Passengers’ Emotions Recognition to Improve Social Acceptance of Autonomous Driving Vehicles

Passengers' Emotions Recognition to Improve Social Acceptance of Autonomous Driving Vehicles

Jacopo Sini(✉)1[0000-0002-2163-9925], Antonio Costantino Marceddu1[0000-0003-4843-1267], Massimo Violante1[0000-0002-5821-3418], and Riccardo Dessì[2]

[1] Politecnico di Torino, Torino, Italy
{jacopo.sini, antonio.marceddu, massimo.violante}@polito.it
[2] General Motors – Global Propulsion Systems, Torino, Italy
riccardo.dessi@gm.com

Abstract. Autonomous driving cars hopefully could improve road safety. However, they pose new challenges, not only on a technological level but also from ethical and social points of view. In particular, social acceptance of those vehicles is a crucial point to obtain a widespread adoption of them. People nowadays are used to owning manually driven vehicles, but in the future, it will be more probable that the autonomous driving cars will not be owned by the end users, but rented like a sort of driverless taxis. Customers can feel uncomfortable while riding an autonomous driving car, while rental agencies will need to differentiate the services offered by their fleets of vehicles. If people are afraid to travel by these vehicles, even if from the technological point of view they are safer with respect to the manually driven ones, customers will not use them, making the safety improvements useless. To prevent the occupants of the vehicle from having bad feelings, the proposed strategy is to adapt the vehicle driving style based on their moods. This requires the usage of a neural network trained by means of facial expressions databases, of which there are many freely available online for research purposes. These resources are very useful, but it is difficult to combine them due to their different structures. To overcome this issue, a tool designed to uniform them, in order to use the same training scripts, and to simplify the application of commonly used postprocessing operations, has been implemented.

Keywords: Autonomous Driving Vehicle, Emotion Recognition, Neural Network, Social Acceptance.

1 Introduction

The development of autonomous driving vehicles is a challenging activity, with technological, ethical, and social implications. The main driver of this development is an expected increase in road safety [1]: the idea is that such vehicles, equipped with complex sensors like LIDAR, RADAR, cameras and so on, will be more able, with respect to human drivers, to avoid (or at least to mitigate the consequences if inevitable) crashes.

People used cars for over a hundred years, and during all this time period the vehicles have been operated by human drivers. Due to this habit, it is simpler for people to trust another person with respect to a computer to drive their vehicles. Moreover, we expect that drivers adopt different styles depending on their emotions [2]. So, it is important to take into account their social implications in order to avoid making the expected safety improvement useless due to a lack of people's trust on those vehicles. These regards ethic issues, social trust on autonomous driving capabilities, and novel commercialization model for the car manufacturers. In this paper, we would like to analyze the problem and propose an idea to improve the trust in these new driverless cars.
We wish to draft a preliminary workflow to perform autonomous vehicles passengers' emotions detection, in order to improve their comfort and so their trust. From a technical perspective, we would like to present a software tool to simplify the training of a neural network that, starting from the passengers' faces pictures, recognize their emotions. It could be used to adapt consequentially the autonomous cars driving style. As evidenced by various studies [3], positive emotions would consider faster or riskier driving as more favorable than people with negative emotions so, regarding the driving style adaptation, we can make more precise clarifications about what we wish to achieve:

- if the people on the car are scared or sad, the car adopts a caring driving style, that slows the speed and takes the curve more accurately (lowering the lateral accelerations);
- if the people on the car are neutral, the car adopts a normal driving style;
- if the people on the car are happy or over enjoyed, the car adopts a sportive driving style (steeper acceleration/braking ramps and curve trajectory with a higher level of lateral accelerations).

To obtain these results, as the first step in the development of the approach, a novel software has been developed. It is designed to allow usage of various third-party facial expressions databases for the training, the validation, and the testing of the emotions recognition neural network. It also simplifies the applications of commonly used post-processing operations on the images, like grayscale conversion and histogram equalization.

2 State of art

In recent years the miniaturization, which allows the creation of even smaller devices, the spread of smartphones with better and better cameras, the continuous improvement of wireless connections, and the advent of neural networks and other technologies have opened the door to the use of these techniques for different purposes. These devices can be easily embedded inside vehicles in a cost-effective way.
Neural networks have been used in recent years to solve various problems. For example, the LeNet-5 network [4] proposed by Yann LeCun, Joshua Bengio and Patrick Haffner has been used to automatically read the digits on bank cheques in the United States and is still used nowadays in the industry. Other examples of these technologies are the artificial intelligence developed by DeepMind [5], a subsidiary of Alphabet

Inc.'s Google, like AlphaGo, AlphaZero, and the last AlphaStar, that was capable of beating different professional StarCraft players.

Around the 1970s, psychologist Paul Ekman [6] identified six basic emotional states common to all cultures: anger, disgust, fear, happiness, sadness, and surprise. Further emotional states were subsequently theorized and then extended by Paul Ekman himself and other researchers [7, 8, 9]. For the sake of this paper, we considered these six states with the addition of the neutrality and contempt states.

Starting from the 2000s, several databases depicting human faces have been published with the aim of improving the algorithms for identifying facial expressions [18].

These databases are composed of posed and non-posed photos, and sometimes show additional data such as, for example, the Facial Action Coding System (FACS) coding [10, 11] and the Active Appearance Model (AAM) facial landmarks [12, 13].

Various authors [14, 15] described in detail that the interactions between cars, drivers, passengers and the ownership of cars are really complex [16]. Autonomous driving will put in place a new era in car ownership and usage, so car manufacturers need to understand how passengers will want to use that kind of vehicles. Other than that, to make things even more complex, it is reasonable to expect a time (from 10 to 30 years) in where driver-owned manually driven cars coexist and compete on the market with rented autonomous driving vehicles.

A lot of effort has been spent in the development of autonomous driving cars, considered in both owners, regulatory bodies and passengers' perspective [1]. But some questions remain open. In the future, the customers will want to own a car or they will become an on-demand transportation service only like taxis, trains or airplanes? If the answer is the second one, how can different carmakers and their new business-to-business customers, in this case transport companies and not the passengers, who will hold autonomous cars fleet, differentiate from each other in order to reach more final users? And finally, how much the customers be willing to pay this kind of service? At the moment, it is difficult to forecast answers to these questions, so they need to invest into the user experience of passengers, as much as airlines and railways companies did.

We claim that an enabling technology to improve the customers experience could be a properly trained neural network able to recognize the passengers' emotions. In this way, it will be possible to modify the vehicle behavior based on the passengers' feelings detected during the ride.

3 Proposed approach

Our proposal is to use cameras mounted in front of each seat of the car, and a properly trained neural network in order to adapt the vehicle driving style basing on the passengers' emotions. We claim that in this way it will be possible to improve the passengers' trustiness on these vehicles, hence their social acceptance. To avoid privacy issues, all the related emotions detection computations will be executed onboard thanks to a dedicated ECU, without any transmission to the external world. Moreover, the algorithm does not need to store into a permanent memory the frames representing the passengers' faces. This device is not directly involved in safety-critical tasks, so it can be implemented in a cost-effective way. The onboard system will be composed of:

- micro-cameras, placed in front of each seat to take pictures of the passengers' faces;
- a neural network, to recognize the emotions of each passenger;
- a decision algorithm, that determines the vehicle configuration on the base of the recognized passengers' emotions;
- a centralized emotions detector and control management computer, to tune vehicle parameters dynamically to improve the passengers' experiences.

The workflow to obtain the dataset needed to train this neural network is the following. Firstly, it is necessary to obtain the databases needed to train the neural network. There are many databases containing facial expressions that can be used for research purposes, which will be discussed later.

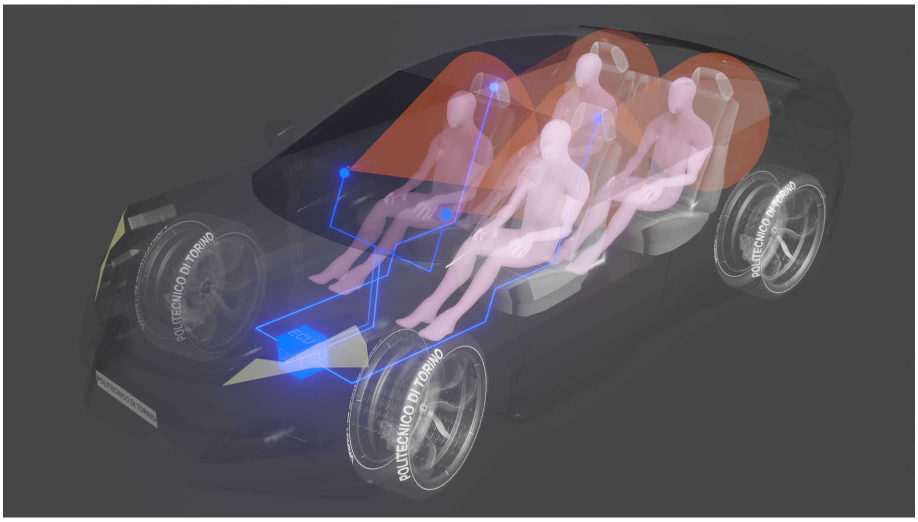

Fig. 1. Graphical representation of the proposed system

After this step, it is possible to start with the training, validation, and testing of the neural network. After that a sufficient emotions recognition accuracy has been achieved, a test campaign on a real vehicle can be started, in order to verify the effectiveness of the network in different lighting conditions (sun position, shadows due to the frame of the car, night driving) and in more realistic scenarios.

The last part of the work will regard how to properly manage the driving style depending on the detected feelings.

For the sake of this paper, we want to focus on the first phase, regarding the datasets preparation activities. To improve the effectiveness of this phase, we chose to develop a novel software that we called Facial Expressions Databases Classifier (FEDC).

FEDC is a program able to automatically classify images of some of the most used databases, depicting posed human faces:

- Extended Cohn-Kanade Database (CK+) [17, 18];

- Facial Expression Recognition 2013 Database (FER2013) [19];
- Japanese Female Facial Expression (JAFFE) [20];
- Multimedia Understanding Group Database (MUG) [21];
- Radboud Faces Database (RaFD) [22].

In all databases, emotions were classified according to Ekman's framework, but from a technical point of view, each author has adopted a different structure to organize the images. Since we would like to use pictures from multiple databases, without having to adapt the neural network training script for each of these, we have chosen to develop a novel program to uniform the classifications.

Since FEDC exploits the labels provided by the databases' creators, the accuracy relies only upon the classifications performed by them. In addition to this, it is also able to do several useful operations on the images, in order to simplify the neural network training operations and enhance its performances:

- scaling of the horizontal and vertical resolutions;
- conversion in grayscale color space;
- histogram equalization;
- face detection to crop the images to face only.

This allows, for the people who make use of these databases, to minimize the time necessary for their classification, so that they can dedicate directly to other tasks, such as training of a neural network.

3.1 FEDC description

FEDC has a clean and essential user interface, consisting of four macro areas:

- in the left column, it is possible to choose the database to be classified;
- in the right column, it is possible to select the operations to be performed on the photos: those available have already been mentioned previously;
- in the lower part of the window, there are the buttons to select the input file, the output folder, and to start and cancel the classification;
- finally, above the buttons, there is the progress bar, that indicates the progress of the current operation.

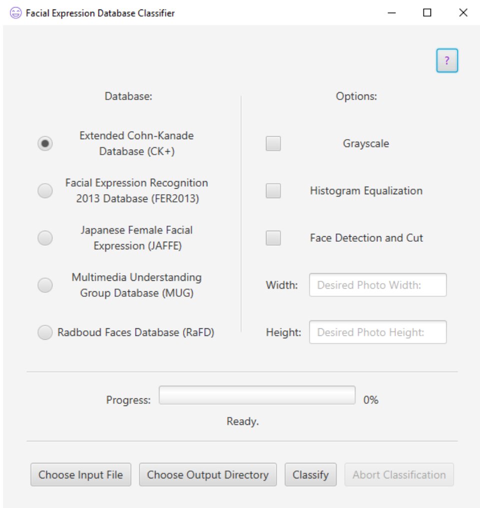

Fig. 2. Graphical User Interface of FEDC

It should be noted that:

- the user must choose a size for the photos to be classified: it must be between 48x48 and 1024x1024 pixels. For the FER2013 database, since the starting images have a 48x48 pixels resolutions, this possibility, alongside to the face cropping feature, is not available;
- the JAFFE and the FER2013 databases contain grayscale-only images;
- the RaFD database also contains photos taken in profile: the program excels in the recognition of frontal photos and allows recognition to be made even for profile photos, although it is likely that it will not be able to classify all these photos.

The images automatically classified with this program can be used, for example, for training a neural network with Keras [23] or similar frameworks: using Python with the Scikit-learn library [24], these images can be subdivided in the training, validation and test datasets or cross-validated.

Antonio Costantino Marceddu, that is one of the authors of this paper, developed FEDC resorting to Eclipse, with Java and the addition of the OpenCV framework [25], and released his software under the MIT license on GitHub [24].

4 Conclusions

Social acceptance implications on autonomous driving cars have to be considered in autonomous driving vehicles development. To achieve this goal, customers' trustiness plays a key role. In this paper, a possible improving methodology focused on the adaptation of the vehicle driving style, based on the passengers' emotions have been proposed.

At the moment, the effort of the authors is directed on the first stage of the methodology development, which is the data preparation for the training, validation, and test of the neural network, using pictures from the various databases of facial expressions available online for research purposes. Once that the neural network is trained, it will be possible to perform tests on cars in order to verify the emotions detection algorithm capabilities from the pictures captured with onboard cameras, in grayscale, in daylight or IR at night. When an acceptable matching rate will be reached, the next steps will be the development of the algorithms for car parameters tuning and for choosing the prevailing emotion, since the passengers can have conflicting moods. The last point will be the finding of an acceptable frequency to update the driving style.

References

1. Burns L.D. (2013), "Sustainable mobility: A vision of our transport future" In: Nature 497(7448), 181-182.
2. Trògolo A.M., Melchior F,., Medrano L.A. (2014), "The Role of Difficulties in Emotion Regulation on Driving Behavior", Journal of Behaviour, Health & Social Issues, DOI: 10.5460/jbhsi.v6.1.47607
3. Hu T, Xie X., Lee, J., (2013) "Negative or Positive? The effect of emotion and mood on risky driving", DOI: 10.1016/j.trf.2012.08.009
4. Y. LeCun, L. Bottou, Y. Bengio and P. Haffner: Gradient-Based Learning Applied to Document Recognition, Proceedings of the IEEE, 86(11):2278-2324, November 1998.
5. DeepMind company website, https://deepmind.com/, last accessed 2019/05/10
6. Ekman, P. & Friesen, W. V. (1971). Constants Across Cultures in the Face and Emotion. Journal of Personality and Social Psychology, 17(2), 124-129.
7. Ekman, P. (1999). Basic Emotions. In Dalgleish, T. & Power, M. J. (Eds.), Handbook of Cognition and Emotion (pp. 45-60). New York, NY: John Wiley & Sons Ltd.
8. Ekman, P., Cordaro, D. (2011). What is meant by calling emotions basic. Emotion Review, 3(4), 364-370.
9. Cordaro, D. T., Sun, R., Keltner, D., Kamble, S., Huddar, N., & McNeil, G. (2018). Universals and cultural variations in 22 emotional expressions across five cultures. *Emotion, 18*(1), 75-93.
10. P. Ekman and W. Friesen. Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto, 1978.
11. Paul Ekman, Wallace V. Friesen, and Joseph C. Hager. Facial Action Coding System: The Manual on CD ROM. A Human Face, Salt Lake City, 2002.
12. G.J. Edwards, C.J. Taylor and T.F. Cootes, "Interpreting face images using active appearance models", Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition, pp. 300-305, 1998.
13. T.F. Cootes, G.J. Edwards, and C.J. Taylor. "Active Appearance Models", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 6, pp. 681–685, 2001.
14. Fraedrich E., Lenz B, "Societal and Individual Acceptance of Autonomous Driving", In: Autonomous Driving, DOI 10.1007/978-3-662-48847-8_29 (2016)
15. Fraedrich E., Lenz B., "Taking a Drive, Hitching a Ride: Autonomous Driving and Car Usage", In: Autonomous Driving, DOI 10.1007/978-3-662-48847-8_31 (2016)
16. Woisetschläger D.M., "Consumer Perceptions of Automated Driving Technologies: An Examination of Use Cases and Branding Strategies", In: Autonomous Driving, DOI 10.1007/978-3-662-48847-8_32 (2016)

17. Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial expression analysis. Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG'00), Grenoble, France, 46-53.
18. Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression. Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010), San Francisco, USA, 94-101.
19. "Challenges in Representation Learning: A report on three machine learning contests." I. Goodfellow, D. Erhan, P.L. Carrier, A. Courville, M. Mirza, B. Hamner, W. Cukierski, Y. Tang, D.H. Lee, Y. Zhou, C. Ramaiah, F. Feng, R. Li, X. Wang, D. Athanasakis, J. Shawe-Taylor, M. Milakov, J. Park, R. Ionescu, M. Popescu, C. Grozea, J. Bergstra, J. Xie, L. Romaszko, B. Xu, Z. Chuang, and Y. Bengio. arXiv 2013.
20. Michael J. Lyons, Shigeru Akemastu, Miyuki Kamachi, Jiro Gyoba. Coding Facial Expressions with Gabor Wavelets, 3rd IEEE International Conference on Automatic Face and Gesture Recognition, pp. 200-205 (1998).
21. N. Aifanti, C. Papachristou and A. Delopoulos, ”The MUG Facial Expression Database” in Proc. 11th Int. Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), Desenzano, Italy, April 12-14 2010.
22. Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D.H.J., Hawk, S.T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. Cognition & Emotion, 24(8), 1377—1388. DOI: 10.1080/02699930903485076
23. Chollet, F. and others, Keras (2015), https://keras.io, last accessed 2019/05/10
24. Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.
25. Bradski, G. (2000) The OpenCV Library. Dr. Dobb’s Journal of Software Tools, 120; 122-125.
26. Facial Expression Database Classifier (FEDC), https://github.com/AntonioMarceddu/Facial_Expression_Database_Classifier, last accessed 2019/05/10