Audiovisual Affect Recognition for Autonomous Vehicles: Applications and Future Agendas

Vincent Karas, Dagmar M. Schuller, and Björn W. Schuller, *Fellow, IEEE*

***Abstract*— Emotion and a broader range of affective and cognitive states play an important role on the road. While this has been predominantly investigated in terms of driver safety, the approaching advent of autonomous vehicles (AVs) is expected to bring a fundamental shift in focus for emotion recognition in the car, from the driver to the passengers. This work presents a number of affect-enabled applications, including adapting the driving style for an emotional experience or tailoring the infotainment to personal preferences. It attempts to foresee upcoming challenges and provides suggestions for multimodal affect modelling, with a focus on the audio and visual modalities. In particular, this includes context awareness, reliable diarisation of multiple passengers, group affect, and personalisation. Finally, we provide some recommendations on future research directions, including explainability, privacy, and holistic modelling.**

***Index Terms*— Autonomous vehicles, interior sensing, human-machine interaction, emotion recognition.**

I. Introduction

CARS are becoming increasingly intelligent through more powerful on-board computational hardware, the ability to communicate with and receive over-the-air (OTA) updates from backend servers, and the integration of novel sensors. While some of these sensors are directed outwards, such as cameras and LIDAR used for parking and driving assistants [1], others are built into the cabin to monitor the occupants.

Among the occupants, the driver has traditionally been the focus of attention in order to improve driving safety [2], [3], [4]. Driver-facing cameras are used to assess distraction and fatigue by monitoring head and eye movements [5]. Steering wheel sensors can determine muscular tension and infer driver stress [6] and takeover readiness (TOR) [7]. Regulators have also proposed to have the car detect alcohol in the driver's breath and intervene or restrict operation if the driver is intoxicated [8].

Beyond driving-related sensory equipment (which also includes monitoring speed, braking behaviour and steering wheel movements [5], [9]), other sensors serve as interfaces with the increasing number of complex customer functions. This enables new forms of interaction, *e.g.*, neuromorphic vision [10] or radar-based [11] gesture control for easily accepting or rejecting incoming calls and adjusting sound volume, or voice-based commands that control an intelligent personal assistant (IPA) [12] like Alexa or Siri, which can engage in conversations, answer questions and make reservations, while adapting its output to the user's needs [13]. These new interaction modes, together with large touch displays, frequently complement or replace more traditional input devices like dials and buttons [14], as it is, *e.g*. easier to state an address than to type it into the navigation system.

As the automotive industry moves towards autonomous driving, tasks originally performed by the driver are increasingly taken over by the vehicle. This means that the driver is free to attend to other tasks, and that the distinction between driver and passenger becomes less important [2], [9]. It has been proposed that the vehicle of the future will be a kind of "living room on wheels" [15], allowing its occupants to perform a wide range of activities, *e.g.*, remote work. Advanced multimedia and infotainment systems are already installed in premium vehicles. Since this requires adding even more sophisticated and interconnected systems into the car, it is important to avoid the users getting overwhelmed. Thus, a smooth and natural Human-Machine-Interaction (HMI) is necessary.

Affective computing, which aims to provide machines with the ability to recognise, interpret, and influence emotions [16], has the potential to greatly improve the transportation experience. An emotionally intelligent car could continuously assess the emotions of its passengers and adjust its various interaction modes, *e.g*. the voice of its assistant, accordingly. It could change its lighting or even physically rearrange the interior, to ensure its occupants' wellbeing.

There have been extensive research efforts towards affect in the car. An early work which surveyed the influence of emotions and the potential to improve user interfaces (UIs) through affect sensing was [17]. Since then, many works have focused on elicitation, measurement, and effects of emotions in an automotive context [18], [19], [20], [21], [22], for instance, on frustration [23], [24], [25], [26]. Among prior literature, the majority focuses on the state of the driver, as their emotions are very relevant to driving performance and therefore safety. In particular, negative affective states can be dangerous, *e.g.*, drivers behaving recklessly as a result of experiencing road

This work was supported by the BMW Group. *(Corresponding author: Vincent Karas.)*

Vincent Karas is with the Chair EIHW, University of Augsburg, 86159 Augsburg, Germany, and also with the BMW AG, 80788 Munich, Germany (e-mail: vincent.karas@informatik.uni-augsburg.de).

Dagmar M. Schuller is with audEERING GmbH, 82205 Gilching, Germany (e-mail: ds@audeering.com).

Björn W. Schuller is with the Chair EIHW, University of Augsburg, 86159 Augsburg, Germany, also with audEERING GmbH, 82205 Gilching, Germany, and also with GLAM, Imperial College London, SW7 2BX London, U.K. (e-mail: bjoern.schuller@uni-a.de).

Digital Object Identifier 10.1109/TITS.2023.3333749

rage. Thus, it is important to recognise when a driver is frustrated or distressed, so the car can take countermeasures. While it is generally desirable to move the driver towards a state of positive valence and balanced arousal, choosing the appropriate mitigation strategy depends on the emotional state [18]. For driver affect, supervised methods analysing the face with convolutional neural networks (CNNs) are popular [27], [28], [29], [30]. Meanwhile, other recent works attempt to capture the cognitive state via unsupervised learning based on vehicle signals or smart devices [31], [32], [33], [34], [35]. Two recent surveys on the state of driver affect recognition and affective UIs respectively are [3] and [36]. Interaction technologies are surveyed in [5] and [9]. A research agenda by Vögel et al. [37] proposes a multi-disciplinary design approach for the development of emotion-aware vehicle assistants (EVAs), combining suggestions of researchers from the industry and academia.

We argue that it is necessary to go beyond the driver-focused perspective and examine the broader role of affect in vehicles where the driving task is increasingly automated. In this paper, we review current research and give an overview on how affective computing could be applied to semi-autonomous and autonomous vehicles (AVs). Our focus is on the audio and visual modalities, for the following reasons:

1) Microphones and cameras are already being installed in vehicles to enable, *e.g.*, speech and gesture command recognition. The same data could be used in emotion recognition, possibly without needing additional sensors, which is attractive from a cost perspective.
2) Both audio and visual data can be used to analyse the entire interior space, providing important context.
3) Passengers are not required to wear additional sensors, unlike *e.g.* for electroencephalogram (EEG) [38].
4) There is a general trend towards multimodal analysis in affective computing, and audio and video are popular choices [39], [40]. This may also appear natural to users, as the same modalities are available and familiar to humans.

In addition, we also touch upon other sensing methods, *e.g.*, physiological signals captured through dedicated devices. This is motivated by the popularity of vital parameters in driver monitoring [3], [41], the increasing availability of compact fitness wearables, and the potential to use this data for evaluating comfort [42].

Other related surveys examine only in-cabin sensing [5], or they are focused on the driver [3], [4], [43] or UI improvements [36] instead of the autonomous driving perspective. In studies that center on AVs, emotions may be considered as a contributor to accepting a new technology [44], [45], [46], [47], but the focus is often more on social and psychological factors [48], [49] or on safety and liability issues [50], [51]. Another related survey is the work of Xing et al. [52], who investigate cognitive factors and make recommendations for collaboration between humans and AVs. By comparison, we focus more on technical challenges of integrating affect into AVs.

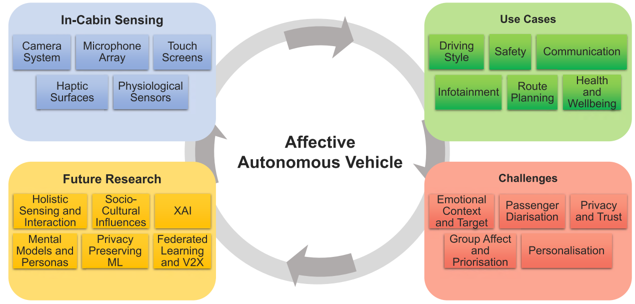

Fig. 1. Aspects concerning the development of affect-aware AVs. In-cabin sensors provide data needed for various use cases, including functions for emotional driving, infotainment and wellbeing. From these applications, challenges such as context-awareness, priorisation among passengers, and personalisation arise, motivating further research, which in turn leads to more sophisticated sensing technologies.

Our contribution is two-fold:

1) We identify existing works on affect recognition in the context of AVs and present applications.
2) We highlight challenges in creating an affect-aware vehicle and offer perspectives on future research directions.

The rest of this paper is structured as follows: We describe advances in sensors and functions for in-cabin monitoring in Sec. II. Applications of affect in AVs, such as use cases related to health and wellbeing, are discussed in Sec. III. Holistic in-cabin monitoring of all occupants creates technical challenges, which are presented in Sec. IV. Recommendations for future research directions are given in Sec. V. An overview of these topics is given in Sec. 1. Finally, Sec. VI sums up this paper.

II. Advances of Sensors and Functions for In-Cabin Monitoring

We present a selection of in-cabin monitoring solutions from the automotive industry, distinguishing between series-production technology and concept cars. Tab. I summarises those systems. An example of in-car interaction technologies is also sketched in Fig. 2.

Many manufacturers are introducing camera-based driver monitoring [53], in response to coming regulatory demands like Vision Zero (*i.e.*, many countries pursuing a zero accident goal [4], [54]), and related consumer protection standards, *e.g.*, Euro NCAP's Roadmap 2025 initiative [55]. Intelligent comfort functions, often showcased in concept vehicles, extend the sensing scope to the passengers.

A. Production Technology

GM uses *Super Cruise* for hands-free semi-autonomous driving on compatible highways. The vehicle tracks the road and independently executes manoeuvres like lane changes. Driver attention is checked through a camera, and the vehicle requests taking over the wheel if necessary [62]. A similar feature named *BlueCruise* is used by Ford and has been approved on selected highways in the US and UK [63].

Volvo has announced a camera-based system that will monitor driver distraction, intoxication, and dangerous behaviour in its upcoming generation of vehicles [64].

TABLE I

Overview of Recent Advances in In-Cabin Monitoring Solutions, Both in Concept Cars and Production Vehicles. Sensors for Audio, Visual, and Physiological Signals Are Abbreviated as A, V, P, Respectively

Company	**Year**	**Name**	**Sensors**	**User Functions**
			Concept Cars	
Audi	2017	Elaine [56]	A,P	personalised recommendations, revitalising programs
Audi	2019	AI:ME [57]	V	gaze-based menu selection, 3D-HUD
Mercedes Benz	2020	Vision ATR [58]	A,V,P	breathing pattern based identification
				adaptive illumination
				EEG-based function selection
Toyota	2018	Concept-i [59]	A,V	empathetic conversation
Toyota	2019	LQ [60]	A,V	emotional conversation
				adaptive illumination and music
				air flow/fragrance regulation
				seat adaptation for alertness/relaxation
Kia	2019	R.E.A.D [61]	V,P	emotion-based interior adaptation
			Production technology	
GM	2019	Super Cruise [62]	V	driver attention monitoring
Ford	2021	BlueCruise [63]	V	driver attention monitoring
Volvo	2019	driver monitoring system [64]	V	driver distraction/intoxication/behaviour monitoring
Nio	2017	NOMI [65]	A,V	personal assistant, emotional animations
BMW	2019	Driver Attention Camera [66]	V	driver distraction/fatigue monitoring
BMW	2020	Snapshot [67]	V	emotion-based selfie trigger
Mercedes Benz	2022	MBUX [68]	A,V	gesture control, adaptive illumination
Mercedes Benz	2023	DRIVE PILOT [69]	V	driver attention monitoring

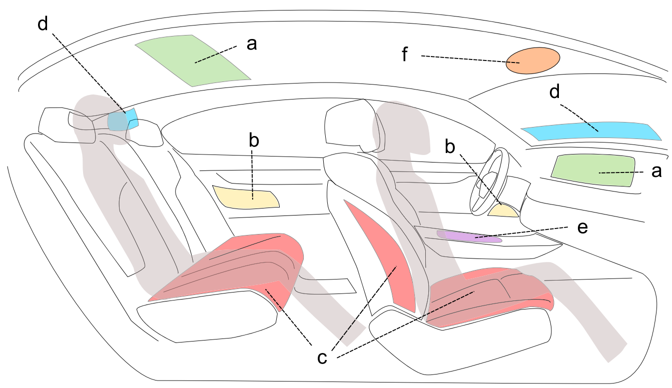

Fig. 2. Sensing and Interaction technologies in future intelligent vehicles: a) Infotainment displays, b) Audio system, c) Seat integrated sensors and comfort functions, d) VR rear seat entertainment / AR head-up display, e) Haptic control interface, and f) Camera system.

Nio introduced *NOMI* [65], an in-vehicle assistant that can understand spoken commands and can take selfies of the passengers with a backward-facing interior camera. NOMI is represented through a small sphere mounted atop the dashboard, which can turn towards the speaker to give the impression of looking at them. In addition, it displays cute animations to evoke an emotional reaction.

BMW has recently integrated a camera mounted overhead above the rear-view mirror in its iX model. The camera's location (*cf*. position e in Fig. 3) combined with its wide-angle lens allows it to see occupants on all seats. Passengers can take selfies, where the command can be given either by voice, touch, or implicitly by smiling. Snaphots can also be taken remotely via an app, or the car can send a picture of the interior if its anti-theft protection is triggered [67].

Mercedes Benz monitors the driver and co-driver with cameras for its *Mercedes-Benz User Experience (MBUX)* interior assistant, which can interpret gesture commands and control the interior lighting based on user activity. MBUX also responds to spoken commands and touch [70].

B. Concept Cars

In 2017, Audi presented *Elaine*, a concept car which contains a voice-controlled assistant that adapts to the driver's behaviour and makes personalised recommendations. The *Audi Fit Driver* function is supposed to improve health and wellbeing by monitoring vital parameters via a wearable. If stress or fatigue are detected, the car may activate revitalising programs [56]. In 2019, Audi AI:ME was introduced, a show car with an UI concept that includes eye-tracking. Infrared cameras monitor the eye movements, which allows for menu selection. In addition, the camera is used to track eye positions, in order to project a three-dimensional head-up display (HUD) [57].

Mercedes-Benz presented its VISION AVTR concept vehicle at CES 2020, which is supposed to recognise drivers by their breathing [58].

Toyota promises multi-modal emotion recognition through face, voice, and behaviour with its *Concept-i* prototype. The car would be capable of having empathetic conversations and offering assistance if the driver is stressed [59]. In 2019, an updated version called the *LQ* was presented. It includes an emotional AI named *Yui*, which can deliver a personalised mobility experience. Yui is capable of having conversations and playing music matching the drive environment, as well as changing illumination, air flow, and air fragrance. In addition, the driver seat can change shape to promote alertness or relaxation [60].

Kia presented *Real-Time Emotion Adaptive Driving (R.E.A.D)* at CES 2018, a concept that monitors facial expression and heart rate through camera and electrocardiogram (ECG) sensors to determine the emotional state and adapt the interior in real time [61].

TABLE II

USE CASES FOR AFFECT RECOGNITION IN AVS. THEY ARE BROADLY SEPARATED INTO CLUSTERS RELATED TO EXPERIENCING THE JOURNEY AND NON-DRIVING RELATED ACTIVITIES IN THE CAR

Cluster	Use case
Driving Experience	Natural and engaging driving style Promote trust and avoid anxiety in passengers Improve acceptance with other traffic participants Emotion-based route planning
Adaptive Interior	Customisation of cabin style and configuration Empathetic voice assistant and conversation partner Guide group interactions and mediate conflicts Provide immersive entertainment Create calm and productive work environment Optimise wellbeing programs

III. APPLICATIONS OF AFFECT IN AUTONOMOUS VEHICLES

We discuss selected applications for affective systems in the vehicle cabin. Some of these arise from the driving task, which is handled by the vehicle. When the vehicle is semi-automated, affect and cognitive states play a role in deciding when to take or cede control. In automated vehicles, affect is more related to driving behaviour, where the vehicle makes decisions and checks their impact on the passengers' emotional states. In addition, there are affective components in non-driving related functions such as infotainment and communications. We summarise the use cases in Tab. II.

A. Driving Style Adaptation

Emotions play an important role in driving, and the style of the driver can in turn have an impact on the emotions of the passengers, causing joy or thrill but also surprise or fear with "sporty" high acceleration manoeuvres, or calm and relaxation via a smooth ride. Current vehicles allow for manual selection of driving dynamics with different UIs, *e.g.*, BMW offers modes geared towards performance, comfort, and fuel efficiency [71]. In automated vehicles, driving is controlled not by a human but by algorithms that process data captured by external sensors, *e.g.*, cameras or LIDAR. Most existing research is focused on making these algorithms more reliable, whereas adaptation to drivers' preferences receives little attention [72]. Careful consideration needs to be given to the question of how and to which extent emotion is integrated into the driving style [73]. Otherwise, passengers might feel uneasy because the vehicle moves in a way that seems robotic and unnatural [74]. In addition, insecurity might arise from a lack of transparency [75], leaving the passengers wondering whether the vehicle judges the traffic situation correctly and knows when to perform emergency manoeuvres. To increase acceptance, Sini et al. [76] propose to adapt driving style based on passengers' facial expressions.

Issues of emotions towards and trust into an automated actor not only arise for the passengers inside the vehicle, but also for other traffic participants attempting to judge the AV's actions and intent from the outside [77]. A study by Dey et al. [78], examined the interactions of pedestrians with a 3-series BMW and a Renault Twizy. The cars were chosen for contrasting appearance: familiar and aggressive design (BMW) and futuristic and less threatening (Renault). Each was operated by a visible driver and by a hidden operator to simulate autonomous driving. The results suggested that speed and driving behaviour have a larger impact on pedestrian reactions than the physical appearance of the car. Wang et al. [79] examined how AVs could convey emotions to pedestrians, suggesting that the AV should combine emotional displays with movement cues, be not just proactive but responsive, and consider contextual and environmental information, *e.g.*, the weather. AVs could share data with each other and react faster than humans, allowing them to move in highly coordinated formations. Bera et al. [80] examined group emotions related to autonomous navigation, aiming to make it socially aware. They argued that *entitativity*, *i.e.*, similarity in appearance and behaviour, as well as physical proximity, would make fleets of autonomous cars stand out and elicit negative reactions from pedestrians. In a user study, emotional reactions to different entitativity profiles were examined, then, a navigation algorithm was developed to minimise negative reactions and increase social invisibility of vehicle groups.

Driver Affect already plays a role in customising the interior in current generation vehicles. For instance, BMW offers *MyModes*, allowing the driver to select visual and sound themes fitting their current mood. At CES 2022, two new modes named Relax and Expressive were presented [81]. AVs could extend this concept by adapting both driving and interior styles to create a more immersive emotional experience.

From this, it can be concluded that similar to a human driver, an AV should monitor the affect of its passengers and integrate that information into its driving behaviour to ensure a comfortable ride. Simultaneously, a more natural movement might help build trust with other traffic actors and improve acceptance.

B. Driving Safety

Apart from adjusting its driving style to build trust implicitly, an AV may also interact directly with the passengers. Here, affect again plays an important role to make the occupants feel safe.

Lee et al. conducted a simulated autonomous driving study, in which three different agents interacted with the drivers before and after events. Some participants perceived the conversational agent, which imitated natural human interaction, as more friendly and competent than agents which merely conveyed information. However, others complained that the agent was too verbose and distracting [82].

A study on semi-autonomous driving by Koo et al. investigated the effects of the information given on driving safety and driver satisfaction. Participants drove in a simulator where the car had an automatic braking function. The car would inform the drivers of an imminent action (*how*) or the reason for an action (*why*). The former decreased driving performance, while the latter improved it and was preferred. Giving both pieces of information caused negative emotions but the best driving performance. Thus, there is a trade-off between safety and comfort when the driving is not fully automated. Giving the full information may be necessary in a critical situation, but in non-critical situations, limiting to the *why* could help drivers process the situation better and feel more involved [83].

For highly or fully automated vehicles, human driver intervention with respect to road safety is rarely or never necessary. However, the car should offer explanations or comments on its actions to promote passenger trust and avoid anxiety [84]. Emotion recognition can be used to adjust the communication to convey the proper amount of information in an engaging or calming voice.

In case of an emergency, *e.g.*, an accident in which the vehicle is damaged, or an imminent danger, evacuation may be necessary. Here, the vehicle could act as coordinator of the evacuation procedure, and should take the affective states of its occupants into account when doing so. Literature on emotions in evacuations is mainly concerned with other transportation modes or crowds [85], however, insights might be transferred to emergency protocols for AVs.

An emotional vehicle may also help improve safety in non-driving related scenarios, when passengers are angry and upset with each other. Consider, *e.g.*, small children arguing or fighting on the rear seats, who may pose a risk to injuring themselves or damaging the interior. The vehicle may choose to either defer to an adult, or attempt to deescalate itself, depending on its reading of the situation.

C. Route Planning

Route Planning is another application where passengers' affect can play an important role. A human driver would consider the feelings and preferences of other people in the car when deciding on a route, and so should the AV. Here, affect recognition can help improve the ride experience by selecting routes that lead to positive emotions.

A user study by Braun et al. sought to identify use case clusters for affect in the car, based on interviews with German and Chinese testers of demo functions in a parked car. They found that navigation-related functions were in highest demand among participants of both cultures. Proposed use cases included finding routes based on positive experiences of other drivers, avoiding negative emotions by re-routing around difficult situations, and recommending parking spots for stress-free parking [86].

A navigation algorithm considering affective responses to the environment was proposed in [87]. Routes were generated from crowd-sourced data that linked map locations to emotional responses. A user study showed that routes which considered affect were preferred to taking the shortest path.

Besides selecting destinations based on aggregated journeys from other drivers, a personalised route planning system could also learn to revisit destinations the passengers liked in the past. Two examples of personalised tour recommendation algorithms integrating point-of-interest information and user behaviour are [88] and [89]. Passenger affect could also influence the willingness to take shorter or longer routes [90]. However, this might co-depend on other, non-affect related factors, making an accurate assessment difficult. Most of the participants in [86] cited desire to explore a new region, *e.g.*, on vacation, as motivation to try an affective navigation system. Chinese subjects in particular associated the function with family trips.

Acceptance of these route planning features is expected to be greater when they are seen as optional suggestions [86]. When it comes to selecting the route, the passengers should be involved in making the choice, and the AV might participate in the conversation to help guide them towards a consensus.

Knowledge of the affective states and preferences of passengers has the potential to enrich autonomous driving by tailored, scenic routes [91]. However, literature on the topic of affect-based route planning is still scarce, and the efficacy of the suggestions above is yet to be evaluated in studies on the road.

D. Communication

The AV can serve as a mediator between the passengers, and possibly also between passengers and people outside the vehicle. Being able to assess and convey affect in these interactions is of great importance.

During regular operation, the vehicle could help guide conversations, lighten mood and help solve conflicts between passengers. The extent of the car's interaction plays a crucial role here. Previous studies on driver emotions show that drivers benefit from and favour empathetic assistants [23], [92], [93]. However, they may reject help if they feel patronised [93]. In the broader context of affective UIs, the vehicle should know when and how to communicate [86], [92].

AVs can be thought of as service robots, which must meet wants and needs of users to build trust and acceptance. Emotional design can help with this goal, but careful considerations have to be made to account for the diversity of users [94].

E. Infotainment

With the reduced role of the driver, other activities become more important. Current vehicles already include a wide range of systems for information and entertainment, which are summed up into the concept of *infotainment*. Infotainment is commonly provided via touch displays mounted to or integrated inside the dashboard. There is a trend towards larger, merged displays, *e.g.*, Mercedes-Benz's *Hyperscreen*, which combines multiple displays under a common glass cover [95], and BMW's *iDrive Curved Display* [96]. In combination with the sound system and the vehicle's car-to-x network connectivity, these allow for reading news, listening to music, or playing games. Premium brands also include extensive rear seat infotainment options, *e.g.*, the retractable *Theatre Screen* presented by BMW at CES 2022 [97].

Infotainment systems in AVs could benefit greatly from incorporating passengers' affect. The vehicle may help sift through the large amount of available content by providing personalised recommendations for music playlists [98], video games, or movies and TV shows. The user preferences may be inferred implicitly from the browsing behaviour. Alternatively, the system may derive the current emotional state from a conversation and adapt its recommendations accordingly [99].

Future vehicles will allow their occupants to perform work-related tasks, *e.g.*, conference calls or writing and reviewing documents, while on the road. For these use cases, the car may provide an empathetic assistant that adjusts the UI and presents information based on the passenger's emotional state. This could help improve productivity by minimising stress and increasing focus.

F. Health and Wellbeing

Health and wellbeing are important factors in driving safety and enjoyment, and closely linked to the emotional state.

As discussed above, existing approaches focus on camera-based detection of the driver's state, *e.g.*, distraction, stress and fatigue. This functionality could be extended to all occupants with camera systems that can see the entire cabin. In addition, these systems could be used to detect major health risks, such as an imminent heart attack. The car could then stop and call for help, or navigate to a nearby hospital.

Camera based systems could also be used to monitor infants on the back seats. A direct camera stream is preferred to an abstract state description, so that parents can observe and react to the child's actions [86].

Physiological sensors are also used to estimate a variety of driving-related vital and cognitive states. Examples from the literature include stress level via EEG [100], [101] and heart rate variability (HRV) derived from ECG or blood volume pulse (BVP) from wearables [102], cognitive load [103], fatigue [104] and autonomic arousal [105] via HRV, comfort [42] via EEG and HRV, and vigilance and take-over readiness via EEG [106], [107]. Note that while those sensors may become more lightweight and comfortable to wear, there are competing research efforts to use cameras for remote sensing of vital signs [20], [108], [109]. Provided that they work reliably, car manufacturers may prefer these unobtrusive methods.

Currently, premium vehicles incorporate programs for improving driver wellbeing. For instance, BMW offers the *Caring Car* program, which either vitalises or relaxes the driver by changing ambient lighting, temperature, air flow, and playing music [110]. Mercedes Benz has created a suite of programs called *ENERGIZING*, which can adjust lighting and sound, activate seat massages, and scent the air. The feature can take drive duration, weather and traffic data into account, as well as integrating vital parameters from a wearable [68].

Future AVs could extend these wellbeing functions to all passengers. Affect recognition could then help the car detect the impact of the wellbeing program in real time and optimise its effectiveness. The vehicle may also adjust its interior, *e.g.*, the lighting, proactively to improve the mood of the occupants, or actively guide them towards a calm state of mind using, *e.g.*, music [111]. A study by Paredes et al. investigated controlled breathing exercises in a driving simulator. Participants were guided by voice commands or haptic stimuli in the seat. The interventions showed a lasting effect in reducing arousal, with haptic feedback being preferred as less obtrusive [112].

Another technology that has significant potential for both infotainment and wellbeing is virtual reality (*VR*). Displaying calming content like an ocean dive has been shown to reduce autonomic arousal, while the dynamic nature of the simulation helped reduce motion sickness in a moving vehicle [113].

IV. Challenges in Holistic in-Cabin Monitoring

We now discuss technical implications and challenges arising when moving from driver state monitoring towards passenger monitoring in future AVs.

A. Emotional Target and Context

The transition from manual to autonomous driving invalidates one of the central assumptions in many current systems and studies on in-vehicle affect: That the driving task is the target of the emotions. In an AV, affect may arise from passengers using the infotainment systems to watch movies or play games, talking to each other or having conversations with other persons remotely, speaking to the in-car assistant, built-in comfort functions attempting to enhance their wellbeing, or reacting to the vehicle dynamics and the view outside.

Thus, in order to respond to passenger affect in any of the use cases described in Sec. III, it is of paramount importance for the vehicle to understand the target of the affective sensation. For instance, when deciding on a driving behaviour adjustment to promote relaxation or excitement, the vehicle must know whether the passengers' affect is actually influenced by the driving style. If a passenger's feeling of distress or unhappiness is not directed towards vehicle movements, other actions such as changing the interior lighting or seat shape may be more effective for improving that person's wellbeing.

Separating the different influences on affect and determining the targets requires the intelligent vehicle to possess context awareness. However, there are still comparatively few works on context in affective computing, audiovisual or otherwise.

Learning contextual information requires datasets that contain rich scene information. In [114] and [115], the EMCO and EMOTIC corpora for visual emotion recognition in context are introduced. A larger dataset named CAER, compiled from TV show scenes, together with an architecture that encodes and fuses facial and context information, is presented in [116]. Joint emotion and game context recognition on live-streams of the popular e-sports title *League of Legends* is performed in [117].

For vehicles, the majority of work is again focused on situations that trigger negative emotions, since these are most relevant to driving safety [36]. An evaluation of 531 self-reports of 33 participants, recorded while driving, found four main clusters related to affect. Traffic-related events and issues with navigation systems primarily caused negative emotions, while the vehicle's equipment and nice environments were mostly leading to positive emotions. An asynchronous online survey with 170 participants by Braun et al. found, unsurprisingly, that dangerous situation and drivers being unsatisfied with their own performance triggered negative emotions, while positive emotions arose from satisfaction with vehicle performance and driving skills [118].

B. Passenger Diarisation

Another challenge arises due to the possibility of multiple people being in the car simultaneously. In this case, an audio-visual system needs to perform a separation in terms of who

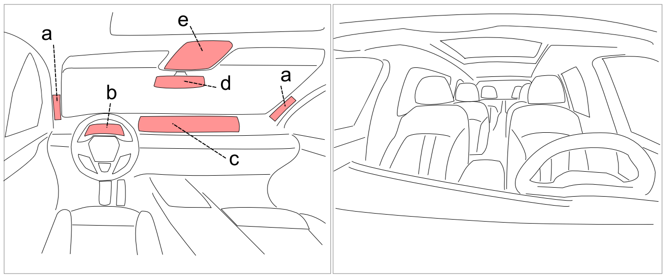

Fig. 3. In-cabin camera system for a current generation vehicle interior. Left: several possible locations for backward-facing cameras; a) A-pillars, b) driver-facing, c) dashboard, d) rearview mirror, e) car roof. Right: field of view for a camera mounted near the top center of the windscreen. Based on camera placement and passenger activity, partial occlusions and extreme angles may occur. In AVs, the camera systems will likely need to account for greater freedom of movement, as well as different interior configurations.

is speaking when (*speaker diarisation*), as well as identifying multiple persons in the camera footage (*video segmentation*).

While passenger diarisation can seem rather trivial today, it may become more challenging as we move towards autonomous cars. In present day vehicles, there is a clear assignment of roles based on the seat occupation (driver, co-driver, rear seat passengers). This distinction is expected to disappear with autonomous driving. Advances in safety may allow passengers to be less restricted by their seats than is the case today. In addition, the cabin itself may transform to accommodate various user needs, *e.g.*, moving and rotating seats, or retracting consoles for greater freedom of movement. Thus, the assumption that passengers occupy fixed positions, which change only between drives, no longer holds.

Fortunately, re-identification of persons in video is a fairly mature area of research [119], [120], [121], as is speaker diarisation in combination with speech recognition systems [122]. For vehicles, identification via smartphone data has also been proposed [123], although this is more error-prone if devices and persons can move freely in the cabin. The simultaneous detection of multiple faces is also reasonably well addressed in current computer vision algorithms. However, the vehicle environment presents challenges that dynamically impact how well each passenger's emotion can be recognised. These can involve illumination (*e.g.*, rapid changes due to entering a tunnel or driving along a road partially shadowed by trees, glare from low sun), unfavourable angles (*e.g.*, looking out the side window with the face in profile), or partial occlusions due to camera placement and occupants' movements and activities (*e.g.*, leaning across seats, gestures). All of these factors can negatively impact detection and tracking [4], leading to performance losses in face and body pose based emotion recognition. The in-cabin monitoring system will need to be designed with these issues in mind, *e.g.*, by switching to infrared cameras in low light conditions or integrating additional sensors like LIDAR. Possible camera placements at the front of the cabin, as well as an example field of view are illustrated in Fig. 3.

Speech Emotion Recognition (SER) in the vehicle faces issues of noise, due to external sources like vibrations from wind and the road surface, as well as internal sources like people talking simultaneously or music [124]. External noise may be reduced effectively due to quiet electric drives and improved acoustic design. However, manufacturers may add artificial sounds instead to make the driving more immersive. Despite advances due to deep learning, SER inside vehicles is currently still suffering from performance issues [125].

Algorithms for automatic in-car affect and emotion recognition will need to solve the challenges outlined above to identify each individual in the cabin and access their affective states even when data from some modalities is temporarily degraded, as is the case with occlusions in video streams. Detecting affect in such challenging situations is, however, still largely unexplored in the literature today.

C. Group Affect and Priorisation

Another challenge related to the presence of multiple individuals in the car is the existence of a group-level affect state, which a holistic in-cabin sensing system should be able to detect. Automatic group emotion recognition received comparatively little attention in the past [126], but has garnered increasing interest in recent years. A comprehensive survey on current approaches is given by Veltmeijer et al. [127].

Group-level affect can be considered bottom-up, also called local context, arising from emotions of individuals, or top-down derived influenced by the environment and called global context, or as a hybrid mixture of the two [127]. Hybrid approaches frequently combine face and scene information using CNNs or visual attention [128], [129], [130].

While advances in group emotion recognition have been made, existing approaches have shortcomings, such as assuming the group to share one common emotion, largely relying on just one modality and not considering the evolution of emotions over time [127]. Group affect in the car is rarely investigated, one such work being a study of Alyuz et al., who examine the emotions of driver-passenger dyads in simulated autonomous driving [131].

Beyond detecting group affect, the presence of multiple occupants necessitates some form of priorisation strategy which enables the car to weigh their wants and needs before making a decision. For instance, the car could attempt to find a course of action that maximises satisfaction for all occupants, or it could strive to specifically improve the wellbeing of passengers that are feeling very uncomfortable, *e.g.*, when choosing the driving style.

Affect-based priorisation among multiple users is mostly a white space in the literature. In cars, the final decision is usually made by the driver. However, some works exist for other forms of transportation. Passengers of Chinese high-speed rail (HSR) were surveyed on factors influencing in cabin comfort in [132]. A fuzzy linguistic approach was then used to reach a consensus on the most important factors.

D. Personalisation

Current affect recognition systems mostly follow the paradigms of *(semi-)supervised learning*, in which a model is trained on a set of (partially) human-labelled samples and then deployed. The model is then expected to generalise well across diverse users. However, in the use cases described in Sec. III, there exists a significant potential to use affective feedback to continuously improve the vehicle. Passengers spend extended periods of time travelling, and may regularly use the same car for years. This gives plenty of opportunities for the empathetic car to get to know its occupants and learn to better recognise their feelings and preferences.

Learning on a daily basis in a wide range of situations has the additional benefit of contributing to overcome the bottleneck of limited data in affective computing. When collecting data during the drive, instead of just performing inference, the model can be adapted. One way to do this is by *reinforcement learning* (RL). Here, the vehicle could use the reactions of the passengers to better assess the affect. For instance, in the driving style adaptation use case, the vehicle may detect fear from a passenger and decide to drive more conservatively. If it then detects that the passenger's affect is not improving or becoming more negative, indicating dissatisfaction, the vehicle may rethink its assessment of fear for future situations. RL has the advantage that it can function without the need for explicit human labelling. Instead, the agent makes decisions based on a reward function and the positive or negative feedback it receives. The learning process could also be considered weakly supervised or semi-supervised when the interaction context is used to guide the model. As an example, the reward function could be long-term maximisation of wellbeing. The car could then observe how the passengers' affect evolves, and optimise its actions towards that goal. At the same time, these interactions may help improve its recognition of a user's emotions. While reinforcement learning for AVs has recently attracted much interest [133], research focuses predominantly on problems like motion planning [134], routing [135], decision-making [136], vehicle control [137], [138], understanding pedestrians [139], and cyber-physical safety [140], [141]. Applications to affective computing are still scarce, but have shown promising results [142]. For instance, Ling et al. propose a framework for emotionally adaptive AVs using a driving simulator and EEG sensors. Fuzzy logic is used to infer driver preferences from emotional states. The vehicle behaviour is controlled by a RL algorithm, with a composite reward function combining safety with personalisation [72].

An alternative to reinforcement would be *active learning*, where the vehicle gathers data, then asks a human for feedback to improve its decision-making in cases where it is uncertain. This may be less convenient than the implicit feedback used by RL. Users constantly asked for comment would be distracted and annoyed [37]. On the other hand, given that users value having personal choice in their interactions with the vehicle [86], they may welcome it asking for their input, provided it is not done in an obnoxious or patronising manner.

Another option to utilise the large amounts of data collected in the vehicle would be *self-supervised learning*, which does not require explicit labelling. In this learning paradigm, models are trained on pretext tasks that teach them information implicitly contained in the data, *e.g.*, the spatial orientation of an object. The pre-trained model can then be adapted to a variety of downstream tasks. This approach has been proven to be effective in natural language processing with BERT [143] and other Transformer-based architectures [144], and there are works on using self-supervision for autonomous navigation [145], [146]. In affective computing, it has improved the performance of SER [147]. Combining visual and audio data for self-supervision can help generate more robust features for emotion recognition in noisy environments by encoding complementary information, outperforming unimodal approaches [148]. Self-supervision is of particular interest for affective computing since large amounts of audiovisual data are readily available, but high-quality annotation of affect is expensive and therefore only done on small datasets [149]. This encourages cross-modal, cross-domain or cross-corpus approaches which pre-train on large unlabeled corpora and then fine-tune on a small labeled dataset [148], [150].

E. Privacy

Personalisation of the in-vehicle experience also raises issues of privacy. Given the growing amount of sensors in the car, and the networked nature of modern vehicles, which will only increase with the shift to autonomous driving [151], protecting the passengers' personal data becomes more important than ever. With holistic in-cabin monitoring, the vehicle could potentially learn an enormous amount of information on its users, including highly sensitive health data, habits, aspects of their personal life, *etc*. At the same time, the automotive industry is incentivised to monetise this data, while third parties, *e.g.* insurance companies, seek to gain access to it [53]. Data from built-in sensors can potentially reveal a lot more than was originally intended [151], *e.g.*, a camera system for object detection and passenger identification may also be used to detect health and affect. Audiovisual sensing through camera systems and microphone arrays also has the inherent disadvantage that these are potentially always on, capturing information. This might lead some users to feel they are under constant surveillance by their car [152], and fear that their recording may be stored and analysed somewhere beyond their control.

V. Discussion and Future Development Recommendations

We have given a selection of potential use cases for affect in AVs in Sec. III, and presented some technical challenges in Sec. IV. This list does not claim to be comprehensive or complete, since technological advances and social trends will certainly open up new opportunities and issues. Here, we discuss our findings and make suggestions for future research directions, which we summarise in Tab. III.

A. Safety and Comfort

We have touched upon the relation of safety and affect in particular in Sec. III-A and Sec. III-B. However, this was through the lens of passengers' perceptions of trust and safety, and how integrating affective feedback can be beneficial. This discussion would not be complete without examining how affect might be harmful, if inappropriately used.

Before any affective use cases from Sec. III can be implemented in series production, these functions will need to be

TABLE III

FUTURE RESEARCH OPPORTUNITIES FOR AFFECTIVE AUTONOMOUS VEHICLES. THEY ARE BROADLY ORGANISED IN CHALLENGE CLUSTERS, RELATED TO: ATTITUDE TOWARDS AVS (BY PASSENGERS AND THE PUBLIC), PERSONALISED VEHICLES AND IMMERSIVE USER EXPERIENCE DRIVEN BY HOLISTIC SENSING AND MODELS OF THE CAR INTERIOR

Challenge cluster	**Research direction recommendations**
Improve acceptance of AVs	Anthropomorphisation to build trust [153], [154], [155] Intuitive explanations for passengers [83], [156], [52], [84] Intent communication to other road users [78], [77], [157], [158], [159], [79]
Create a personalised vehicle	Emotional personal assistant [37], [160], [13], [161], [12], [72], [162] Privacy preservation [163], [164], [165], [166]
Provide an immersive, intelligent in-cabin experience	Affect-responsive UI [167], [168], [169], [118], [93], [36] New sensing and interaction technologies [170], [82], [14], [5], [9] Context awareness [80], [131], [171], [172], [22]

carefully evaluated in terms of how they interact with the primary goal of of a (semi-)autonomous vehicle, which is to ensure the safety of its occupants and other members of traffic.

The affective components should be at least neutral in regards to safety, *i.e.*, there should be no way for them to impact critical functions in a dangerous manner. For instance, in case of driving style adaptation, there should obviously be no way for extreme emotions to override basic heuristics for safe driving. In other words, the AV should not behave so human-like that it starts road-raging or thrill seeking.

For use cases that are not directly related to driving but more to comfort, *e.g.*, infotainment or health and wellbeing, safety issues appear in the context of semi-autonomous driving, when there is the potential of the driver having to take over in case of an emergency but being distracted. The current solution is to bar the driver from using such functions, at least while the vehicle is on the road, or to limit their availability.

In a fully autonomous vehicle, comfort functions have less of an immediate impact on transportation safety, and more on the sense of safety of the passengers. In particular, the interaction with the emotional vehicle assistant that governs the functions has to be carefully designed so that is does not cause anxiety or distress. For instance, it may monitor the passengers' vital parameters to improve their wellbeing, but it should not make insensitive comments about their health. Having multiple passengers in the cabin also brings the risk of the vehicle inadvertently causing harm by revealing personal information about one of them. A better understanding of group affect might help prevent such situations. Knowing when and how to convey information is a key competency of an empathetic conversation partner, and should be comprehensively investigated in future work.

B. Trust and Explainability

As described in Sec. III-C and Sec. III-B, trust plays a key role in the acceptance of AVs. This is closely linked to affect, since passengers can be expected to feel apprehensive and interact more hesitantly with the car when they do not fully trust it. In order to provide the best possible experience to customers, vehicle manufacturers need to be aware of the interplay of psychological factors like trust and affect, and incorporate this knowledge into the design [36], [160]. An empathetic vehicle that has a concept of trust could deliberately adjust its actions, *e.g.*, by initially exhibiting slower, seemingly more careful driving to prevent anxiety, and moving towards a more daring, but still safe style once its occupants have learnt to be confident in its capabilities.

Likewise, the vehicle may integrate trust and cognitive load, together with affect estimates, into its infotainment and comfort functions. This could help strike a balance in terms of the amount and type of information presented [156]. While passengers might get frustrated with an overwhelming amount of customisation options, they may, on the other hand, resent the vehicle for trying to explain everything to them, which may seem overbearing or patronising.

Explainability and trust are closely linked, especially when dealing with HMI. Passengers may be inclined to mistrust a disembodied intelligence, whose reasoning they see as non-human. Besides making the design more anthropomorphic to build trust [94], [153], [154], having the car be able to explain its decisions, would be desirable. Simply displaying the results of in-cabin analysis, such as the emotional state, may cause negative reactions [160]. Instead, the car should preferably communicate in an intuitive manner such as natural speech, to appear more as a helpful companion than a detached observer.

Explainable AI (XAI) is an active and growing research field, due to the pervasiveness of deep models in many aspects of life and the need to make sense of their decisions [173], [174], [175], [176]. Using XAI to investigate intelligent assistants for AVs would be an interesting research avenue.

C. Personalisation and Privacy

As shown in Sec. IV-D, personalisation can play an important role in helping the vehicle understand the subjective preferences of its passengers and improving the user experience. In particular, emotional personalisation is promising [36]. Beyond the technical challenges of updating the on-board intelligence with, *e.g.*, reinforcement learning, there are also psychological and ethical questions.

While personalisation can be very beneficial, it also involves making specific inferences about individuals. How these insights are gathered and how they are applied will be of critical importance for how the intelligent vehicle is perceived. As demonstrated by the literature, users welcome opportunities for personalising their car, but they want personal choice [86]. One possible solution would be to run the adaptation of the affect recognition system largely implicitly, but not changing the behaviour of use cases without more explicit input. The vehicle should involve the users in the determination of their preferences by presenting them a selection of options, so the users feel that they have agency. Nevertheless, the built-in assistant could still be proactive in its suggestions and perform a certain level of filtering options, if desired.

For privacy protection, it is recommended that the analysis of in-cabin sensor streams be run on on-board hardware. Note that this may not always be possible for computationally heavy tasks, such as question answering by intelligent assistants. However, there is significant potential in making, *e.g.*, computer vision and audio algorithms run efficiently on embedded hardware [177], which has the added benefit of saving costs and energy consumption. In situations where visual data has to be sent to a backend, facial anonymisation via GANs [165], [166], [178] is an interesting research direction.

D. Cultural and Social Influences

Cultural and social factors play an important role when designing affect-based functions for AVs. Manufacturers who want to appeal to global markets need to consider this to improve sales and customer satisfaction.

On the one hand, acceptance and demand towards certain clusters of functions may be impacted by socio-cultural context [44], [48]. For instance, functions that let users share emotional experiences may be highly sought in some cultures or social groups but considered at best an unnecessary gimmick in others. While highly dependent of the individual, the desire for personalisation of the vehicle is also likely to be impacted by social norms regarding personal expression. Cultural attitudes may also change over time, defying initial hypotheses. As reported in [86], German subjects were more willing to share emotion detection data than Chinese subjects, indicating shifts on privacy.

On the other hand, cultural background and origin also has a significant impact on how people experience and express emotions [179], which needs to be considered for automatic affect recognition. Systems trained on insufficiently diverse data, *e.g.*, only on a small subset of cultures, will be less reliable when performing inference on people from a previously unseen culture. Cross-cultural affective computing is explored by relatively few works yet, but it is an important area of research to make emotion-aware systems ready for large-scale use [177].

E. Theoretical Frameworks

As the industry progresses towards autonomous driving and better in-cabin sensing, much thought will need to be devoted to solving challenges like the ones listed in Sec. IV. While we have given a high-level overview here, manufacturers will need to address numerous low-level technical issues, within the constraints of commercial product development. Since affect-enabled intelligent vehicles will be highly complex, future work should also focus on developing broader theoretical frameworks.

For instance, one way to make an autonomous car more appealing and improve acceptance would be to consider certain profiles, or personas, in the design process. However, what these personas should look like is still largely unexplored for AVs, and would need to be carefully and systematically validated across large populations [94].

As discussed above, turning the vehicle into an empathetic, human-like companion should include a mental model of the passengers [156] that not only encompasses affect and specific emotions, but a broader psychological picture involving, *e.g.*, trust and cognitive load. Future research should be multi-disciplinary, involving psychologists and engineers. Emotions should be studied in detail under realistic conditions, and across many contexts [179].

To create a truly context-aware vehicle, as touched upon in Sec. IV-A, there needs to be a sophisticated cabin model, including the various UIs, built-in and connected sensors, customer functions, as well as passenger states and external influences. Such a model could also help combine interactive technologies in new and interesting ways. For instance, haptic surfaces may replace traditional buttons while maintaining tactile feedback and providing distinctive sensations [5], [14], [170]. Furthermore, soft materials like seats can be enhanced via stretchable electronics [180]. Combining haptics with affect can create a more immersive user experience, and affect detection through touch remains an open research area [181].

In addition, means of sharing information between vehicles (V2V), while protecting personal data, should be investigated. Manufacturers could use rich data from vehicle fleets to improve algorithms and provide OTA updates. Edge computing is highly relevant here [30], [182], as are distributed methods like federated learning with differential privacy [164]. Privacy preservation in AVs is a highly topical research field [166], [183], as is fairness in machine learning [184].

VI. Conclusion

This paper examined automatic affect recognition and its applications to future generations of autonomous vehicles. It highlighted challenges and offered perspectives on future research opportunities. We based our discussion on the observation that in-car sensing technologies and customer functions are becoming increasingly numerous and sophisticated and extrapolate that, as we move towards large-scale autonomous driving, the focus will shift from driver affect to passenger affect, enabling emotionally intelligent vehicles to adapt to their occupants. An exemplary selection of affect-enabled use cases and technical challenges was presented. Since the majority of prior work is focused on driver affect, some of the claims made herein will need further studies to substantiate them. As new technologies emerge and the wants and needs of passengers in autonomous driving become clearer, parts of this work will undoubtedly need to be revised. Nevertheless, one can assume a major shift in how affect is recognised and used in (semi-)autonomous vehicles of the future.

Acknowledgment

The authors would like to thank Prajval Kumar Murali for fruitful discussion and advice on preparing the figures in this article.

References

[1] S. Mozaffari, O. Y. Al-Jarrah, M. Dianati, P. Jennings, and A. Mouzakitis, "Deep learning-based vehicle behavior prediction for autonomous driving applications: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 1, pp. 33–47, Jan. 2022.

[2] A. Koesdwiady, R. Soua, F. Karray, and M. S. Kamel, "Recent trends in driver safety monitoring systems: State of the art and challenges," *IEEE Trans. Veh. Technol.*, vol. 66, no. 6, pp. 4550–4563, Jun. 2017.

[3] S. Zepf, J. Hernandez, A. Schmitt, W. Minker, and R. W. Picard, "Driver emotion recognition for intelligent vehicles: A survey," *ACM Comput. Surveys*, vol. 53, no. 3, pp. 1–30, May 2021, doi: 10.1145/3388790.

[4] J. Wang et al., "A survey on driver behavior analysis from in-vehicle cameras," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 10186–10209, Aug. 2022.

[5] P. K. Murali, M. Kaboli, and R. Dahiya, "Intelligent in-vehicle interaction technologies," *Adv. Intell. Syst.*, vol. 4, no. 2, Feb. 2022, Art. no. 2100122, doi: 10.1002/aisy.202100122.

[6] P. E. Paredes, F. Ordonez, W. Ju, and J. A. Landay, "Fast & furious: Detecting stress with a car steering wheel," in *Proc. CHI Conf. Human Factors Comput. Syst.* New York, NY, USA: Association for Computing Machinery, Apr. 2018, pp. 1–12, doi: 10.1145/3173574.3174239.

[7] C. Lv et al., "Human–machine collaboration for automated driving using an intelligent two-phase haptic interface," *Adv. Intell. Syst.*, vol. 3, no. 4, Apr. 2021, Art. no. 2000229, doi: 10.1002/aisy.202000229.

[8] US Government Publishing Office. (Nov. 2021). *H.R.3684—Infrastructure Investment and Jobs Act*. Accessed: Jun. 10, 2022. [Online]. Available: https://www.congress.gov/bill/117th-congress/house-bill/3684/text

[9] Z. Tan et al., "Human–machine interaction in intelligent and connected vehicles: A review of status quo, issues, and opportunities," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 9, pp. 13954–13975, Sep. 2022.

[10] G. Chen et al., "NeuroIV: Neuromorphic vision meets intelligent vehicle towards safe driving with a new database and baseline evaluations," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 2, pp. 1171–1183, Feb. 2022.

[11] L. Zheng et al., "Dynamic hand gesture recognition in in-vehicle environment based on FMCW radar and transformer," *Sensors*, vol. 21, no. 19, p. 6368, Sep. 2021.

[12] A. de Barcelos Silva et al., "Intelligent personal assistants: A systematic literature review," *Express Syst. Appl.*, vol. 147, Jun. 2020, Art. no. 113193. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417420300191

[13] J. Park, M. Abdel-Aty, Y. Wu, and I. Mattei, "Enhancing in-vehicle driving assistance information under connected vehicle environment," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 9, pp. 3558–3567, Sep. 2019.

[14] S. J. Breitschaft, A. Pastukhov, and C.-C. Carbon, "Where's my button? Evaluating the user experience of surface haptics in featureless automotive user interfaces," *IEEE Trans. Haptics*, vol. 15, no. 2, pp. 292–303, Apr. 2022.

[15] Audi. (Sep. 2021). *Audi Grandsphere Concept*. Accessed: Jun. 10, 2022. [Online]. Available: https://www.audi-mediacenter.com/en/audi-grandsphere-concept-2021-14199

[16] R. W. Picard, "Affective computing: From laughter to IEEE," *IEEE Trans. Affect. Comput.*, vol. 1, no. 1, pp. 11–17, Jan. 2010.

[17] F. Eyben et al., "Emotion on the road—Necessity, acceptance, and feasibility of affective computing in the car," *Adv. Hum.-Comput. Interact.*, vol. 2010, pp. 1–17, 2010.

[18] M. Jeon, B. N. Walker, and J.-B. Yim, "Effects of specific emotions on subjective judgment, driving performance, and perceived workload," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 24, pp. 197–209, May 2014.

[19] X. Wang, Y. Guo, C. Bai, Q. Yuan, S. Liu, and J. Han, "Driver's intention identification with the involvement of emotional factors in two-lane roads," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 11, pp. 6866–6874, Nov. 2021.

[20] G. Du, Z. Wang, B. Gao, S. Mumtaz, K. M. Abualnaja, and C. Du, "A convolution bidirectional long short-term memory neural network for driver emotion recognition," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4570–4578, Jul. 2021.

[21] H. Kim, S. Kim, H. Kim, Y. Ji, and C.-H. Im, "Modulation of driver's emotional states by manipulating in-vehicle environment: Validation with biosignals recorded in an actual car environment," *IEEE Trans. Affect. Comput.*, vol. 13, no. 4, pp. 1783–1792, Oct. 2022.

[22] L. Mou et al., "Driver emotion recognition with a hybrid attentional multimodal fusion framework," *IEEE Trans. Affect. Comput.*, early access, Feb. 28, 2023, doi: 10.1109/TAFFC.2023.3250460.

[23] H. Harris and C. Nass, "Emotion regulation for frustrating driving contexts," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.* New York, NY, USA: Association for Computing Machinery, May 2011, pp. 749–752, doi: 10.1145/1978942.1979050.

[24] K. Ihme, C. Dömeland, M. Freese, and M. Jipp, "Frustration in the face of the driver: A simulator study on facial muscle activity during frustrated driving," *Interact. Stud.*, vol. 19, no. 3, pp. 487–498, Dec. 2018.

[25] A. Joshi, Y. Attia, and T. Mishra, "Protocol for eliciting driver frustration in an in-vehicle environment," in *Proc. 8th Int. Conf. Affect. Comput. Intell. Interact. (ACII)*, Sep. 2019, pp. 620–626.

[26] M. Oehl, M. Lienhop, and K. Ihme, "Mitigating frustration in the car: Which emotion regulation strategies might work for different age groups?" in *HCI International 2021—Posters*, C. Stephanidis, M. Antona, and S. Ntoa, Eds. Cham, Switzerland: Springer, 2021, pp. 273–280.

[27] Y.-L. Wu, H.-Y. Tsai, Y.-C. Huang, and B.-H. Chen, "Accurate emotion recognition for driving risk prevention in driver monitoring system," in *Proc. IEEE 7th Global Conf. Consum. Electron. (GCCE)*, Oct. 2018, pp. 796–797.

[28] H. Xiao et al., "On-road driver emotion recognition using facial expression," *Appl. Sci.*, vol. 12, no. 2, p. 807, Jan. 2022.

[29] W. Li et al., "CogEmoNet: A cognitive-feature-augmented driver emotion recognition model for smart cockpit," *IEEE Trans. Computat. Social Syst.*, vol. 9, no. 3, pp. 667–678, Jun. 2022.

[30] G. Muhammad and M. S. Hossain, "Light deep models for cognitive computing in intelligent transportation systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 1, pp. 1144–1152, Jan. 2023.

[31] E. G. Mantouka, E. N. Barmpounakis, and E. I. Vlahogianni, "Identifying driving safety profiles from smartphone data using unsupervised learning," *Saf. Sci.*, vol. 119, pp. 84–90, Nov. 2019.

[32] S. Ansari, H. Du, F. Naghdy, and D. Stirling, "Unsupervised patterns of driver mental fatigue state based on head posture using Gaussian mixture model," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Dec. 2020, pp. 2699–2704.

[33] Y. Xie and Y. L. Murphey, "Unsupervised driver workload learning through domain adaptation from temporal signals," in *Proc. IEEE Conf. Evolving Adapt. Intell. Syst. (EAIS)*, May 2020, pp. 1–8.

[34] B. Li et al., "A new unsupervised deep learning algorithm for fine-grained detection of driver distraction," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 10, pp. 19272–19284, Oct. 2022.

[35] E. de Salis et al., "Clustering of drivers' state before takeover situations based on physiological features using unsupervised machine learning," in *Human Interaction, Emerging Technologies and Future Systems*, T. Ahram and R. Taiar, Eds. Cham, Switzerland: Springer, 2022, pp. 550–555.

[36] M. Braun, F. Weber, and F. Alt, "Affective automotive user interfaces–reviewing the state of driver affect research and emotion regulation in the car," *ACM Comput. Surveys*, vol. 54, no. 7, pp. 1–26, Sep. 2021.

[37] H.-J. Vögel et al., "Emotion-awareness for intelligent vehicle assistants: A research agenda," in *Proc. IEEE/ACM 1st Int. Workshop Softw. Eng. AI Auto. Syst. (SEFAIAS)*. Gothenburg, Sweden: IEEE, May 2018, pp. 11–15.

[38] C. Park, S. Shahrdar, and M. Nojoumian, "EEG-based classification of emotional state using an autonomous vehicle simulator," in *Proc. IEEE 10th Sensor Array Multichannel Signal Process. Workshop (SAM)*, Jul. 2018, pp. 297–300.

[39] S. K. D'mello and J. Kory, "A review and meta-analysis of multimodal affect detection systems," *ACM Comput. Surveys*, vol. 47, no. 3, p. 43, 2015.

[40] S. Poria, E. Cambria, R. Bajpai, and A. Hussain, "A review of affective computing: From unimodal analysis to multimodal fusion," *Inf. Fusion*, vol. 37, pp. 98–125, Sep. 2017.

[41] A. J. Béquet, A. R. Hidalgo-Muñoz, and C. Jallais, "Towards mindless stress regulation in advanced driver assistance systems: A systematic review," *Frontiers Psychol.*, vol. 11, Dec. 2020, Art. no. 609124. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.609124

[42] H. Su and Y. Jia, "Study of human comfort in autonomous vehicles using wearable sensors," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 11490–11504, Aug. 2022.

[43] M. Q. Khan and S. Lee, "A comprehensive survey of driving monitoring and assistance systems," *Sensors*, vol. 19, no. 11, p. 2574, Jun. 2019.

[44] N. Charness, J. S. Yoon, D. Souders, C. Stothart, and C. Yehnert, "Predictors of attitudes toward autonomous vehicles: The roles of age, gender, prior knowledge, and personality," *Frontiers Psychol.*, vol. 9, p. 9, Dec. 2018, Art. no. 2589. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02589

[45] M. Raue, L. A. D'Ambrosio, C. Ward, C. Lee, C. Jacquillat, and J. F. Coughlin, "The influence of feelings while driving regular cars on the perception and acceptance of self-driving cars," *Risk Anal.*, vol. 39, no. 2, pp. 358–374, Feb. 2019.

[46] X. Wang, Y. D. Wong, K. X. Li, and K. F. Yuen, "This is not me! Technology-identity concerns in consumers' acceptance of autonomous vehicle technology," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 74, pp. 345–360, Oct. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1369847820304423

[47] M. A. Ribeiro, D. Gursoy, and O. H. Chi, "Customer acceptance of autonomous vehicles in travel and tourism," *J. Travel Res.*, vol. 61, no. 3, pp. 620–636, Feb. 2021, doi: 10.1177/0047287521993578.

[48] T. Zhang et al., "Automated vehicle acceptance in China: Social influence and initial trust are key determinants," *Transp. Res. C, Emerg. Technol.*, vol. 112, pp. 220–233, Mar. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X19312537

[49] P. Jing, G. Xu, Y. Chen, Y. Shi, and F. Zhan, "The determinants behind the acceptance of autonomous vehicles: A systematic review," *Sustainability*, vol. 12, no. 5, p. 1719, Feb. 2020.

[50] A. Taeihagh and H. S. M. Lim, "Governing autonomous vehicles: Emerging responses for safety, liability, privacy, cybersecurity, and industry risks," *Transp. Rev.*, vol. 39, no. 1, pp. 103–128, Jan. 2019, doi: 10.1080/01441647.2018.1494640.

[51] K. Othman, "Public acceptance and perception of autonomous vehicles: A comprehensive review," *AI Ethics*, vol. 1, no. 3, pp. 355–387, Aug. 2021, doi: 10.1007/s43681-021-00041-8.

[52] Y. Xing, C. Lv, D. Cao, and P. Hang, "Toward human-vehicle collaboration: Review and perspectives on human-centered collaborative automated driving," *Transp. Res. C, Emerg. Technol.*, vol. 128, Jul. 2021, Art. no. 103199.

[53] A. McStay and L. Urquhart, "In cars (are we really safest of all?): Interior sensing and emotional opacity," *Int. Rev. Law, Comput. Technol.*, vol. 36, no. 3, pp. 470–493, Sep. 2022, doi: 10.1080/13600869.2021.2009181.

[54] M. Martínez-Díaz and F. Soriguera, "Autonomous vehicles: Theoretical and practical challenges," *Transp. Res. Proc.*, vol. 33, pp. 275–282, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352146518302606

[55] Euro NCAP. (2017). *Roadmap 2025*. [Online]. Available: https://cdn.euroncap.com/media/30700/euroncap-roadmap-2025-v4.pdf

[56] Audi. (2017). *Audi Elaine*. Accessed: May 5, 2022. [Online]. Available: https://www.audi.com/de/innovation/concept-cars/audi-elaine.html

[57] (2019). *Audi AI: ME*. Accessed: Mar. 6, 2022. [Online]. Available: https://aime.audi/ecosystem

[58] Mercedes-Benz. (2020). *VISION AVTR*. Accessed: May 6, 2022. [Online]. Available: https://www.mercedes-benz.com/en/vehicles/passenger-cars/mercedes-benz-concept-cars/vision-avtr/

[59] Toyota. (Jun. 2018). *Concept-I*. Accessed: May 5, 2022. [Online]. Available: https://newsroom.toyota.eu/2018-toyota-concept-i/

[60] (Oct. 2019). *LQ*. Accessed: May 10, 2022. [Online]. Available: https://global.toyota/en/newsroom/corporate/30063126.html

[61] KIA. (2019). *Real Time Emotion Adaptive Driving*. Accessed: May 5, 2022. [Online]. Available: https://www.kia.com/in/discover-kia/innovation/future-tech.html

[62] GM. (2019). *Cadillac SuperCruise*. Accessed: May 6, 2022. [Online]. Available: https://www.cadillac.com/ownership/vehicle-technology/super-cruise

[63] Ford. (2021). *Ford Bluecruise*. Accessed: May 6, 2022. [Online]. Available: https://www.ford.com/support/how-tos/ford-technology/driver-assist-features/what-is-ford-bluecruise-hands-free-driving/

[64] Volvo. (Mar. 2019). *Volvo Driver Monitoring*. Accessed: May 5, 2022. [Online]. Available: https://www.media.volvocars.com/global/en-gb/media/pressreleases/250015/

[65] Nio. (2017). *Nio's Nomi—World's First in-Vehicle Artificial Intelligence*. Accessed: May 3, 2022. [Online]. Available: https://www.electricvehiclesnews.com/TopNews/articles/nio-nomi-02122020.html

[66] BMW. (2020). *Driver Attention Camera*. Accessed: Jan. 11, 2023. [Online]. Available: https://www.youtube.com/watch?v=LKxgOXuhBUg

[67] (Jun. 2021). *The First Ever BMW IX*. Accessed: May 3, 2022. [Online]. Available: https://www.press.bmwgroup.com/global/article/attachment/T0333569EN/483906

[68] Mercedes-Benz. *ENERGIZING*. Accessed: May 9, 2022. [Online]. Available: https://www.mercedes-benz.de/passengercars/mercedes-benz-cars/energizing/energizing-comfort.module.html

[69] *Drive Pilot*. Accessed: May 22, 2023. [Online]. Available: https://www.mercedes-benz.de/passengercars/technology/drive-pilot.html

[70] *Mbux Interior Assistant*. Accessed: May 5, 2022. [Online]. Available: https://www.mercedes-benz.de/passengercars/mercedes-benz-cars/models/eqs/saloon-v297/specifications/digital-interior.module.html

[71] BMW. *BMW Driving Dynamics Control*. Accessed: Jan. 14, 2023. [Online]. Available: https://faq.bmwusa.com/s/article/BMW-3-Series-2022-Drive-modes-Driving-Dynamics-Control-saQp9

[72] J. Ling, J. Li, K. Tei, and S. Honiden, "Towards personalized autonomous driving: An emotion preference style adaptation framework," in *Proc. IEEE Int. Conf. Agents (ICA)*, Dec. 2021, pp. 47–52.

[73] A. Alsaid, J. D. Lee, S. I. Noejovich, and A. Chehade, "The effect of vehicle automation styles on drivers' emotional state," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 4, pp. 3963–3973, Apr. 2023.

[74] Z. Li et al., "Towards human-compatible autonomous car: A study of non-verbal Turing test in automated driving with affective transition modelling," *IEEE Trans. Affect. Comput.*, early access, May 23, 2023, doi: 10.1109/TAFFC.2023.3279311.

[75] J. K. Choi and Y. G. Ji, "Investigating the importance of trust on adopting an autonomous vehicle," *Int. J. Hum.-Comput. Interact.*, vol. 31, no. 10, pp. 692–702, Oct. 2015, doi: 10.1080/10447318.2015.1070549.

[76] J. Sini, A. C. Marceddu, M. Violante, and R. Dessì, "Passengers' emotions recognition to improve social acceptance of autonomous driving vehicles," in *Progresses in Artificial Intelligence and Neural Systems*, A. Esposito, M. Faundez-Zanuy, F. C. Morabito, and E. Pasero, Eds. Singapore: Springer, 2021, pp. 25–32, doi: 10.1007/978-981-15-5093-5_3.

[77] K. Saleh, M. Hossny, and S. Nahavandi, "Towards trusted autonomous vehicles from vulnerable road users perspective," in *Proc. Annu. IEEE Int. Syst. Conf. (SysCon)*, Apr. 2017, pp. 1–7.

[78] D. Dey, M. Martens, B. Eggen, and J. Terken, "The impact of vehicle appearance and vehicle behavior on pedestrian interaction with autonomous vehicles," in *Proc. 9th Int. Conf. Automot. User Interfaces Interact. Veh. Appl. Adjunct*, in AutomotiveUI. New York, NY, USA: Association for Computing Machinery, 2017, pp. 158–162, doi: 10.1145/3131726.3131750.

[79] Y. Wang, L. Hespanhol, and M. Tomitsch, "How can autonomous vehicles convey emotions to pedestrians? A review of emotionally expressive non-humanoid robots," *Multimodal Technol. Interact.*, vol. 5, no. 12, p. 84, Dec. 2021.

[80] A. Bera, T. Randhavane, A. Wang, D. Manocha, E. Kubin, and K. Gray, "Classifying group emotions for socially-aware autonomous vehicle navigation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 1152–11528.

[81] BMW. (Jan. 2022). *BMW MyModes*. Accessed: May 13, 2022. [Online]. Available: https://www.bmw.com/en/events/ces2022/mymodes.html

[82] S. C. Lee, H. Sanghavi, S. Ko, and M. Jeon, "Autonomous driving with an agent: Speech style and embodiment," in *Proc. 11th Int. Conf. Automot. User Interfaces Interact. Veh. Appl., Adjunct*. New York, NY, USA: Association for Computing Machinery, 2019, pp. 209–214.

[83] J. Koo, J. Kwac, W. Ju, M. Steinert, L. Leifer, and C. Nass, "Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance," *Int. J. Interact. Design Manuf. (IJIDeM)*, vol. 9, no. 4, pp. 269–275, Nov. 2015.

[84] M. Seet, J. Harvy, R. Bose, A. Dragomir, A. Bezerianos, and N. Thakor, "Differential impact of autonomous vehicle malfunctions on human trust," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 1, pp. 548–557, Jan. 2022.

[85] R. Zhou, Y. Ou, W. Tang, Q. Wang, and B. Yu, "An emergency evacuation behavior simulation method combines personality traits and emotion contagion," *IEEE Access*, vol. 8, pp. 66693–66706, 2020.

[86] M. Braun, J. Li, F. Weber, B. Pfleging, A. Butz, and F. Alt, "What if your car would care? Exploring use cases for affective automotive user interfaces," in *Proc. 22nd Int. Conf. Hum.-Comput. Interact. with Mobile Devices Services*. New York, NY, USA: Association for Computing Machinery, 2020, doi: 10.1145/3379503.3403530.

[87] H. Huang et al., "AffectRoute–considering people's affective responses to environments for enhancing route-planning services," *Int. J. Geographical Inf. Sci.*, vol. 28, no. 12, pp. 2456–2473, Dec. 2014, doi: 10.1080/13658816.2014.931585.

[88] K. H. Lim, J. Chan, C. Leckie, and S. Karunasekera, "Personalized trip recommendation for tourists based on user interests, points of interest visit durations and visit recency," *Knowl. Inf. Syst.*, vol. 54, no. 2, pp. 375–406, Feb. 2018, doi: 10.1007/s10115-017-1056-y.

[89] L. Chen, L. Zhang, S. Cao, Z. Wu, and J. Cao, "Personalized itinerary recommendation: Deep and collaborative learning with textual information," *Exp. Syst. Appl.*, vol. 144, Apr. 2020, Art. no. 113070. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417419307870

[90] E. A. Morris and E. Guerra, "Are we there yet? Trip duration and mood during travel," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 33, pp. 38–47, Aug. 2015.

[91] A.-E. Taha and N. AbuAli, "Route planning considerations for autonomous vehicles," *IEEE Commun. Mag.*, vol. 56, no. 10, pp. 78–84, Oct. 2018.

[92] S. M. Sarala, D. H. Sharath Yadav, and A. Ansari, "Emotionally adaptive driver voice alert system for advanced driver assistance system (ADAS) applications," in *Proc. Int. Conf. Smart Syst. Inventive Technol. (ICSSIT)*, Dec. 2018, pp. 509–512.

[93] M. Braun, J. Schubert, B. Pfleging, and F. Alt, "Improving driver emotions with affective strategies," *Multimodal Technol. Interact.*, vol. 3, no. 1, p. 21, Mar. 2019. [Online]. Available: https://www.mdpi.com/2414-4088/3/1/21

[94] P. Kong, H. Cornet, and F. Frenkler, "Personas and emotional design for public service robots: A case study with autonomous vehicles in public transportation," in *Proc. Int. Conf. Cyberworlds (CW)*, Oct. 2018, pp. 284–287.

[95] Mercedes-Benz. (Dec. 2021). *Mbux Hyperscreen*. Accessed: May 11, 2022. [Online]. Available: https://group.mercedes-benz.com/magazine/technology-innovation/hyperscreen.html

[96] BMW. (Mar. 2021). *BMW iDrive*. Accessed: May 6, 2022. [Online]. Available: https://www.bmwgroup.com/en/news/general/2021/new_bmwidrive.html

[97] *Theatre Screen*. Accessed: May 9, 2022. [Online]. Available: https://www.bmw.com/en/events/ces2022/theaterscreen.html

[98] D. Ayata, Y. Yaslan, and M. E. Kamasak, "Emotion based music recommendation system using wearable physiological sensors," *IEEE Trans. Consum. Electron.*, vol. 64, no. 2, pp. 196–203, May 2018.

[99] A. Nair, S. Pillai, and G. S. Nair, "Emotion based music playlist recommendation system using interactive chatbot," in *Proc. 6th Int. Conf. Commun. Electron. Syst. (ICCES)*, Jul. 2021, pp. 1767–1772.

[100] L. Mou et al., "Driver stress detection via multimodal fusion using attention-based CNN-LSTM," *Exp. Syst. Appl.*, vol. 173, Jul. 2021, Art. no. 114693. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417421001342

[101] H. Yang, J. Wu, Z. Hu, and C. Lv, "Real-time driver cognitive workload recognition: Attention-enabled learning with multimodal information fusion," *IEEE Trans. Ind. Electron.*, early access, Jun. 26, 2023, doi: 10.1109/TIE.2023.3288182.

[102] P. Wintersberger, A. Riener, and A.-K. Frison, "Automated driving system, male, or female driver: Who'd you prefer? Comparative analysis of passengers' mental conditions, emotional states & qualitative feedback," in *Proc. 8th Int. Conf. Automot. User Interfaces Interact. Veh. Appl.* New York, NY, USA: Association for Computing Machinery, 2016, pp. 51–58, doi: 10.1145/3003715.3005410.

[103] A. K. Sriranga, Q. Lu, and S. Birrell, "A systematic review of in-vehicle physiological indices and sensor technology for driver mental workload monitoring," *Sensors*, vol. 23, no. 4, p. 2214, Feb. 2023.

[104] K. Lu, A. Sjörs Dahlman, J. Karlsson, and S. Candefjord, "Detecting driver fatigue using heart rate variability: A systematic review," *Accident Anal. Prevention*, vol. 178, Dec. 2022, Art. no. 106830. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0001457522002652

[105] A. Lanatà et al., "How the autonomic nervous system and driving style change with incremental stressing conditions during simulated driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 3, pp. 1505–1517, Jun. 2015.

[106] Z. Guo, Y. Pan, G. Zhao, S. Cao, and J. Zhang, "Detection of driver vigilance level using EEG signals and driving contexts," *IEEE Trans. Rel.*, vol. 67, no. 1, pp. 370–380, Mar. 2018.

[107] M. M. Van Miltenburg, D. J. Lemmers, A. Tinga, M. Christoph, and R. Zon, "Can EEG measurements be used to estimate the performance of taking over control from an autonomous vehicle for different levels of distracted driving? An explorative study," in *Proc. Adjunct 14th Int. Conf. Automot. User Interfaces Interact. Veh. Appl.* New York, NY, USA: Association for Computing Machinery, 2022, pp. 20–24, doi: 10.1145/3544999.3552324.

[108] V. Selvaraju et al., "Continuous monitoring of vital signs using cameras: A systematic review," *Sensors*, vol. 22, no. 11, p. 4097, May 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/11/4097

[109] D. McDuff, "Camera measurement of physiological vital signs," *ACM Comput. Surv.*, vol. 55, no. 9, p. 40, Jan. 2023, Art. no. 176, doi: 10.1145/3558518.

[110] BMW. *Connected Drive*. Accessed: May 9, 2022. [Online]. Available: https://www.bmw-me.com/en/topics/fascination-bmw/connected-drive/bmw-connected-drive-overview.html#

[111] D. Ouyang, G. Li, Q. Li, X. Sui, X. Qu, and G. Guo, "Effect of music intervention strategies on mitigating drivers' negative emotion in post-congestion driving," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2023, pp. 1–6.

[112] P. E. Paredes et al., "Just breathe: In-car interventions for guided slow breathing," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 2, no. 1, pp. 1–23, Mar. 2018, doi: 10.1145/3191760.

[113] P. E. Paredes et al., "Driving with the fishes: Towards calming and mindful virtual reality experiences for the car," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 2, no. 4, pp. 1–21, Dec. 2018, doi: 10.1145/3287062.

[114] R. Kosti, J. M. Alvarez, A. Recasens, and A. Lapedriza, "Emotion recognition in context," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1960–1968.

[115] R. Kosti, J. M. Alvarez, A. Recasens, and A. Lapedriza, "Context based emotion recognition using EMOTIC dataset," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 11, pp. 2755–2766, Nov. 2020.

[116] J. Lee, S. Kim, S. Kim, J. Park, and K. Sohn, "Context-aware emotion recognition networks," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 10142–10151.

[117] C. Ringer, J. A. Walker, and M. A. Nicolaou, "Multimodal joint emotion and game context recognition in league of legends livestreams," in *Proc. IEEE Conf. Games (CoG)*, London, U.K., Aug. 2019, pp. 1–8, doi: 10.1109/CIG.2019.8848060.

[118] M. Braun, B. Pfleging, and F. Alt, "A survey to understand emotional situations on the road and what they mean for affective automotive UIs," *Multimodal Technol. Interact.*, vol. 2, no. 4, p. 75, Oct. 2018.

[119] N. McLaughlin, J. Martinez del Rincon, and P. Miller, "Recurrent convolutional network for video-based person re-identification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1325–1334.

[120] Z. Wang et al., "Zero-shot person re-identification via cross-view consistency," *IEEE Trans. Multimedia*, vol. 18, no. 2, pp. 260–272, Feb. 2016.

[121] S. Li, S. Bak, P. Carr, and X. Wang, "Diversity regularized spatiotemporal attention for video-based person re-identification," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 369–378.

[122] T. J. Park, N. Kanda, D. Dimitriadis, K. J. Han, S. Watanabe, and S. Narayanan, "A review of speaker diarization: Recent advances with deep learning," *Comput. Speech Lang.*, vol. 72, Mar. 2022, Art. no. 101317.

[123] B. I. Ahmad, P. M. Langdon, J. Liang, S. J. Godsill, M. Delgado, and T. Popham, "Driver and passenger identification from smartphone data," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 4, pp. 1278–1288, Apr. 2019.

[124] W. Li et al., "Global-local-feature-fused driver speech emotion detection for intelligent cockpit in automated driving," *IEEE Trans. Intell. Vehicles*, vol. 8, no. 4, pp. 2684–2697, Apr. 2023.

[125] L. Tan et al., "Speech emotion recognition enhanced traffic efficiency solution for autonomous vehicles in a 5G-enabled space–air–ground integrated intelligent transportation system," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 3, pp. 2830–2842, Mar. 2022.

[126] A. Dhall, R. Goecke, J. Joshi, J. Hoey, and T. Gedeon, "EmotiW 2016: Video and group-level emotion recognition challenges," in *Proc. 18th ACM Int. Conf. Multimodal Interact.* New York, NY, USA: ACM, 2016, pp. 427–432, doi: 10.1145/2993148.2997638.

[127] E. A. Veltmeijer, C. Gerritsen, and K. V. Hindriks, "Automatic emotion recognition for groups: A review," *IEEE Trans. Affect. Comput.*, vol. 14, no. 1, pp. 89–107, Jan. 2023.

[128] L. Tan, K. Zhang, K. Wang, X. Zeng, X. Peng, and Y. Qiao, "Group emotion recognition with individual facial emotion CNNs and global image based CNNs," in *Proc. 19th ACM Int. Conf. Multimodal Interact.*, E.Lank, E. Hoggan, S. Subramanian, A. Vinciarelli, and S. Brewster, Eds. New York, NY, USA: Association for Computing Machinery, Jan. 2017, pp. 549–552.

[129] X. Guo, B. Zhu, L. F. Polanía, C. Boncelet, and K. E. Barner, "Group-level emotion recognition using hybrid deep models based on faces, scenes, skeletons and visual attentions," in *Proc. 20th ACM Int. Conf. Multimodal Interact.* New York, NY, USA: Association for Computing Machinery, Oct. 2018, pp. 635–639.

[130] K. Wang et al., "Cascade attention networks for group emotion recognition with face, body and image cues," in *Proc. 20th ACM Int. Conf. Multimodal Interact.* New York, NY, USA: Association for Computing Machinery, 2018, pp. 640–645, doi: 10.1145/3242969.3264991.

[131] N. Alyuz, S. Aslan, J. Healey, I. J. Alvarez, and A. A. Esme, "Towards understanding emotional reactions of driver-passenger dyads in automated driving," in *Proc. 13th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, May 2018, pp. 585–592.

[132] Z.-S. Chen et al., "Identifying and prioritizing factors affecting in-cabin passenger comfort on high-speed rail in China: A fuzzy-based linguistic approach," *Appl. Soft Comput.*, vol. 95, Oct. 2020, Art. no. 106558.

[133] B. R. Kiran et al., "Deep reinforcement learning for autonomous driving: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 6, pp. 4909–4926, Jun. 2022.

[134] S. Aradi, "Survey of deep reinforcement learning for motion planning of autonomous vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 2, pp. 740–759, Feb. 2022.

[135] M. Nazari, A. Oroojlooy, L. Snyder, and M. Takac, "Reinforcement learning for solving the vehicle routing problem," in *Advances in Neural Information Processing Systems*, vol. 31, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds. Red Hook, NY, USA: Curran Associates, 2018. [Online]. Available: https://proceedings.neurips.cc/paper/2018/file/9fb4651c05b2ed70fba5afe0b039a550-Paper.pdf

[136] X. He, B. Lou, H. Yang, and C. Lv, "Robust decision making for autonomous vehicles at highway on-ramps: A constrained adversarial reinforcement learning approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 4, pp. 4103–4113, Apr. 2023.

[137] Y. Shan, B. Zheng, L. Chen, L. Chen, and D. Chen, "A reinforcement learning-based adaptive path tracking approach for autonomous driving," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 10581–10595, Oct. 2020.

[138] J. Wu, Z. Huang, W. Huang, and C. Lv, "Prioritized experience-based reinforcement learning with human guidance for autonomous driving," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jun. 10, 2022, doi: 10.1109/TNNLS.2022.3177685.

[139] J. Li, L. Yao, X. Xu, B. Cheng, and J. Ren, "Deep reinforcement learning for pedestrian collision avoidance and human-machine cooperative driving," *Inf. Sci.*, vol. 532, pp. 110–124, Sep. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0020025520302851

[140] A. Ferdowsi, U. Challita, W. Saad, and N. B. Mandayam, "Robust deep reinforcement learning for security and safety in autonomous vehicle systems," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 307–312.

[141] T. T. Nguyen and V. J. Reddi, "Deep reinforcement learning for cyber security," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 8, pp. 3779–3795, Nov. 2023.

[142] O. Rudovic, M. Zhang, B. Schuller, and R. Picard, "Multi-modal active learning using reinforcement learning," in *Proc. 21st ACM Int. Conf. Multimodal Interact.* Suzhou, China: ACM, Oct. 2019, p. 10.

[143] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol.*, vol. 1. Minneapolis, MN, USA: Association for Computational Linguistics, 2019, pp. 4171–4186. [Online]. Available: https://aclanthology.org/N19-1423

[144] F. A. Acheampong, H. Nunoo-Mensah, and W. Chen, "Transformer models for text-based emotion detection: A review of BERT-based approaches," *Artif. Intell. Rev.*, vol. 54, no. 8, pp. 5789–5829, Dec. 2021, doi: 10.1007/s10462-021-09958-2.

[145] C. Luo, X. Yang, and A. Yuille, "Self-supervised pillar motion learning for autonomous driving," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 3182–3191.

[146] F. Chiaroni, M.-C. Rahal, N. Hueber, and F. Dufaux, "Self-supervised learning for autonomous vehicles perception: A conciliation between analytical and learning methods," *IEEE Signal Process. Mag.*, vol. 38, no. 1, pp. 31–41, Jan. 2021.

[147] M. Macary, M. Tahon, Y. Estève, and A. Rousseau, "On the use of self-supervised pre-trained acoustic and linguistic features for continuous speech emotion recognition," in *Proc. IEEE Spoken Lang. Technol. Workshop (SLT)*, Jan. 2021, pp. 373–380.

[148] A. Shukla, S. Petridis, and M. Pantic, "Does visual self-supervision improve learning of speech representations for emotion recognition?" *IEEE Trans. Affect. Comput.*, vol. 14, no. 1, pp. 406–420, Jan. 2023.

[149] Z. Zhang, F. Ringeval, B. Dong, E. Coutinho, E. Marchi, and B. Schüller, "Enhanced semi-supervised learning for multimodal emotion recognition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2016, pp. 5185–5189.

[150] S. Latif, R. Rana, S. Khalifa, R. Jurdak, and B. W. Schuller, "Self supervised adversarial domain adaptation for cross-corpus and cross-language speech emotion recognition," *IEEE Trans. Affect. Comput.*, vol. 14, no. 3, pp. 1912–1926, Sep. 2022.

[151] S. Karnouskos and F. Kerschbaum, "Privacy and integrity considerations in hyperconnected autonomous vehicles," *Proc. IEEE*, vol. 106, no. 1, pp. 160–170, Jan. 2018.

[152] H. Lim and A. Taeihagh, "Autonomous vehicles for smart and sustainable cities: An in-depth exploration of privacy and cybersecurity implications," *Energies*, vol. 11, no. 5, p. 1062, Apr. 2018.

[153] A. Waytz, J. Heafner, and N. Epley, "The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle," *J. Experim. Social Psychol.*, vol. 52, pp. 113–117, May 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022103114000067

[154] D. Niu, J. Terken, and B. Eggen, "Anthropomorphizing information to enhance trust in autonomous vehicles," *Human Factors Ergonom. Manuf. Service Industries*, vol. 28, no. 6, pp. 352–359, Nov. 2018, doi: 10.1002/hfm.20745.

[155] Z. Xu, K. Zhang, H. Min, Z. Wang, X. Zhao, and P. Liu, "What drives people to accept automated vehicles? Findings from a field experiment," *Transp. Res. C, Emerg. Technol.*, vol. 95, pp. 320–334, Oct. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X18302316

[156] G. Wiegand, M. Eiband, M. Haubelt, and H. Hussmann, "'I'd like an explanation for that! 'Exploring reactions to unexpected autonomous driving," in *Proc. 22nd Int. Conf. Hum.-Comput. Interact. Mobile Devices Services*. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–11, doi: 10.1145/3379503.3403554.

[157] R. Zimmermann and R. Wettach, "First step into visceral interaction with autonomous vehicles," in *Proc. 9th Int. Conf. Automot. User Interfaces Interact. Veh. Appl.* New York, NY, USA: Association for Computing Machinery, Sep. 2017, pp. 58–64, doi: 10.1145/3122986.3122988.

[158] S. Deb, M. M. Rahman, L. J. Strawderman, and T. M. Garrison, "Pedestrians' receptivity toward fully automated vehicles: Research review and roadmap for future research," *IEEE Trans. Hum.-Mach. Syst.*, vol. 48, no. 3, pp. 279–290, Jun. 2018.

[159] A. Rasouli and J. K. Tsotsos, "Autonomous vehicles that interact with pedestrians: A survey of theory and practice," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 3, pp. 900–918, Mar. 2020.

[160] M. Braun, A. Mainz, R. Chadowitz, B. Pfleging, and F. Alt, "At your service: Designing voice assistant personalities to improve automotive user interfaces," in *Proc. CHI Conf. Hum. Factors Comput. Syst.* New York, NY, USA: Association for Computing Machinery, 2019, pp. 1–11, doi: 10.1145/3290605.3300270.

[161] B. S. Alpers et al., "Capturing passenger experience in a ride-sharing autonomous vehicle: The role of digital assistants in user interface design," in *Proc. 12th Int. Conf. Automot. User Interfaces Interact. Veh. Appl.* New York, NY, USA: Association for Computing Machinery, 2020, pp. 83–93, doi: 10.1145/3409120.3410639.

[162] M. Wang, S. C. Lee, H. Kamalesh Sanghavi, M. Eskew, B. Zhou, and M. Jeon, "In-vehicle intelligent agents in fully autonomous driving: The effects of speech style and embodiment together and separately," in *Proc. 13th Int. Conf. Automot. User Interfaces Interact. Veh. Appl.* New York, NY, USA: Association for Computing Machinery, Sep. 2021, pp. 247–254, doi: 10.1145/3409118.3475142.

[163] I. Krontiris et al., "Autonomous vehicles: Data protection and ethical considerations," in *Proc. Comput. Sci. Cars Symp.* New York, NY, USA: Association for Computing Machinery, Dec. 2020, doi: 10.1145/3385958.3430481.

[164] K. Wei et al., "Federated learning with differential privacy: Algorithms and performance analysis," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 3454–3469, 2020.

[165] A. Mishra, J. Cha, and S. Kim, "HCI based in-cabin monitoring system for irregular situations with occupants facial anonymization," in *Intelligent Human Computer Interaction: 12th International Conference*, Daegu, South Korea, Cham, Switzerland: Springer, Nov. 2020, pp. 380–390.

[166] A. Mishra, J. Cha, and S. Kim, "Privacy-preserved in-cabin monitoring system for autonomous vehicles," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–15, Apr. 2022, doi: 10.1155/2022/5389359.

[167] F. Nasoz, C. L. Lisetti, and A. V. Vasilakos, "Affectively intelligent and adaptive car interfaces," *Inf. Sci.*, vol. 180, no. 20, pp. 3817–3836, Oct. 2010.

[168] A. Löcken, K. Ihme, and A. Unni, "Towards designing affect-aware systems for mitigating the effects of in-vehicle frustration," in *Proc. 9th Int. Conf. Automot. User Interfaces Interact. Veh. Appl. Adjunct*. New York, NY, USA: Association for Computing Machinery, Sep. 2017, pp. 88–93, doi: 10.1145/3131726.3131744.

[169] W. Wang, R. Li, Z. M. Diekel, and Y. Jia, "Intuitive maneuver of autonomous vehicles without physical control interfaces using wearable sensing devices," in *Proc. IEEE 7th Annu. Int. Conf. CYBER Technol. Autom., Control, Intell. Syst. (CYBER)*, Jul. 2017, pp. 1451–1456.

[170] H. Culbertson, S. B. Schorr, and A. M. Okamura, "Haptics: The present and future of artificial touch sensation," *Annu. Rev. Control, Robot., Auto. Syst.*, vol. 1, no. 1, pp. 385–409, May 2018, doi: 10.1146/annurev-control-060117-105043.

[171] J. Li, M. Braun, A. Butz, and F. Alt, "Designing emotion-aware in-car interactions for unlike markets," in *Proc. 11th Int. Conf. Automot. User Interfaces Interact. Veh. Appl., Adjunct*. New York, NY, USA: Association for Computing Machinery, 2019, pp. 352–357, doi: 10.1145/3349263.3351324.

[172] R. Fernandez-Rojas et al., "Contextual awareness in human-advanced-vehicle systems: A survey," *IEEE Access*, vol. 7, pp. 33304–33328, 2019.

[173] S. Mathews, "Explainable artificial intelligence applications in NLP, biomedical, and malware classification: A literature review," in *Proc. Intell. Comput. Comput. Conf.*, in Advances in Intelligent Systems and Computing, vol. 998, 2019, pp. 1269–1292.

[174] A. Baird, S. Hantke, and B. Schuller, "Responsible and representative multimodal data acquisition and analysis: On auditability, benchmarking, confidence, data-reliance & explainability," 2019, *arXiv:1903.07171*.

[175] A. Adadi and M. Berrada, "Peeking inside the black-box: A survey on explainable artificial intelligence (XAI)," *IEEE Access*, vol. 6, pp. 52138–52160, 2018.

[176] A. Barredo Arrieta et al., "Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Inf. Fusion*, vol. 58, pp. 82–115, Jun. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1566253519308103

[177] B. W. Schuller, "Speech emotion recognition: Two decades in a nutshell, benchmarks, and ongoing trends," *Commun. ACM*, vol. 61, no. 5, pp. 90–99, Apr. 2018.

[178] Y. Nirkin, Y. Keller, and T. Hassner, "FSGAN: Subject agnostic face swapping and reenactment," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 7183–7192.

[179] L. F. Barrett, R. Adolphs, S. Marsella, A. M. Martinez, and S. D. Pollak, "Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements," *Psychol. Sci. Public Interest*, vol. 20, no. 1, pp. 1–68, Jul. 2019, doi: 10.1177/1529100619832930.

[180] K. K. Kim, Y. Suh, and S. H. Ko, "Smart stretchable electronics for advanced human–machine interface," *Adv. Intell. Syst.*, vol. 3, no. 2, Feb. 2021, Art. no. 2000157.

[181] M. A. Eid and H. Al Osman, "Affective haptics: Current research and future directions," *IEEE Access*, vol. 4, pp. 26–40, 2016.

[182] C. Chen, C. Wang, B. Liu, C. He, L. Cong, and S. Wan, "Edge intelligence empowered vehicle detection and image segmentation for autonomous vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 11, pp. 13023–13034, Jan. 2023.

[183] R. Bi, J. Xiong, Y. Tian, Q. Li, and K. R. Choo, "Achieving lightweight and privacy-preserving object detection for connected autonomous vehicles," *IEEE Internet Things J.*, vol. 10, no. 3, pp. 2314–2329, Feb. 2023.

[184] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan, "A survey on bias and fairness in machine learning," *ACM Comput. Surveys*, vol. 54, no. 6, pp. 1–35, Jul. 2021, doi: 10.1145/3457607.

Vincent Karas received the bachelor's degree in engineering and the master's degree in biomedical engineering from the Technical University of Munich in 2014 and 2018, respectively. He is currently pursuing the Ph.D. degree with the Chair of Embedded Intelligence for Health Care and Wellbeing, University of Augsburg. Since 2018, he has been with the BMW Group, where he joined the doctoral candidate researchers program and submitted a patent for an in-vehicle camera system. In 2019, he was a Researcher with the Chair of Embedded Intelligence for Health Care and Wellbeing, University of Augsburg. His research interests include affective computing in the wild and digital health and wellbeing. He is a reviewer of IEEE TRANSACTIONS ON AFFECTIVE COMPUTING and IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING.

Dagmar M. Schuller received the master's degree in economic sciences from the Vienna University of Economics and Business. She is the Co-Founder and CEO of audEERING GmbH, an audio AI company, Gilching and Berlin, Germany, that largely deals with voice-based analysis of health and emotions. Her research interests include speech emotion recognition and general audio intelligence and affective computing. She is also the Vice President of the Chamber of Industry and Commerce for Munich and Upper Bavaria, Germany, and the Deputy Head Bavaria of the German AI Association.

Björn W. Schuller (Fellow, IEEE) received the Ph.D. and Habilitation degrees from TUM, Germany. He is currently a Professor of artificial intelligence and the Head of GLAM, Imperial College London, U.K. He is a Full Professor with the Chair of Embedded Intelligence for Health Care and Wellbeing, University of Augsburg, Germany, and the Co-Founding CEO and the CSO of audEERING GmbH, providing speech emotion recognition technology. He is also the Independent Research Leader of The Alan Turing Institute, Royal Statistical Society Laboratory, U.K. Health Security Agency, leading research on COVID-19. He coauthored five books and more than 1 200 publications. He is a fellow of AAAC, BCS, ISCA, and ELLIS. He is a Senior Member of ACM. He was a recipient of the Golden Core Award of the IEEE Computer Society. He is the Field Chief Editor of *Frontiers in Digital Health* and the Former Editor-in-Chief of IEEE TRANSACTIONS ON AFFECTIVE COMPUTING. He is the President-Emeritus AAAC. He frequently consults, including Barclays, GN, Huawei, Informetis, and Samsung.