AI-as-a-Service Toolkit for Human-Centered Intelligence in Autonomous Driving

Valerio De Caro[1§], Saira Bano[2,3§], Achilles Machumilane[2,3§], Alberto Gotta[3], Pietro Cassarà[3],
Antonio Carta[1], Rudy Semola[1], Christos Sardianos[4], Christos Chronis[4], Iraklis Varlamis[4],
Konstantinos Tserpes[4], Vincenzo Lomonaco[1], Claudio Gallicchio[1] and Davide Bacciu[1]
[1]Department of Computer Science, University of Pisa, Pisa, Italy
[2]Department of Information Engineering, University of Pisa, Pisa, Italy
[3]Information Science and Technology Institute "A. Faedo", National Research Council, Pisa, Italy
[4]Harokopio University of Athens, Athens, Greece

Abstract—This paper presents a proof-of-concept implementation of the AI-as-a-Service toolkit developed within the H2020 TEACHING project and designed to implement an autonomous driving personalization system according to the output of an automatic driver's stress recognition algorithm, both of them realizing a Cyber-Physical System of Systems. In addition, we implemented a data-gathering subsystem to collect data from different sensors, i.e., wearables and cameras, to automatize stress recognition. The system was attached for testing to a driving emulation software, CARLA, which allows testing the approach's feasibility with minimum cost and without putting at risk drivers and passengers. At the core of the relative subsystems, different learning algorithms were implemented using Deep Neural Networks, Recurrent Neural Networks, and Reinforcement Learning.

Index Terms—AI-as-a-Service, Human State Monitoring, Autonomous Driving, Recurrent Neural Networks

I. INTRODUCTION & MOTIVATION

The TEACHING project[1] aims at designing a computing platform and the associated software toolkit to support the development and the deployment of adaptive and dependable Cyber-Physical System of Systems (CPSoS) applications, enabling them to exploit sustainable human feedback to drive, optimize and personalize the provisioning of the offered services. The concept of Humanistic Intelligence in autonomous CPSoS is central to the project, where the human and the cybernetic components cooperate in the process of mutual empowerment. The concept of distributed, efficient, and dependable Artificial Intelligence (AI), leveraging edge computing support, is hence fundamental to achieving the goals of the project. One of the essential goals of the project is to design methodologies and tools to create the TEACHING AI-as-a-Service (AIaaS) software infrastructure for CPSoS applications, developing human-centric personalization mechanisms [1]. To this aim, the activities were organized in order to *i)* develop a toolkit that implements the core methodological components of the TEACHING AIaaS; *ii)* develop AI methodologies that are specialized in the recognition and characterization of the human physiological, emotional, and cognitive state from streams of data gathered from heterogeneous sensors; *iii)* develop the necessary functionalities to self-adapt and personalize AI models of autonomous driving, using
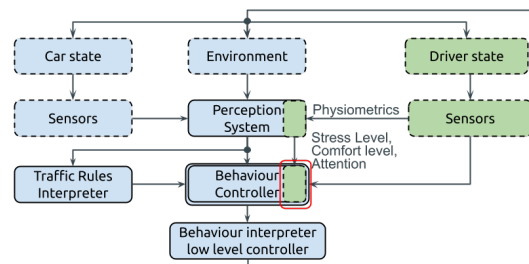
Figure 1. Overview of the presented control system.

the human state information implicitly; *iiii)* develop privacy-aware AI methodologies that can be bundled within the AIaaS toolkit.

In this work, we present a demonstrator which showcases the first three points in an autonomous driving simulation. In particular, the user experiences an autonomous driving simulation, while being monitored by a wearable sensor and a smart camera. The data collected by these devices are gathered into the AIaaS platform, which employs two main learning modules: one for detecting the cognitive state of the user from the physiological data, and the other for personalizing the driving style to one that better suits the state of the user. Data from both sensors and learning modules are then monitored and plotted by an additional application delegated to such purpose.

II. THE AIaaS TOOLKIT

The cognitive state of a human in a driving experience depends on the perception of the vehicle condition and the state of its environment, as well as on the internal beliefs, preferences, and expectations from an autonomous driving scenario. Several studies on perception and the consequent mental state [2], [3] fuse data from a large variety of sensors attached to the vehicle (e.g., accelerometers, gyroscopes, etc.) and the driver in an attempt to detect the driver's driving style and associate it to the driver/passengers' comfort or stress level [4].

Human-centered intelligence in autonomous driving can be achieved by adapting the driving behaviour according the corresponding contextual data. As sketched in Figure 1, such data can be categorized in three main types, depending on their availability along the driving experience:

1) *Vehicle's state*, which is monitored with the use of sensors attached to the vehicle, the driver and the passengers;
2) *Environment's state*, which denotes the outward context that can be used to achieve awareness of the situation occurring at any moment;
3) *Driver's state*, which can be classified by acquiring physiological signals through inward monitoring cameras or physiological sensors embedded in smart wearable devices.

In this context, the adaptation of the driving behaviour consists in performing actions to modify the vehicle's behaviour in order to provide a driving profile which better suites the cognitive state of the driver.

It is clear that the complexity of the task can scale at will, depending on the different HW/SW setups, as well as the different modelling choices from the AI perspective, making the definition of the dedicated software arbitrarily complex. The AIaaS toolkit that we show in this paper aims to streamline the definition of dedicated software by relying on a set of composable, hardware-agnostic app-building blocks called Learning Modules (LM) and data sources, which allow to design reusable and portable AI applications. In particular, the LM building blocks can be separately ported to and optimized for different device HW/SW architectures, increasing their efficiency with respect to common metrics (performance, power consumption), allowing careful debugging and verification, as well as allowing to exploit specific features of the execution platform within the LM. Communications between the modules of the AIaaS toolkit happen in two main ways, depending on whether the interactions are internal (i.e., between modules deployed within the vehicle) or external (i.e., from a model deployed within the vehicle, towards an external one). For the former, we employed a data brokering approach, in which publishers publish their own data on the message broker, while subscribers can gather specific data from it. For the latter, we employed a distributed event streaming platform, which enables the toolkit to high-performance communications towards edge and cloud. Finally, the AIaaS toolkit enables to humanistic intelligence, since, with an appropriate interaction between the modules, it is possible to produce a loop in which the human is implicitly an integral part.

III. DEMO DESCRIPTION

A. Modelling of the Scenario

As mentioned previously, the autonomous driving scenario is subject to a wide variety of tasks, depending on either the HW/SW support and the modelling choices of the driver's state, the vehicle's state and the environment's state. In this work, we showcase a simulated scenario in which the task consists on choosing the best driving style, depending only on the driver's state. Notice that the perception of the vehicle and the environment from the human affects their cognitive state, thus closing the loop with the system. In particular, the driver's state is modelled by two main components:

- *Facial Action Coding System* (FACS), which is a system for classifying facial movements as they appear in various human expressions [5]. Such movements of individual facial muscles are coded by FACS attributing a combination of codes corresponding to certain micro-movements,
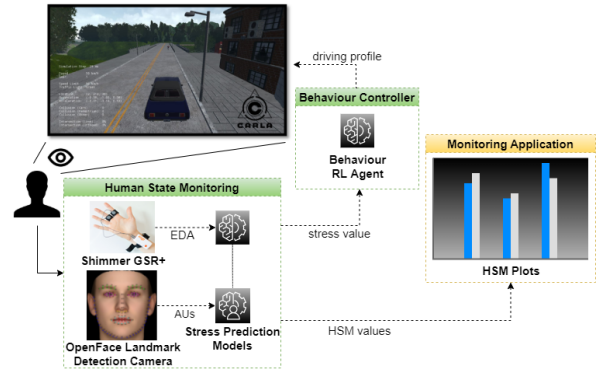

Figure 2. Overview of the demonstrator. Dashed arrows represent communications via MQTT.

called AUs, performed by the person. FACS can systematically encode almost all anatomically possible facial expressions, reducing them to a specific subset of AUs;
- *Cognitive state* of the driver, which we simplified with the notion of *cognitive stress*. Several papers in the literature employ physiological signals for identifying drivers' cognitive state [2] and comfort levels [4] under varying driving conditions.

Instead, the adaptation of the driving style consists in leveraging on predictions on the driver's state to predict the best driving profile for the vehicle's Advanced Driver-Assistance Systems (ADAS) from a predefined set. In the context, the task of personalizing the driving style is typically performed with the use of a model that can "learn from humans" how a real human would adjust the car's controls in various conditions. Recently, authors in [6] designed a deep Reinforcement Learning (RL) algorithm and used the deterministic policy gradients algorithm for training an asynchronous Advantage Actor-Critic (A3C) RL agent to maximize a reward that corresponds to the distance travelled by the vehicle without human interference. Nevertheless, most models are based on an expected average conscious user, and their control parameters are predefined and cannot be altered to match the different driver profiles. Thus, in our setting, we opted for a different approach to learn only with the use of a data coming from sensors and simulated scenarios.

B. Design and Components

In Figure 2, the demo platform is briefly presented: at the top of the figure, CARLA emulator [7] is the autonomous driving simulator, which provides the interaction with the driver. Data extracted by visual and physiological sensors provide the input to a ML model that yields the driver's stress level. According to this output, a RL algorithm provides a set of actions that personalize the driving profile simulated in CARLA aiming to resolve the driver's stress level. In the following, we will discuss the details of each of the interacting components of our demo.

1) Physiological Data: these data were collected with the use of a Shimmer3 GSR+ Unit, which monitors galvanic skin response between two reusable electrodes attached to the two fingers of one hand. A stimulus can activate the sweat glands, increasing moisture on the skin and allowing
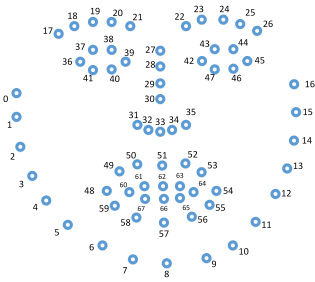
Figure 3. The 68 landmarks mapped onto the face of the subject.

the current to flow more readily by changing the balance of positive and negative ions in the secreted fluid (increasing skin conductance). The Shimmer3 GSR+ Unit can also provide a PPG signal from a finger, ear lobe, or other body part by adding an Optical Pulse probe. This signal is used as a proxy to compute the Heart Rate (HR) of the individual, which, in this demo, we used for a monitoring purpose only.

2) Visual Data: for collecting and processing these data, we employed OpenFace, which is an open-source software developed for computer vision applications based on facial behaviour analysis [8]. Such analysis is performed by detecting 68 *facial landmark locations* (Figure 3), which are employed for understanding emotional states; *head pose*, which is related to the expression of emotions and social cues; *gaze direction*, which is crucial when assessing attention, mental health, and emotion intensity. With the use of dedicated learning algorithms [8], Openface can determine the features mentioned above by analyzing faces in video files, in computer images or even through a live camera, as for this demo, without any specialized hardware.

3) Data-Gathering and Storage: sensors are accessed through a common, dedicated Sensors API, but not directly, in order to provide a place for adapter components that insulate the application support from the specifics of different kind of sensors. The data is then collected by the Data Brokering component, which takes the role of main distributor of data inside the AIaaS platform through a publish/subscribe messaging paradigm, allowing the selection of data of interest to each client. To this aim, we employed the MQTT protocol, and instantiated it with a Mosquitto[2] broker. Data from the physiological sensor were gathered thanks to its dedicated SDK, while an interface was developed to integrate OpenFace into the data-gathering subsystem.

4) Stress Detection: the goal of this module is to recognize the level of stress from physiological sensors. For this demo, we monitored the EDA to provide a continuous measure of the current level of stress through the inference of an Echo State Network (ESN), trained over a public dataset (WESAD). ESNs are models which belong to the family of recurrent neural networks, and are known for their efficiency in both training and inference, which makes them suitable for their use in embedded systems. This module showcases how trivial it is, with the AIaaS framework, to insert the human in the loop by adding dedicated modules (i.e., sensors and Learning Modules) to the application.

5) Driving Style Personalization: an RL method implements an Actor-Critic algorithm to train an agent that chooses driving profile which best suits the cognitive state of the driver [9]. More specifically, the value network predicts the driver's stress values with respect to the vehicle's behaviour, while a policy network outputs an action for CARLA's agent. The action consists in choosing between three driving profiles, i.e., conservative, normal and aggressive, which differ in terms of driving parameters like maximum speed, maximum steering angle and acceleration. These profiles are implemented with the use of the CARLA API, which allows to set the behaviour of the vehicle and the environment, as well as collecting the data from simulated sensors along the simulation.

IV. CONCLUSION

In this paper we present a proof-of-concept of the AI-as-a-Service toolkit, developed within the H2020 TEACHING project. We showcase an autonomous driving scenario in which a driving personalization system predicts a driving profile according to the driver's stress. In particular, a learning algorithm evaluates the driver's stress level with respect to the physiological data captured by a wearable sensor, while a RL agent adapts CARLA's driving profile to reduce the driver's stress. Furthermore, the driver is monitored by a smart camera, which extracts features about the facial behaviour. The demo, consisting in a driving session simulated through the CARLA software, highlights how easily the AIaaS toolkit allows to insert humanistic intelligence in AI applications.

REFERENCES

[1] D. Bacciu, S. Akarmazyan, E. Armengaud, M. Bacco, G. Bravos, C. Calandra, E. Carlini, A. Carta, P. Cassarà, M. Coppola *et al.*, "Teaching-trustworthy autonomous cyber-physical applications through human-centred intelligence," in *2021 IEEE International Conference on Omni-Layer Intelligent Systems (COINS)*. IEEE, 2021, pp. 1–6.

[2] K. A. Brookhuis and D. De Waard, "Monitoring drivers' mental workload in driving simulators using physiological measures," *Accident Analysis & Prevention*, vol. 42, no. 3, pp. 898–903, 2010.

[3] S. Begum, M. U. Ahmed, P. Funk, and R. Filla, "Mental state monitoring system for the professional drivers based on heart rate variability analysis and case-based reasoning," in *2012 Federated Conference on Computer Science and Information Systems (FedCSIS)*. IEEE, 2012, pp. 35–42.

[4] H. Bellem, T. Schönenberg, J. F. Krems, and M. Schrauf, "Objective metrics of comfort: developing a driving style for highly automated vehicles," *Transportation research part F: traffic psychology and behaviour*, vol. 41, pp. 45–54, 2016.

[5] P. Ekman and E. L. Rosenberg, *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*. Oxford University Press, USA, 1997.

[6] A. Kendall, J. Hawke, D. Janz, P. Mazur, D. Reda, J.-M. Allen, V.-D. Lam, A. Bewley, and A. Shah, "Learning to drive in a day," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 8248–8254.

[7] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "Carla: An open urban driving simulator," in *Conference on robot learning*. PMLR, 2017, pp. 1–16.

[8] T. Baltrusaitis, A. Zadeh, Y. C. Lim, and L.-P. Morency, "Openface 2.0: Facial behavior analysis toolkit," in *2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018)*. IEEE, 2018, pp. 59–66.

[9] C. Chronis, C. Sardianos, I. Varlamis, D. Michail, and K. Tserpes, "A driving profile recommender system for autonomous driving using sensor data and reinforcement learning," in *25th ACM Panhellenic Conference on Informatics (PCI 2021)*. ACM, 2021.