European Transport \ Trasporti Europei (2021) Issue 84, Paper n° 5, ISSN 1825-3997

https://doi.org/10.48295/ET.2021.84.5

Designing in-car emotion-aware automation

Silvia Ceccacci[1], Maura Mengoni[1], Andrea Generosi [11*], Luca Giraldi[2], Roberta Presta[3], Giuseppe Carbonara[3], Andrea Castellano[3], Roberto Montanari[3]

[1]Department of Industrial Engineering and Mathematical Sciences, Università Politecnica delle Marche - Ancona, Italy
[2]Emoj Srl – Ancona, Italy
[3]RE-Lab Srl – Reggio Emilia, Italy

Abstract
Driver behaviour recognition is of paramount importance for in-car automation assistance. It is widely recognized that not only attentional states, but also emotional ones have an impact on the safety of the driving behaviour.
This research work proposes an emotion-aware in-car architecture where it is possible to adapt driver's emotions to the vehicle dynamics, investigating the correlations between negative emotional states and driving performances, and suggesting a system to regulate the driver's engagement through a unique user experience (e.g. using music, LED lighting) in the car cabin.
The relationship between altered emotional states induced through auditory stimuli and vehicle dynamics is investigated in a driving simulator. The results confirm the need for both types of information to improve the robustness of the driver state recognition function and open up the possibility that auditory stimuli can modify driving performance somehow.

Keywords: Emotion recognition, facial expression recognition, driver monitoring system.

1. Introduction

Equipping vehicles with intelligent driver assistant systems seems to be a promising way of preventing road traffic accidents since most of them are due to driver's performances (Panou, 2018; Stephens and Groeger, 2009). Therefore, driver Monitoring and Assistance Systems (DMAS), acting on the user's status, are increasingly used (Saulino et al., 2015). Driver monitoring is the key function allowing for the adaptation of the assistance provided by the automation to the driver. To date, psychophysiological monitoring of human states has the chance of exploiting a plethora of sensors, and the detection results can be combined with the analysis of the driving context to provide by means of adaptive HMI (Human Machine Interfaces) the best driving experience (Khan & Lee, 2019). Partially autonomous driving still needs the human "in the loop" in some circumstances, and control switch from the automation to the human must ensure his / her suitability of being able to drive, thus calling even more for the critical role of human

[1*] Corresponding author: Andrea Generosi (a.generosi@univpm.it)

behavior monitoring and assessment (Davoli et al., 2020). Besides distraction, the detection of driver's emotions got momentum, given the potentially dangerous effects emotions can have on the driver's performance (Zepf et al., 2020; Jeon, 2017). Research demonstrates that attention and emotions are linked with driving performances (Pêcher et al., 2009), and aggressive driving, related to the difficulty in managing human emotions, is one of the primary cause of cars accidents (Özkan et al., 2011; Sârbescu, 2012). Negative emotions, like anxiety and anger, can affect perception and decision-making and, sometimes, even alter physical capabilities (Lisetti and Nasoz, 2005; Matthews, 2002). According to (Jeon et al., 2017), some emotions have different effects on others: for example, anger and happiness significantly reduce the driving performances and safety level compared to neutral and especially fear. Equipping cars with an emotion-aware system would provide warning and proper stimuli to regulate emotions, ranging from ambient light modulation to empathic vocal interactions with the assistance systems or acting on the vehicle's dynamics (Braun et al., 2019). AI technologies allow human emotions to be detected and monitored automatically. Today's different technologies are used to recognize people emotions differ mainly at an intrusiveness level: especially biofeedback sensor, like EEG (Electroencephalography), can introduce a strong bias that could affect the subjects' behavior and the experienced emotion itself (Ceccacci et al., 2018). For this reason, this research area is starting to focus on non-intrusive devices in the last year to automatically recognize human emotions, particularly for speech and facial coding analysis. Most of the facial expression recognition systems today make use of Deep Neural Networks (especially Convolutional Neural Networks), like the one presented by (Generosi et al., 2018; Generosi et al., 2019), taking pictures of human faces in input and providing a prediction of the relative Ekman's primary emotions (i.e., happiness, surprise, sadness, anger, disgust, and fear) (Ekman and Friesen, 1978), just like most of the state-of-the-art systems that involve this kind of technologies (Li et al., 2018). The literature proposes different emotion-aware car systems, some using voice analysis (Jones and Jonsson, 2007) others wearable devices (Nasoz et al., 2010; Katsis et al., 2008). Many of the proposed Driver Monitoring Systems (DMS) aim to recognize:

1) Inappropriate driving behaviour and abnormal driver’s state through car’s dynamic data
2) Driver’s looking direction and so every kind of visual distraction and drowsiness, through cameras.

This basics equipment already allows for integrating emotion awareness in the interactive dialogue between the driver and the automation and constitutes the infrastructure considered for the architecture proposed and discussed in the following chapters.

2. Research aim

Extending the research work proposed in (Ceccacci et al., 2020), with a more structured tool design proposal and experimental phase, this project aims to provide a model and implement an in-car emotion-aware architecture through a reactive and “symbiotic” human-computer interface able to:

1) Analyze driver's facial expression and his/her emotional state;
2) Increase the driver's comfort and safety by creating a link between emotions and the car interface

An overall architecture with preliminary implementation and assessment results are presented. In particular, a comparative study, within a driving simulation environment, has been carried out in order to:

1) Observe the relationship between driving performance parameters and detected emotions, induced through acoustic stimuli;
2) Evaluate the impact of the driving task on emotions elicited through acoustic stimulation.

3. The proposed system

The architecture presented in Figure 1 is characterized by:
1) A Driving Monitoring System, leveraging (i) a Driving Style Detection module and (ii) an Emotion Recognition module;
2) A smart car interface, controlling led lights and music within the car environment.

The functionalities of such modules and the implementation details of some of them are described in the following sub-paragraphs.

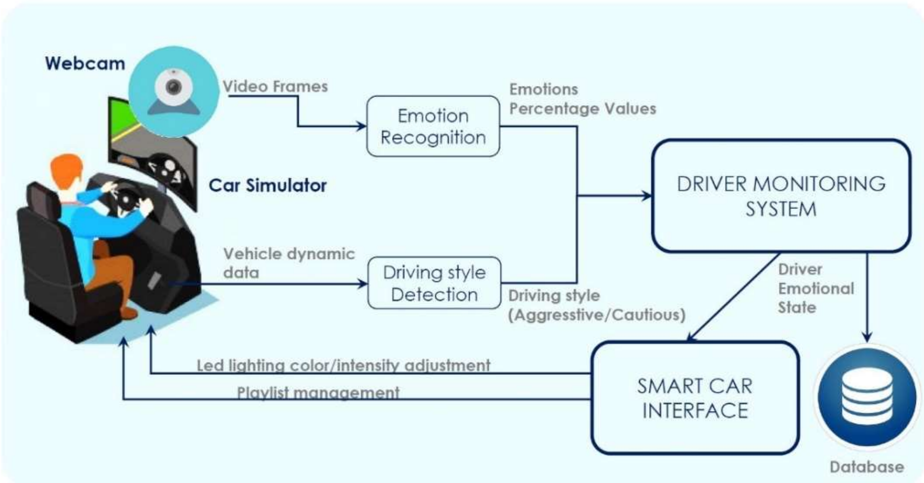

Figure 1: The proposed system architecture.

3.1 Driving Style Detection module

Different research works (Toledo and Lotan, 2006; ROSPA, 2013; Verster and Roth, 2011) demonstrated that driving data, often acquired from the vehicle CAN (Controller Area Network), are significant and objective indicators to assess any driver's impairment: these include steering frequency (expressed in Steering Wheel Reversal Rate, SWRR) (Macdonald and Hoffmann, 1980) and the driver steering respect to the center of the lane (Standard Deviation from Lateral Position, SDLP, i.e. the results of the vehicle's movement induced by driver steering actions with respect to the road environment) (Verster and Roth, 2011). Another crucial data for this approach is time required for two vehicles to collide at a certain speed, from the start of braking, on the same path (Time to Collision, TTC) (Van Der Horst and Hogema, 1993).

The Driving Style Detection module follows the trend dictated by today's DMS and DMAS, designed to use and depend on this kind of data acquired from the vehicle CAN, to assess the driving behavior (Saulino et al., 2015).

3.2 Emotion Recognition module

The Emotion Recognition module is based on a Convolutional Neural Network (CNN) trained using a merged dataset with both “in the wild" and “in lab” properties, and the Python version of Tensorflow and Keras frameworks (Talipu et al., 2019). In particular, this CNN has been trained with the public dataset CK+ (Lucey et al., 2010) and FER+ (Barsoum et al., 2016) built in laboratory, and the “in the wild " dataset provided by Affectnet. Combining and cleaning these datasets with different properties, has been possible to improve the resulting model reliability, obtaining a dataset composed of 250k photos. The merged dataset has been splitted using the 80-20 proportion, so using 80% of the dataset for the training phase and 20% for the validation phase. The deploy script has been developed in Python using Dlib, Tensorflow and Keras frameworks: giving the face images to the trained model’s network input layer, it returns in output the six main Ekman’s Emotions (happiness, surprise, anger, disgust, sadness, fear and neutral) classification probabilities. Different Keras model architectures such as Inception (Szegedy et al., 2016), VGG13, VGG16 and VGG19 (Simonyan and Zisserman, 2014) have been tested. Considering the test results, VGG13 has been chosen as the best one, with a 75.48% of accuracy.

3.3 Smart Car Interface

The Smart Car Interface is the module that manages the adjustment of dashboard lights and radio music playlists based on the detected driver's affective state and rules generated by the application of predictive Machine Learning models. The objective is to adjust the driver's emotional state as soon as specific conditions, considered dangerous for the driving style, are detected. This goal is achieved through the activation of lights and sounds/music judged suitable to bring the driver's detected emotional condition to a neutral state. A classic example that can be taken into consideration is when a driver is in an altered emotional state due to an incorrect driving style by other drivers, in this case the objective of the Driver Monitoring System becomes to detect as soon as possible such a condition and activate through the Smart Car Interface a different tone/color of the dashboard lights and a musical playlist suitable for the driver emotion regulation. As a first step, how to match five Ekman’s basic emotions (Joy, Surprise, Fear, Sadness and Anger) with lighting colors and most common musical genres, has been investigated. Following the main approach described in (Altieri et al., 2019) and (Altieri et al., 2019), to follow this purpose it has been necessary to define 5 color transitions (for every investigated Ekman’s emotion) using a survey carried out involving about 300 Italian people (58.4% females and 41.6% males), older than 18 years (27.3% aged between 18 and 24, 61.1% aged between 25 and 35, 11.6% older then 35) so to get an association between Ekman’s emotions and color. Results are shown in Table 1.

Table 1. Predominant emotions-color associations

EMOTION	COLOR	RESULT (%)
Joy	C3 (Yellow)	46,6
Surprise	C2 (Orange)	41,1
Fear	C7 (Purple)	42,7
Anger	C1 (Red)	71,5
Sadness	C6 (Blue)	62,2

About music tracks, it has been considered to map seven musical genres (Pop, Rock, Classical, Latin, R&B, Jazz, and Metal) with the bi-dimensional valence-arousal model proposed by Russel in (Russel et al., 1980). To associate this scale of values with the most popular music tracks the Spotify Web APIs has been used: these APIs provide metadata for songs belonging to a Spotify playlist, including genre, loudness, energy, bpm, valence etc. In this context, following also results discussed in (Kim et al., 2011) it has been possible to notice that songs characterized by "high-valence/high-arousal" values are strongly related with exciting sensations, while "low-valence/low-arousal" values are associated with sad, melancholic, and boring music. Accordingly with this approach, and as showed in (Altieri et al., 2019), five areas have been identified in this valence-arousal space, so to map Ekman's emotions with the Russel quantitative system.

As investigated in this paper and as will be further explored in future research works, there exists a statistical correlation between some of Ekman's emotions and an aggressive driving style. To bring the driver's altered emotional state back to acceptable threshold levels for a correct driving style, the most immediate association could be to modify the dashboard lights by proposing blue or purple (associated with emotions of sadness and fear) as the dominant colors, together with playlists composed of genre songs and valence/arousal values associated with the same emotions. For example, it is possible to notice how the Jazz genre, or many tracks of the Classic genre are associated with the emotions of Sadness and Fear, while songs belonging to the Rock or Metal genre, are often associated with emotions such as Surprise or Anger. Although the system described in paragraph 4.2 can apply the proposed solutions, it is not the purpose of this paper to investigate at an experimental level whether these types of feedback actually have the effect of changing the emotional state of the user driving and how they may affect his/her driving style, but it is instead to show the rationale that is behind its functioning.

4. Experimental case study

The availability of driving performance data via the car CAN network and the possibility of leveraging low-cost video-based analysis systems for the emotion recognition encourages the development of driver state detection techniques exploiting such kind of information.The study presented in this section of the paper investigates within a driving simulation environment the relationship between driving performance parameters and detected emotions, that are the objectives of the two fundamental components of the proposed driving monitoring system, i.e., the driving style detection module and the emotion recognition module.

Experimental design and procedure, data collection and the statistical methods used in the analysis are described in detail below.

4.2 Experimental design.

The experiment setting involved a driving simulator equipped with a camera capturing the frontal video of the driver.

To the aim of inducing emotions in participants, sound stimuli were considered because of their proved efficacy and efficiency, i.e., their capacity to induce emotion despite their short duration time (Jeon, 2017). A set of seven easily recognizable, clearly distinguishable, and strongly connotated audio tracks of few seconds were used: car crash, child's laughter, vomiting, fart, zombie, car horn, and scream of pain.

Participants were randomly assigned to two groups: a control one, where no sound stimuli were provided during the driving task, and an experimental one, where the sound stimuli were randomly provided every 45 seconds. The control group allowed for the collection of a baseline of driving performance data, while the experimental group allowed for the collection driving performance data under elicited emotional conditions.

This first study is indeed conducted to understand the impact of altered emotional states on driving performance (dependent variable), i.e., how different are driving parameters between neutral emotional states (condition: no sounds provided while driving) and non-neutral emotional states (condition: sounds provided while driving).

Revealing significant differences in the dependent variable implies that sound-elicited emotions can change driving parameters and that, in turn, driving parameters are important sources of information to detect altered emotional states in the driver.

A further analysis was conducted within the same study to understand if the driving task (independent variable) altered the emotional response (dependent variable). Participants of the control group, indeed, after the driving task, were asked to stay sit down in front of the shut-down simulator screen (condition: without driving) while listening to the randomized sequence of the 45''- separated sounds.

This second analysis is conducted to understand if the emotions detected under the two different conditions (i.e., while driving, for the experimental group, and without driving, for the control group) are different.

If no differences are revealed between the two groups, we can conclude that:
1) The driving task does not impact on the emotional response to the considered stimuli;
2) We can consider the emotional responses of subjects from both groups to label the non-neutral emotional states under which the driving simulator data were collected.

Figure 2. The driving simulator (on the left) and the circuit shape used for the driving scenario (on the right).

The driving simulator used for the experiment (Figure 2) was built on Oktal SCANeR II platform, enriched with real car commands (e.g., gearbox, pedals, wipers and indicators). The simulation software engine was SCANeR Studio 1.7, including ADAS (Advanced Driver Assistance Systems) and Autonomous Driving functionalities and the driving scenario was visualized through a projector. The simulation engine ran on a Windows 10 pc, while the drivers of the commands ran on a Windows 7 pc connected with an acquisition board able to digitalize the signal coming from the commands. Two audio sources were located on both sides of the driver's seat, in order to ensure an immersive and realistic audio experience. The steering wheel was a SensoDrive force-feedback steering wheel, connected through Peak-CAN to the simulator. The driving scenario was based on a two-lane highway, with no traffic, road signs with speed limit, and a countryside landscape as background. The shape of the circuit was similar to two circles connected by a straight line for a total length of 12 km (Figure 2). For the emotion recognition software, a Logitech Brio 4K webcam has been used.

4.3 Participants

A total of 20 voluntary subjects (9 females and 11 males) have been involved and randomly splatted into the two groups. The control group counts a total of 10 subjects (5 females and 5 males) aged between 24 and 30 (Mean = 26.6, SD = 1.34). The experimental group consists of 10 subjects (4 females and 6 males) aged between 25 and 34 (Mean = 29.4, SD = 2.83). All the participants had a valid driving license since at least 3 years and no particular hearing problems.

4.4 Experimental Procedure and Data Collection

Before starting the test, all participants were instructed about the objectives of the experiment, were asked to sign the informed consent, and to complete a 5-minute training in order to become familiar with the driving simulator.

After that, all the participants (both of the control and experimental groups), one per time, had to execute a driving task lasting 6 minutes. They were required to respect the code of the road as they would in a natural setting. Their speed information was shown on the dashboard speedometer.

For all the participants, several driving parameters were monitored and used as dependent variables for the driving performance evaluation. In particular, the simulator was able to record the Standard Deviation of the Lane Position (SDLP) and the Standard Deviation of the Steering Wheel Angle (SDSTW). These parameters have been considered as metrics to evaluate later control performance. On the other hand, the indicators of longitudinal control performance, Standard Deviation of Pressure (SDP) for the gas pedal and Standard Deviation of Speed (SDS), have instead been considered.

During the driving task, only the participants of the experimental group were asked to listen to the seven acoustic stimuli (i.e., car crash, child's laughter, vomiting, fart, zombie, car horn, and scream of pain) of approximately 5 seconds each and delivered with a delay of 45 seconds between each other. The order the stimuli have been administered has been counterbalanced across subjects. For the participants of the experimental group, video captured from the camera was processed through the Emotion Recognition system and resulted information data were also collected and synchronized with driving performance data. In particular, all the main Ekman's emotions, in terms of percentage probability that a photo belongs to a particular emotional category, have been recorded. After the driving task, only participants of the control group were asked also to complete a listening task equal to that performed by participants in the experimental group during the driving task (i.e., 7 acoustic stimuli, 5 seconds each and delivered every 45 seconds, order of stimuli counterbalanced across subjects). Participants of the control group listened to the audio stimuli sitting on the simulator chair without driving, so focusing on the stimulus itself. In this case, only emotional information has been recorded. The test for the control group lasted about 20 minutes, while the experimental group test lasted 15 minutes.

5. Results

5.1 Emotion aroused by acoustic stimuli.

The emotions detected in the six seconds following the start of each acoustic stimulus were compared between groups (Figure 3) using the non-parametric Mann Whitney U test due to the categorial of the variables and the lack of normality of the distributions.

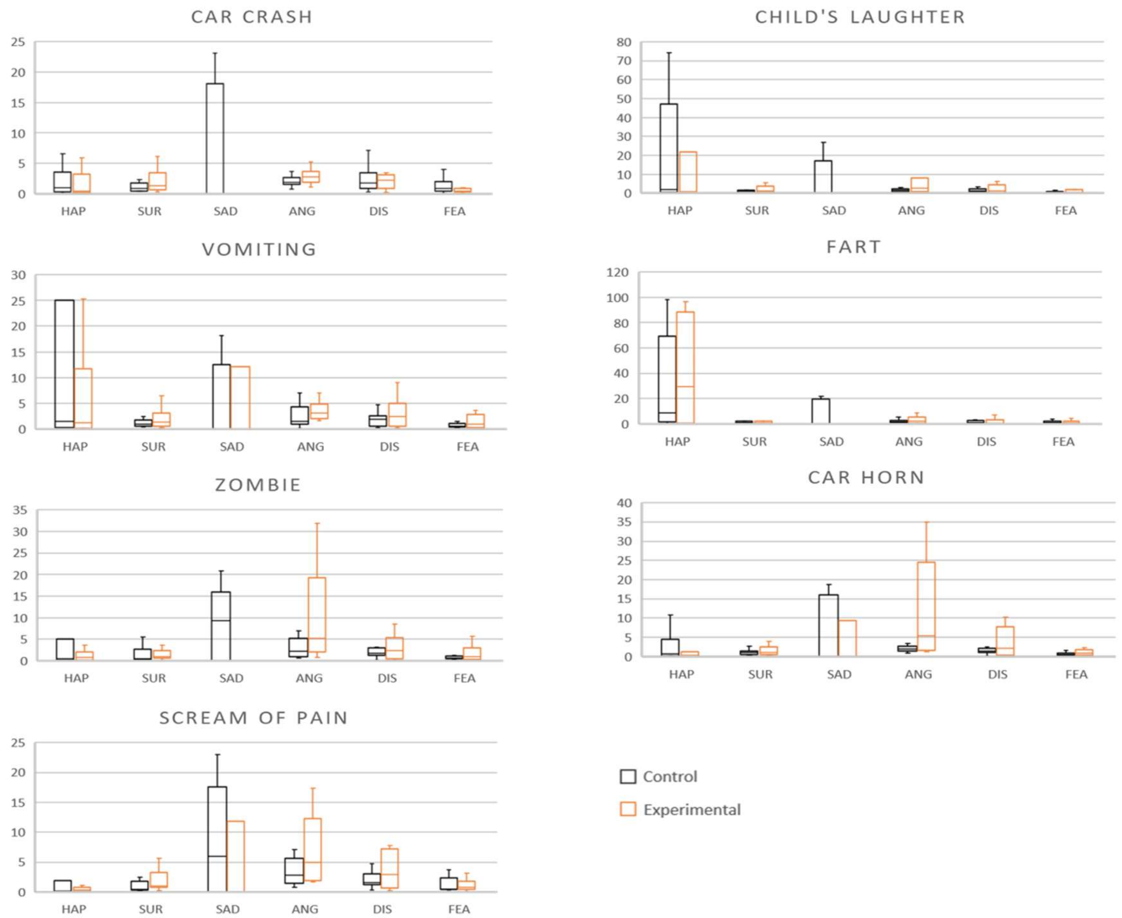

Figure 3: Emotions detected in the six seconds following the start of each acoustic stimulus.

Results revealed only a statistically significant difference ($U = 24$, $p < .005$) between the level of anger elicited by stimulus 1 (car crash) while driving (*Mdn* = 7,99) compared to non-driving (*Mdn* = 2,44). There are no significant differences between the emotions elicited while driving and non-driving for all the other stimuli. This suggests that, in general, the driving task does not impact on the emotional response to the considered stimuli, so that the emotions induced by the sounds do not differ in the driving context compared to the non-driving context. Therefore, we used the responses of all participants (both control group and experiment group) to "label" the sounds used.

To determine which prevailing emotions (if any) were elicited, a within-group assessment was conducted for each stimulus using Wilcoxon signed-rank test.

Results revealed that the **stimulus 1 (car crash)** most aroused both anger and disgust. In fact, the detected level of anger (*Mdn* = 2,75) is statistically significantly higher than the level of fear (*Mdn* = 0,41), $Z = 2.805$, $p = .005$.

Also, the level of disgust (*Mdn* = 2,25) is statistically significantly higher than the level of fear, $Z = 2.701$, $p = .007$. While there is no statistically significant difference between the levels of anger and disgust and between the levels of fear, happiness (*Mdn* = 0,4575) and surprise (*Mdn* = 1,38). The **stimulus 2 (child's laughter)** most aroused anger. In fact, the detected level of anger (*Mdn* = 2,63) is statistically significantly higher than the level of fear (*Mdn* = 0,52), $Z = 2.599$, $p = .009$. There are no statistically significant differences between the levels of fear, happiness (*Mdn* = 0,865), sadness (*Mdn* = 0) and surprise (*Mdn* = 1,03) and disgust (*Mdn* = 1,065). The **stimulus 3 (vomiting)** most aroused both anger and disgust. The level of fear (*Mdn* = 0,98) is statistically lower than the level of anger (*Mdn* = 3,05), $Z = 2.497$, $p = .013$, and disgust (*Mdn* = 2,445), $Z = 2.090$, $p = .037$. Whereas there are no statistically significant differences between the levels of fear, happiness (*Mdn* = 1,35), sadness (*Mdn* = 0) and surprise (*Mdn* = 1,24). There are no statistically significant differences between anger and disgust. The emotion most aroused by the **stimulus 4 (fart)** is happiness. In fact, the detected level of happiness (*Mdn* = 29,45) is statistically significantly higher than the level of fear (*Mdn* = 0,46), $Z = 2.191$, $p = .028$. There are no statistically significant differences between the levels of fear, disgust (*Mdn* = 0,45), sadness (*Mdn* = 0), anger (*Mdn* = 1,90) and surprise (*Mdn* = 0,57).

Stimulus 5 (zombie) predominantly elicited anger, and secondarily disgust. In fact, the detected level of anger (*Mdn* = 5,20) is statistically significantly higher than the level of fear (*Mdn* = 0,92), $Z = 2.803$, $p = .005$. Besides, the level of disgust (*Mdn* = 2,34) is statistically significantly higher than fear, $Z = 2.395$, $p = .017$, but it is statistically significantly lower than anger, $Z = 2.599$, $p = .009$. There are no statistically significant differences between the levels of fear, sadness (*Mdn* = 0), happiness (*Mdn* = 0,75) and surprise (*Mdn* = 0,86). Anger resulted in the emotion most aroused by the **stimulus 6 (car horn)**. The level of anger (*Mdn* = 5,32) is statistically significantly higher than the level of fear (*Mdn* = 0,90), $Z = 2.652$, $p = .008$. There are no statistically significant differences between the levels of fear, disgust (*Mdn* = 2,16), sadness (*Mdn* = 0), happiness (*Mdn* = 0,33) and surprise (*Mdn* = 0,98). **Stimulus 7 (scream of pain)** mainly aroused both anger and disgust. The level of fair (*Mdn* = 0,73) is statistically significantly lower than the levels of anger (*Mdn* = 4,94), $Z = 2.701$, $p = .007$, and disgust (*Mdn* = 2,93), $Z = 2.293$, $p = .022$. There are no statistically significant differences between the levels of anger and disgust. There are no statistically significant differences between fear, sadness (*Mdn* = 0), happiness (*Mdn* = 0,36) and surprise (*Mdn* = 0,98).

5.2 Driving performance

The analysis of the driving performance in the different conditions was conducted firstly at the level of the two groups with all the stimuli aggregate, to investigate the effect of general acoustic emotional distraction on the driving performance. Secondly was investigated the impacts of each single stimulus (each characterized for a specific predominant emotion) on driving. Thirdly, at the level of the single stimulus as within-group analysis only for the experimental group to assess if, on the same subject, different emotional sounds can generate a different driving performance.

Regarding the driving performance between the two driving test groups (with and without acoustic stimuli), only SDLP is statistically relevant (Figure 4), also if not normally distributed. However, comparing these two groups with an unpaired twosamples Wilcoxon test, it showed a statistically higher value (W = 81, p = 0.018) in the tests without acoustic stimuli (Median=0.56) concerning the other group (Median = 0.28).

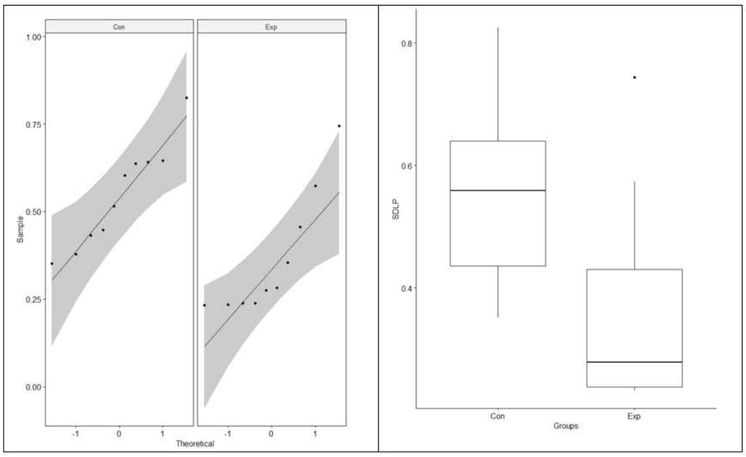

Figure 4: Comparison of the Standard Deviation of Lane Position (SDLP) between groups

At the level of each considered sound, different stimuli resulted to be statistically different in the effects on the three considered driving performance indicators. Indeed, **the stimulus 1 (sound of car crash)** generated a statistically different lateral driving performance t(18) = 2.746, p = 0.01328, for what concern the SDSTW (Figure 5), which also showed a normal distribution and resulted to be lower in the control group (mean = 0.46, SD = 0.21) than in the experimental group (mean = 0.72, SD = 0.20); hence, highlighting a lower driving performance in the control group respect to the experimental group. Similarly, also the **stimulus number 4 (Fart/raspberry)** had caused an impairing driving performance in the experimental group for what concerns the SDS (Figure 6), which resulted to be not a normal distribution, but significantly higher (W = 79, p = 0.02881) for the experimental group (median = 2.48), respect to the control group (median = 0.57).

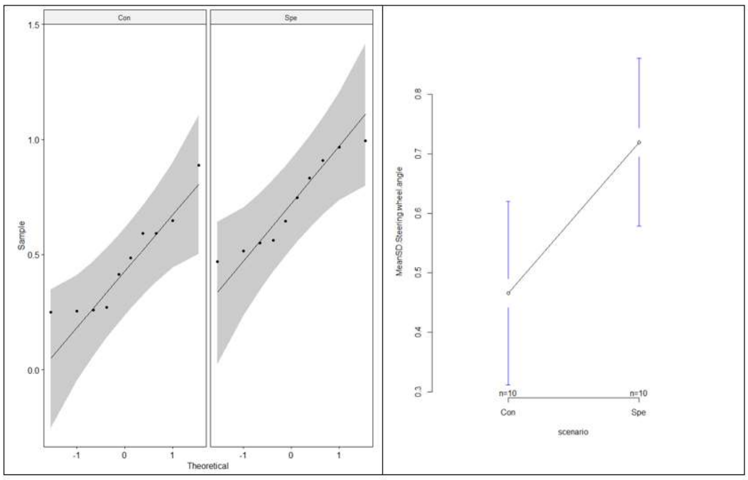

Figure 5: Comparison of the Standard Deviation of Steering Wheel angle (SDSTW) for stimulus 1

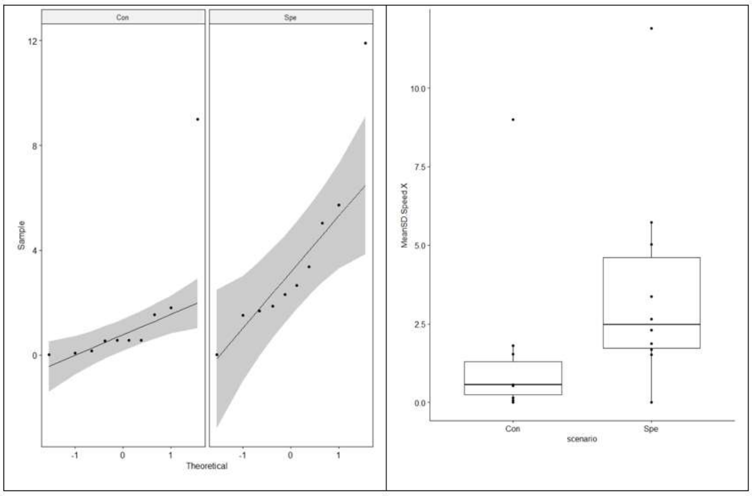

Figure 6: Comparison of the standard deviation of Standard Deviation of Speed (SDS) for stimulus 4

The **stimulus number 5 (Noise of zombie)**, from its part was associated with two contradictory outcomes (Figure 7). For what concern the SDLP (as lateral performance index) it resulted to be significantly higher (W = 23, p = 0.0432) for the control group (median = 0.21) (hence a lower driving performance) respect to the experimental group (median = 0.1). Contrarily, for what concern the standard SDS (as longitudinal performance index), the experimental group evicted a significantly (W = 78, p = 0.04) lower performance (median = 1.19) respect to the control group (median = 0.10).

The last stimulus that was associated with a statistically significant difference (t(18) = -- 3.5192, p = 0.00245) in terms of driving performance was the **number 7 (scream of pain)** for what concern the SDLP (Figure 8); this resulted to be higher in the control group (mean = 0.28, SD = 0.071) respect to the experimental group (mean = 0.15, SD = 0.03). Showing a better lateral performance for the experimental group respect to the control group.

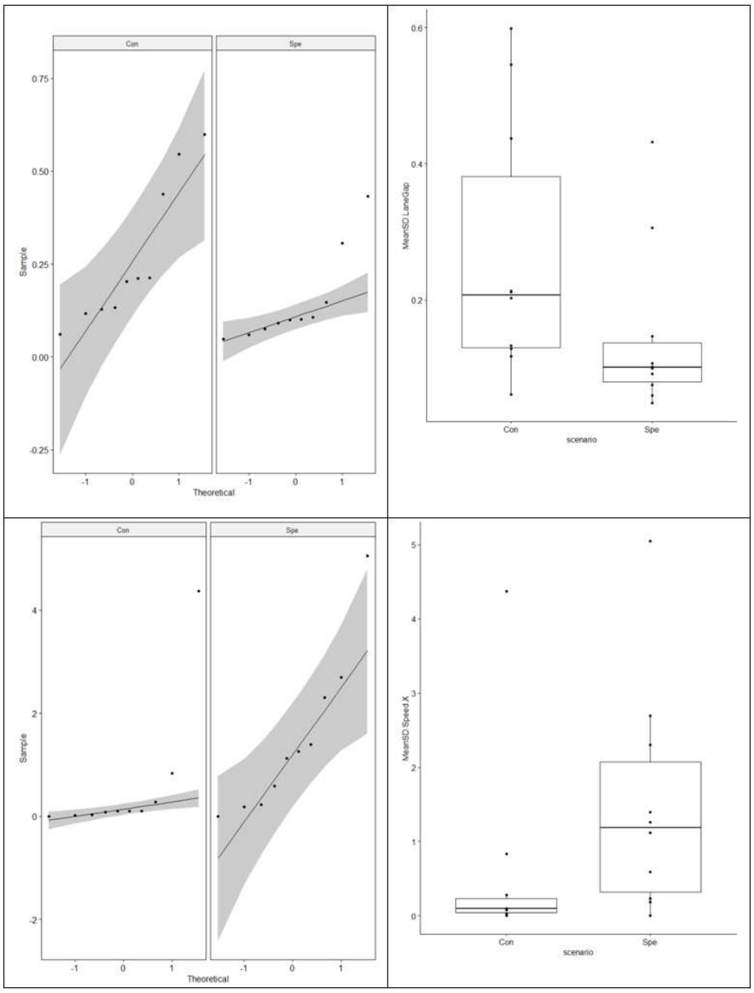

Figure 7: Comparison of the SDLP (on the top) and SDS (on the bottom) for stimulus 5

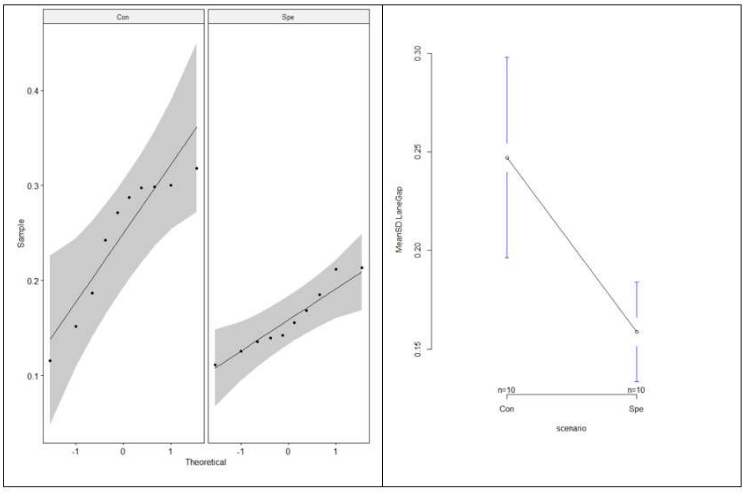

Figure 8: Comparison of the standard deviation of lane gap for stimulus 7

For what concern the within group analysis, the Friedman test showed a statistically significant difference for the indexes related to the SDLP (X2(6) = 13.3, p = 0.039) and the SD of the speed (X2(6) = 21.8, p = 0.00134). However, the effect size of both the indexes was small (respectively 0.111 and 0.181) and the pairwise Wilcoxon signed rank test between groups did not reveal statistically significant differences.

6 Conclusion

This research work proposed an emotion-aware Driving Monitoring System supported by a low-cost Deep Learning-based emotion recognition tool and driving performance parameters. A preliminary experiment has been carried out to investigate the relation between altered emotions and driving parameters, showing the need for both the sources of information to better monitor and understand the driver's state.

The experiment herein presented raises the possibility of leveraging sounds to elicit/regulate emotions and, contextually, altering driving performances. Indeed, the exposure to the auditory stimuli has somehow modified driving performance in terms of SDLP. It seems that the driving task with emotional stimulation gave better lateral control of the trajectory. This behavior could be related to the characteristics of the driving route, really linear and repetitive, and consequently boring. According to (Jeon et al., 2017), boredom is often related to the low activation of emotional states during driving activities. Considering this, acoustic stimuli could have improved the enjoyment of the experience, so the positive engagement, during the driving activities. However, this result deserves to be better investigated through future studies aimed at better understanding the effects of acoustic stimulation on driving performance. Furthermore, given that the research in the sector is still in its early stages, the results must obviously still be taken with caution for an application on the road. In particular, future studies should be conducted to better evaluate how certain acoustic stimuli can change the driver's emotional state and affect driving performance, so that a knowledge base can be built to automatically manage emotional state induction/regulation functions to increase driver comfort and improve driving performance. In the same way, it will be necessary to investigate how in a real car context the activation of lights proposed for the Smart Car Interface can cause distraction and inevitably affect driving performance negatively.

References

Altieri A., Ceccacci S., Mengoni M. (2019) "Emotion-Aware Ambient Intelligence: Changing Smart Environment Interaction Paradigms Through Affective Computing". *In Streitz N., Konomi S. (eds) Distributed, Ambient and Pervasive Interactions. HCII 2019.* Lecture Notes in Computer Science, vol 11587. Springer, Cham.

Altieri, A., Ceccacci, S., Ciabattoni, L., Generosi, A., Talipu, A., Turri, G., & Mengoni, M. (2019, January). An Adaptive System to Manage Playlists and Lighting Scenarios Based on the User's Emotions. In 2019 IEEE International Conference on Consumer Electronics (ICCE) (pp. 1-2). IEEE.

Barsoum, E., Zhang, C., Ferrer, C. C., and Zhang, Z. (2016) "Training Deep Networks for Facial Expression Recognition with Crowd-Sourced Label Distribution" 2016.

Braun, M., Schubert, J., Pfleging, B., & Alt, F. (2019). Improving driver emotions with affective strategies. *Multimodal Technologies and Interaction*, 3(1), 21.

Ceccacci, S., Generosi, A., Giraldi, L., & Mengoni, M. (2018, June). An emotion recognition system for monitoring shopping experience. In Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference (pp. 102-103).

Ceccacci, S., Mengoni, M., Andrea, G., Giraldi, L., Carbonara, G., Castellano, A., & Montanari, R. (2020, July). A Preliminary Investigation Towards the Application of Facial Expression Analysis to Enable an Emotion-Aware Car Interface. In International Conference on Human-Computer Interaction (pp. 504-517). Springer, Cham.

Davoli, L.; Martalò, M.; Cilfone, A.; Belli, L.; Ferrari, G.; Presta, R.; Montanari, R.; Mengoni, M.; Giraldi, L.; Amparore, E.G.; Botta, M.; Drago, I.; Carbonara, G.; Castellano, A.; Plomp, J. On Driver Behavior Recognition for Increased Safety: A Roadmap. *Safety* 2020, 6, 55.

Ekman, P., and Friesen, W. V. (1978) "*Manual for the Facial Action Coding System*," Consulting Psychologists Press.

Generosi, A., Ceccacci, S., & Mengoni, M. (2018, September). A deep learning-based system to track and analyze customer behavior in retail store. *In 2018 IEEE 8th International Conference on Consumer Electronics-Berlin (ICCE-Berlin)* (pp. 1-6). IEEE.

Generosi, A., Altieri, A., Ceccacci, S., Foresi, G., Talipu, A., Turri, G., ... & Giraldi, L. (2019, January). MoBeTrack: A Toolkit to Analyze User Experience of Mobile Apps in the Wild. I*n 2019 IEEE International Conference on Consumer Electronics (ICCE)* (pp. 1-2). IEEE.

Jeon, M., (2017) "Emotions in Driving", in: Jeon, M. (eds) *Emotions and Affect in Human Factors and Human-Computer Interaction*, Academic Press, 437-474.

Jones, C. M., & Jonsson, M. (2007, July). Performance analysis of acoustic emotion recognition for in-car conversational interfaces. *In International Conference on Universal Access in Human-Computer Interaction* (pp. 411-420). Springer, Berlin, Heidelberg.

Katsis, C. D., Katertsidis, N., Ganiatsas, G., & Fotiadis, D. I. (2008). Toward emotion recognition in car-racing drivers: A biosignal processing approach. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 38(3), 502-512.

Khan, M. Q., & Lee, S. (2019). A comprehensive survey of driving monitoring and assistance systems. *Sensors*, 19(11), 2574.

Kim, J., Lee, S., Kim, S., & Yoo, W. Y., (2011), "Music mood classification model based on arousal-valence values," *In Advanced Communication Technology (ICACT)*, 13th International Conference on, 292-295.

Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., Matthews, I., (2010)"The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression," *in 2010 ieee computer society conference on computer vision and pattern recognition-workshops*.

Lisetti, C. L., & Nasoz, F. (2005, July). "Affective intelligent car interfaces with emotion recognition." *In Proceedings of 11th International Conference on Human Computer Interaction*, Las Vegas, NV, USA.

Macdonald, W. A., and Hoffmann, E. R. (1980). "Review of relationships between steering wheel reversal rate and driving task demand". *Human Factors*, 22(6), 733-739.

Nasoz, F., Lisetti, C. L., & Vasilakos, A. V. (2010). Affectively intelligent and adaptive car interfaces. *Information Sciences*, 180(20), 3817-3836.

Özkan, T., Lajunen, T., Parker, D., Sümer, N., & Summala, H. (2011). "Aggressive driving among british, dutch, finnish and turkish drivers." *International journal of crashworthiness*, 16(3), 233-238.

Pêcher, C., Lemercier, C., & Cellier, J. M. (2009). "Emotions drive attention: Effects on driver's behaviour." *Safety Science*, 47(9), 1254-1259.

Russell, J. A., 1980, "A circumflex model of affect, Journal of Persoanlity and Social Psy-chology, 37(3), 1161.

Sârbescu, P. (2012). "Aggressive driving in Romania: Psychometric properties of the driving anger expression inventory." *Transportation research part F: traffic psychology and behaviour*, 15(5), 556-564.

Saulino, G., Persaud, B., & Bassani, M. (2015). Calibration and application of crash prediction models for safety assessment of roundabouts based on simulated conflicts. *In Proceedings of the 94th Transportation Research Board (TRB) Annual Meeting*, Washington, DC, USA (pp. 11-15).

Simonyan, K., Zisserman, A., "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., (2016) "Rethinking the inception architecture for computer vision," *in Proceedings of the IEEE conference on computer vision and pattern recognition*.

Talipu, A. Generosi, M. Mengoni, L Giraldi, (2019) "Evaluation of Deep Convolutional Neural Network architectures for Emotion Recognition in the Wild," in IEEE 23rd International Symposium on Consumer Technologies, 25-27, IEEE

Toledo, T., and Lotan, T. (2006). In-vehicle data recorder for evaluation of driving behavior and safety. Transportation Research Record, 1953(1), 112-119.

Van Der Horst, R., and Hogema, J. (1993, October). "Time-to-collision and collision avoidance systems". *In Proceedings of the 6th ICTCT workshop: Safety evaluation of traffic systems: Traffic conflicts and other measures* (pp. 109-121).

Verster, J. C., and Roth, T. (2011). "Standard operation procedures for conducting the on-the-road driving test, and measurement of the standard deviation of lateral position (SDLP)". *International journal of general medicine*, 4, 359.

Zepf, S., Hernandez, J., Schmitt, A., Minker, W., & Picard, R. W. (2020). Driver Emotion Recognition for Intelligent Vehicles: A Survey. *ACM Computing Surveys* (CSUR), 53(3), 1-30.