is the core of the system. It is composed of 3 activities that are executed for each input data in a correlative way. These activities are:
  1. *Preprocessing*. In this phase, the input image is transformed from RGB to Lab and the Gaussian and descend gradient transformation are performed on it to enhance the edges.
  2. *Segmentation*. Using the preprocessed image, the HoughtLines algorithm is applied in order to detect edges and extract possible signals. From here on, we will work with these clippings to determine whether or not it contains an image, and which image it is.
  3. *Identification*. This last phase uses the model to be defined below to determine which sign the image contains if any.
- **Output control module**. This module controls the output of the system and keeps a history of the processed data in a database.

These 3 modules described above must communicate and pass information to each other, as shown in Figure 2. As each module will be running in a different process of the Intel NUC, it is necessary to define a protocol for communication. The communication will be done by means of JSON messages transmitted through an IPC queuing system already implemented in the operating system. Each message will contain information about the sender, a title, a content –that could be an image, number, text–, a description of the content, and a reference to an image.

The deployment of the system is very simple. All the developed software will be installed on an Intel NUC, in which a camera is connected to take images and a screen to display the information.

2.2 Deep Learning models

The proposal included in this work is based on the DL topology presented in [4]. Ciresan's work proposed a DL architecture consisting of two parts: a first filter part and a second part that computes the results.

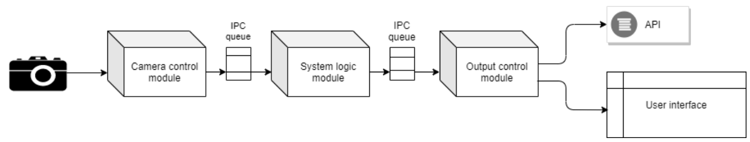

Fig. 2: Interfaces between subsystems

In the first filter part, there are 3 blocks in which each block is composed of 2 layers: one convolutional and one max pooling. In each block, the number of filters, both in the convolutional layer and in the Max Pooling layer, is increased, with filter values of 100, 150, and 250 for each block, respectively. In addition, the kernel of the convolutional layers is also reduced in each block.

The second part is in charge of the model output and is composed of 2 densely connected layers, the first one with 300 neurons and the second one with as many neurons as different signals it can classify. The latter has softmax activation. A graphical representation of the model construction can be seen in Figure 3.

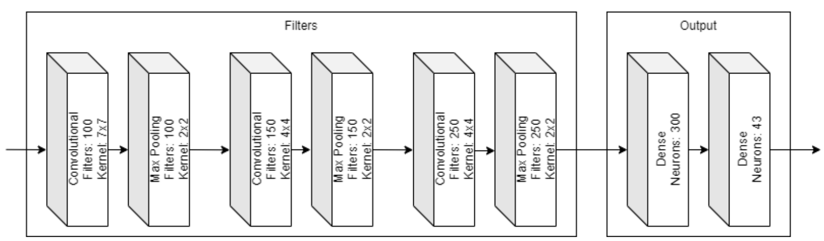

Fig. 3: Graphical representation of the network used

This topology is further implemented in the form of a Multicolumn Deep Neural Network (MCDNN). MCDNN are neural networks that are formed from the use of several deep neural networks that are placed in parallel, all these sub-networks process the model input and provide an output, which processed by a subsequent phase in which the model output is determined by a consensus algorithm.

Besides, the initialization of the weights has been performed with random values in the interval $[-0.5, 0.5]$. Furthermore, the majority algorithm was established as the consensus algorithm. In case of a tie, it will be determined that there is no signal. Figure 4 used includes two examples of how the consensus algorithm works.

2.3 Data set description

In the field of traffic signals, there are not a large number of datasets and all those found have the problem that they do not cover enough of the reality. For

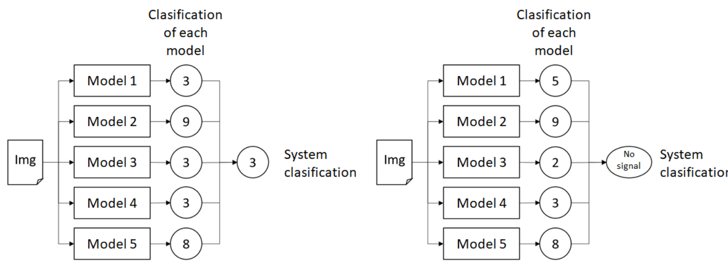

Fig. 4: Exemplification of the consensus algorithm used

a dataset to be considered very complete it must include instances that cover all the conditions that affect visibility:

- Meteorology
- Vegetation
- Time of day
- Additional illumination, such as streetlights or lights from other vehicles
- Infrastructure
- Other vehicles

By way of example, Figure 5 show the clear difference that exists when taking an image from the same point on the road between the morning of a rainy day and the afternoon of a sunny day.

(a) Image of the road on a sunny afternoon

(b) Image of the road on a rainy morning

Fig. 5: Comparison of images according to different environmental conditions

This project has used one of the most comprehensive datasets ever found in a search, consisting of 43 different types of traffic sign images captured in Germany [22]. To try to reduce the limitations of the dataset in terms of a variety of situations, Data Augmentation has been used, making zoom and rotation modifications on each image. In this way, the network can take the same image several times with different transformation settings.

2.4 Experimental setup

With the above dataset, training of 5 classification models is performed, which will be the ones that will be part of the classifier set. This training will be carried out by means of cross-validation with 10 groups and the duration of the training will be determined manually, analyzing the evolution of the accuracy of the dataset.

After the 5 models have been trained, they are merged in the classification system and deployed in the final model for the last test. Once the system was developed, 4 case studies were proposed to observe and evaluate its performance:

1. Incorporation to a conventional road
2. Driving on and exit from a highway.
3. Driving along the highway until it ends at a traffic circle.
4. Driving in the city

We will drive through the 4 scenarios proposed, recording the execution of the system and then analyzing the video and making a count of the images that appear in the video and that the system recognizes, those that it does not recognize, and those that it claims to recognize but do not exist in the real scenario.

3 Results and Discussion

3.1 Model training results

Following the established experimental protocol, 5 models were obtained that will form the classification set and the result of their metrics can be seen in Table 1.

Table 1: Accuracy of the trained nets

	Network 1	**Network 2**	**Network 3**	**Network 4**	**Network 5**
Mean in cross validation	87.851 %	87.516 %	89.807 %	87.322 %	87.962 %
Cross-validation variance	0.1394	0.1462	0.1458	0.1488	0.1429
Test value obtained	93.666 %	93.381 %	92.391 %	94.394 %	92.352 %

3.2 Validation with the physical deployment

As mentioned above, the last step will be to test the operation in a real scenario, using the four scenarios that have been proposed. Figure 6 shows a capture of the prototype's output in each of the 4 planted scenarios.

With this test, it was determined that the system was able to recognize all the signals present on the road, but it is not able to determine whether these

(a) Image taken from the first test case

(b) Image taken from the second test case

(c) Image taken from the third test case

(d) Image taken from the fourth test case

Fig. 6: Screenshots taken during prototype testing.

signals affect the vehicle or not, since it detects signals from other lanes or from other roads on which the vehicle is not traveling, but which it is able to see. In addition, the system also has false positives, detecting signals that do not exist in reality.

4 Conclusions

With this project, a prototype has been designed and developed for the detection of traffic signs on the road. The prototype has a good hit rate with the images received as long as they are captured in good environmental conditions.

To improve the performance of the system, it is proposed as a future study the development of a dataset with a large number of data, including different environmental conditions that may affect the recognition of images.

Acknowledgements

This research has been funded by the Spanish Ministry of Science and Innovation under project MINECO-TIN2017-84804-R, PID2020-112726RB-I00.

References

1. Ambarak, J.M., Ying, H., Syed, F., Filev, D.: A neural network for predicting unintentional lane departures. In: Proceedings of the IEEE International Confer-

ence on Industrial Technology. pp. 492–497. Institute of Electrical and Electronics Engineers Inc. (apr 2017). https://doi.org/10.1109/ICIT.2017.7913281

2. Chen, X., Leng, S., He, J., Zhou, L.: Deep Learning Based Intelligent Inter-Vehicle Distance Control for 6G-Enabled Cooperative Autonomous Driving. IEEE Internet of Things Journal (2020). https://doi.org/10.1109/JIOT.2020.3048050
3. Cheon, K., Kim, J., Hamadache, M., Lee, D.: On Replacing PID Controller with Deep Learning Controller for DC Motor System. Journal of Automation and Control Engineering **3**(6), 452–456 (2015). https://doi.org/10.12720/joace.3.6.452-456
4. Cireşan, D., Meier, U., Masci, J., Schmidhuber, J.: Multi-column deep neural network for traffic sign classification. Neural Networks **32**, 333–338 (aug 2012). https://doi.org/10.1016/j.neunet.2012.02.023
5. Fairfield, N., Urmson, C.: Traffic light mapping and detection. In: Proceedings - IEEE International Conference on Robotics and Automation. pp. 5421–5426 (2011). https://doi.org/10.1109/ICRA.2011.5980164
6. Geng, K., Dong, G., Yin, G., Hu, J.: Deep Dual-Modal Traffic Objects Instance Segmentation Method Using Camera and LIDAR Data for Autonomous Driving. Remote Sensing **12**(20), 3274 (oct 2020). https://doi.org/10.3390/rs12203274, https://www.mdpi.com/2072-4292/12/20/3274
7. Heinzler, R., Schindler, P., Seekircher, J., Ritter, W., Stork, W.: Weather influence and classification with automotive lidar sensors. In: IEEE Intelligent Vehicles Symposium, Proceedings. vol. 2019-June, pp. 1527–1534. Institute of Electrical and Electronics Engineers Inc. (jun 2019). https://doi.org/10.1109/IVS.2019.8814205
8. Jonsson, R., Kollmats, A.: Unintended Lane Departure Prediction using Neural Networks Master's thesis in System, Control and Mechatronic / Complex Adaptive Systems. Ph.D. thesis, Chalmers, Gothernburg (2018)
9. Jung, J.W., Leu, V.Q., Do, T.D., Kim, E.K., Choi, H.H.: Adaptive PID speed control design for permanent magnet synchronous motor drives. IEEE Transactions on Power Electronics **30**(2), 900–908 (feb 2015). https://doi.org/10.1109/TPEL.2014.2311462
10. Kaempchen, N., Schiele, B., Dietmayer, K.: Situation assessment of an autonomous emergency brake for arbitrary vehicle-to-vehicle collision scenarios. IEEE Transactions on Intelligent Transportation Systems **10**(4), 678–687 (dec 2009). https://doi.org/10.1109/TITS.2009.2026452
11. Kumar, V.R., Klingner, M., Yogamani, S., Milz, S., Fingscheidt, T., Mäder, P.: SynDistNet: Self-Supervised Monocular Fisheye Camera Distance Estimation Synergized with Semantic Segmentation for Autonomous Driving. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). pp. 61–71 (2021)
12. Kumpakeaw, S., Dillmann, R.: Semantic road maps for autonomous vehicles. In: Informatik aktuell. pp. 205–211. Springer, Berlin, Heidelberg (2007). https://doi.org/10.1007/978-3-540-74764-2_32
13. Li, D., Zhao, D., Zhang, Q., Chen, Y.: Reinforcement Learning and Deep Learning based Lateral Control for Autonomous Driving (oct 2018), http://arxiv.org/abs/1810.12778
14. Li, P., Qin, T., Shen, S.: Stereo Vision-based Semantic 3D Object and Ego-motion Tracking for Autonomous Driving. In: Proceedings of the European Conference on Computer Vision (ECCV). pp. 646–661 (2018)
15. Liu, L., Li, H., Dai, Y., Pan, Q.: Robust and efficient relative pose with a Multi-Camera system for autonomous driving in highly dynamic environments. IEEE Transactions on Intelligent Transportation Systems **19**(8), 2432–2444 (aug 2018). https://doi.org/10.1109/TITS.2017.2749409

16. Liu, X., Jin, G., Wang, Y., Yin, C.: A Deep Learning-based Approach to Line Crossing Prediction for Lane Change Maneuver of Adjacent Target Vehicles. In: 2021 IEEE International Conference on Mechatronics (ICM). pp. 1–6. IEEE (mar 2021). https://doi.org/10.1109/ICM46511.2021.9385665, https://ieeexplore.ieee.org/document/9385665/
17. Maturana, D., Chou, P.W., Uenoyama, M., Scherer, S.: Real-time semantic mapping for autonomous off-road navigation. In: Hutter, M., Siegwart, R. (eds.) Field and Service Robotics. pp. 335–350. Springer International Publishing, Cham (2018)
18. Nie, L., Guan, J., Lu, C., Zheng, H., Yin, Z.: Longitudinal speed control of autonomous vehicle based on a self-adaptive PID of radial basis function neural network. In: IET Intelligent Transport Systems. vol. 12, pp. 485–494. Institution of Engineering and Technology (aug 2018). https://doi.org/10.1049/iet-its.2016.0293
19. Rosén, E.: Autonomous Emergency Braking for Vulnerable Road Users. In: IRCOBI Conference. pp. 618–627 (2013), http://www.ircobi.org/wordpress/downloads/irc13/pdf_files/71.pdf
20. SAE International: Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems. Tech. rep., On-Road Automated Driving (ORAD) committee (2014). https://doi.org/https://doi.org/10.4271/J3016_201401, https://saemobilus.sae.org/content/j3016_201401
21. Sharma, S., Tewolde, G., Kwon, J.: Behavioral Cloning for Lateral Motion Control of Autonomous Vehicles Using Deep Learning. In: IEEE International Conference on Electro Information Technology. vol. 2018-May, pp. 228–233. IEEE Computer Society (oct 2018). https://doi.org/10.1109/EIT.2018.8500102
22. Stallkamp, J., Schlipsing, M., Salmen, J., Igel, C.: The German Traffic Sign Recognition Benchmark: A multi-class classification competition. In: The 2011 International Joint Conference on Neural Networks. pp. 1453–1460. IEEE (jul 2011). https://doi.org/10.1109/IJCNN.2011.6033395, http://ieeexplore.ieee.org/document/6033395/
23. Yuan, Y., Zhang, J.: A Novel Initiative Braking System with Nondegraded Fallback Level for ADAS and Autonomous Driving. IEEE Transactions on Industrial Electronics **67**(6), 4360–4370 (jun 2020). https://doi.org/10.1109/TIE.2019.2931279
24. Zeng, Y., Hu, Y., Liu, S., Ye, J., Han, Y., Li, X., Sun, N.: RT3D: Real-time 3-D vehicle detection in LiDAR point cloud for autonomous driving. IEEE Robotics and Automation Letters **3**(4), 3434–3440 (oct 2018). https://doi.org/10.1109/LRA.2018.2852843
25. Zhang, G., Furusho, J.: Speed control of two-inertia system by PI/PID control. IEEE Transactions on Industrial Electronics **47**(3), 603–609 (2000). https://doi.org/10.1109/41.847901
26. Zhang, Y., Sun, P., Yin, Y., Lin, L., Wang, X.: Human-like Autonomous Vehicle Speed Control by Deep Reinforcement Learning with Double Q-Learning. In: IEEE Intelligent Vehicles Symposium, Proceedings. vol. 2018-June, pp. 1251–1256. Institute of Electrical and Electronics Engineers Inc. (oct 2018). https://doi.org/10.1109/IVS.2018.8500630
27. Zhou, L., Deng, Z.: Lidar and vision-based real-time traffic sign detection and recognition algorithm for intelligent vehicle. In: 17th International IEEE Conference on Intelligent Transportation Systems (ITSC). pp. 578–583 (2014). https://doi.org/10.1109/ITSC.2014.6957752