A Multimodal Approach to Understand Driver's Distraction for DMS

A Multimodal Approach to Understand Driver's Distraction For DMS

Andrea Generosi[1[0000-0001-6173-9616]], Josè Yuri Villafan[1[0000-0001-6475-2533]], Roberto Montanari[2], and Maura Mengoni[1[0000-0003-2826-7455]]

[1] Department of Industrial Engineering and Mathematical Sciences – Università
[2] Re-lab srl, Italy
Politecnica delle Marche, Italy
a.generosi@univpm.it

Abstract. This study introduces a multimodal approach for enhancing the accuracy of Driver Monitoring Systems (DMS) in detecting driver distraction. By integrating data from vehicle control units with vision-based information, the research aims to address the limitations of current DMS. The experimental setup involves a driving simulator and advanced computer vision, deep learning technologies for facial expression recognition, and head rotation analysis. The findings suggest that combining various data types—behavioral, physiological, and emotional—can significantly improve DMS's predictive capability. This research contributes to the development of more sophisticated, adaptive, and real-time systems for improving driver safety and advancing autonomous driving technologies.

Abstract. Keywords: Driver Monitoring Systems, Deep Learning, Attention Recognition.

1 Introduction

Advanced Driver-Assistance Systems (ADAS) have been thriving and widely deployed in recent years. They are human-machine systems that assist drivers in driving and parking functions, providing partial automation applications like adaptive cruise control, lane departure warning, pedestrian detection, and avoidance, etc. to improve a safe and effortless driving experience [1]. They receive sensor data, compute driving decisions and output control signals to the vehicles. Some open issues in ADAS regard their reliability to reduce malfunctions and critical consequences, precise and real-time measurements, multi-sensor data fusion to avoid that incorrect fusion logic can lead the vehicle to critical accidents.

One of the most challenging applications of ADAS is the Driver Distraction and Drowsiness Recognition (DDR), which warns the driver of drowsiness or other distractions. It is crucial that these systems begin to be incorporated into mass production vehicles, as demonstrated by studies such as [2] and [3].

Real-time Driver Monitoring Systems (DMS) represent a solution to address key behavioural risks as they occur, particularly distraction and fatigue. Their integration in proactive ADAS opens new frontiers in vehicular safety and autonomous driving technologies. DMS and ADAS communication allow multi-sensor data to be fused and processed to assess the level of driver's attention, mental workload and fatigue and then to ensure possible hand-over/hand-back actions to achieve full autonomous and safe driving [4]. These systems commonly aim to offer adaptive Human Machine Interfaces (HMI) to alert the driver to potential risks and take immediate action to prevent a collision based on the level of risk associated with the driver's state [5].

The huge amount of data to be processed by ADAS and DMS includes parameters like the steering wheel movement, the brake pressure, the acceleration patterns, and high-dimensional data extracted by computer vision from images and videos captured by embedded cameras in the cabin, i.e. eye movement, facial expressions, driver's posture etc. The first data allows the DMS to assess the driver's physical interaction with the vehicle, while the second to understand his/her cognitive and emotional state.

This paper examines the correlations between various types of data that can be extracted from an automated DMS and vehicle dynamic data. The aim is to propose a multimodal approach to enhance the reliability of driver recognition of distraction states while driving for ADAS systems.

This research starts with a review of the state-of-art multimodal approaches for multi-sensor data fusion to enhance the assessment of the driver's state, particularly concerning distraction. Most studies explored using signals originating from the vehicle's electronic control unit and various types of sensors, including biometric ones, to analyse the existing correlation between the physical and cognitive aspects of the driver's state. However, previous studies have not assessed the correlations between driving signals extracted from the control unit and behavioural signals processed by RGB-based computer vision software, including facial expression analysis. Despite a similar approach being proposed in [6], where an adaptive Human-Machine Interface capable of responding to user-expressed emotions to enhance driving safety has been introduced, this gap in the literature still needs to be addressed.

This preliminary study results in an overview of current approaches and investigates how to merge physical, cognitive and emotional aspects to have a complete understanding of the driver's status. Testing involves a driving simulator with advanced computer vision and Deep Learning technologies, capturing detailed control unit data and employing state-of-the-art algorithms for facial expression recognition, head rotation analysis. By statistically correlating data from the control unit and computer vision, a better understanding of the driver's state can be achieved. This can lead to advancements in automotive safety systems and autonomous vehicle technology. The main contributions of the paper are as follows:

- A comprehensive analysis of the current state-of-the-art, focusing on the critical parameters and processing techniques essential for the development of effective DMS.
- The introduction of a multimodal approach for data processing. This approach integrates data from the vehicle's control unit and vision-based information, thereby enhancing the accuracy and reliability of the DMS.

- The research also involves the compilation of a substantial dataset, gathered using a driving simulator. This dataset is used to test the potential of Deep Neural Networks (DNN) algorithms with control unit data in the context of distraction recognition.

2 Literature Review

1.1 Driving Monitoring Systems

The evolution of ADAS marks a significant leap in automotive technology, underpinned by the integration of advanced sensors and artificial intelligence. The current trend emphasises the essential role of in-vehicle DMS as a core component of efficient ADAS platforms. However, the journey towards creating an effective and reliable DMS involves numerous challenges and limitations, which are critical to address for the advancement of this field.

The current scientific literature on DMS underscores the critical importance of a multimodal approach to effectively assess driver states. As driving is fundamentally a visual task, the peril posed by visual distractions to safe driving performance is particularly highlighted. Instances of frequent lane deviations, abrupt steering movements, and slow reaction times to braking events underscore the significant risks associated with visual distractions. The European Commission in 2015 emphasised that "activities that cause visual distraction seem to be the most dangerous" [7], underscoring the necessity for systems capable of assessing the driver's visual attention state determining when the driver is visually distracted.

Despite advancements, the state of the art indicates a gap in comprehensive solutions, even among systems utilising advanced artificial intelligence, to accurately evaluate these parameters. Challenges such as occlusions that prevent eye-tracking and the limitations of using only facial behaviour metrics (head movement and/or face orientation) for approximate gaze direction assessment are notable [8].

However, the feasibility of countering distraction through innovative driver analysis systems is supported by research. Direct and non-intrusive psychophysiological monitoring of drivers, such as eye closure, could significantly improve the detection of distraction and drowsiness [9].

At the same time, it is necessary to go beyond the simple use of these systems based on cameras and artificial intelligence software (mostly based on computer vision and deep learning), as is currently the case for a large part of the products integrated by car manufacturers in their cars [10], and to develop multimodal approaches based on different types of data and sensors.

The pioneering work of Daza et al. [11] set a benchmark by introducing non-intrusive methods to monitor driver drowsiness through a mix of physical and driving performance indicators, albeit within the confines of simulated conditions. Sandberg et al. [12] expanded the range of indicators to 35, utilising stochastic optimization algorithms to enhance detection, yet facing the challenge of processing large datasets in real-time.

Zhang et al. [13] provided insights into the correlation between vehicle dynamics and driver fatigue, emphasising the challenge of individual variability, which calls for personalised or adaptable DMS.

In the continued effort to refine DMS [9], introduced a multimodal analysis system that adeptly integrated physiological, behavioural, and vehicle signals. This system's high accuracy rates underscored the efficacy of a holistic approach, considering the full spectrum of a driver's condition, and set a benchmark for future systems.

[14] took a practical approach, analysing one vehicle with different drivers and routes, with the system implemented on a mobile application. This study highlighted the critical role of context variability in DMS and embraced the trend of harnessing mobile technology to broaden the accessibility and integration of DMS applications.

However, a challenge identified in [15] was the lack of a freely available, comprehensive database containing complex driver and vehicle data, including stress or fatigue annotations. This limitation in data availability is a hurdle that continues to impede the development of universally applicable and effective DMS.

1.2 Drowsiness and Distraction parameters

One aim of this study is to investigate the correlation between different signals and their effectiveness in the future development of DMS. After conducting extensive research on the current state of the art, a wide range of parameters has been evaluated to improve the accuracy of detecting drivers' attention states. A key finding from this research highlights the importance of integrating multiple parameters, as singular metrics are insufficient in accurately capturing the complex nature of driver distraction. Consequently, studies are increasingly examining a wider range of parameters, both within the same category (such as physiological, video, vehicle dynamics, or ocular dynamics) and across different categories. These parameters are primarily used to identify states of distraction and/or fatigue, which often overlap. The table categorises each parameter reflecting the nuanced approach required to accurately monitor driver attention and safety.

Table 1. State-of-the-art analysed parameters for distraction detection.

CATEGORY	VARIABLE	DESCRIPTION	REFERENCE
Physiological	Electroencephalogram measures (EEG)	There is good evidence that rising alpha (8–11Hz) and theta (4–7 Hz) EEG activities indicate increasing sleepiness and thus the potential for lapses in attention and behaviour	[16] [17]
Physiological	Heart rate variability (HRV)	Elevated HRV can indicate increased mental workload or stress, while a lower HRV is often associated with fatigue or drowsiness.	[18] [17]

Physiological	Electrodermal activity (EDA)	Refers to the variation of the skin’s electrical conductance in response to sweat secretion. Skin conductance can be a measure of emotional response.	[19]
Physiological	Skin Temperature	Can signify a driver's stress or relaxed state, impacting their attention and alertness	[19]
Ocular dynamics	Visual fixation	Used to indicate a loss of focus on the road. Alert drivers frequently shift their gaze, actively engaging with their surroundings.	[20][21]
Ocular dynamics	Pupil diameter	Changes in pupil size can indicate changes in alertness, helping to detect moments of distraction or drowsiness.	[20]
Ocular dynamics	Eye glance position	Prolonged or frequent glances away from the road, as indicated by eye glance position, suggest the driver is not paying attention to the driving task.	[21]
Vehicle behaviour	Steering wheel variability	Variability is greater as drivers become more distract	[22] [17]
Vehicle behaviour	Steering wheel reversal rate (SWRR)	Variability in SWRR indicates a change in steering behaviour that is symptomatic of both distraction and fatigue	[23]
Vehicle behaviour	Standard deviation of lane position (SDLP)	Lane-tracking variability was observed to be related to the amount of distraction and drowsiness in drivers.	[24] [17]
Vehicle behaviour	Steering wheel rapid movement (SWRM)	High variability in SWRM can indicate a lack of smooth steering control, often due to a driver's divided attention or the onset of fatigue, leading to more frequent and erratic steering corrections.	[22] [17]
Vehicle behaviour	Mean square of the lane deviation	It is considered to be an accurate and reliable measure for the detection of distraction.	[22] [17]

Vehicle behaviour	Time to line crossing (TLC)	Used to predict driver impairment due to drowsiness and distraction, representing how much time before the vehicle drifts out of the lane.	[25] [17]
Vehicle behaviour	Lane keeping offset	When the driver performs a secondary task, her/his lane keeping ability degrades. A road segment is described by speed limit and road curvature	[26][23]
Vehicle behaviour	Standard deviation of speed (SDS)	When the driver performs secondary task, her/his speed maintenance ability degrade	[26]
Vehicle behaviour	Brake pedal angular position	The physical response of braking can be correlated with attention levels while performing various driving manoeuvres	[19]
Vehicle behaviour	Gas pedal angular position	Changes in gas pedal position, such as sudden acceleration, can reflect the driver's response to external stimuli or a lapse in attention.	[19]
Video based	Observation of body motions	Driver alertness decreases over time; initially, drivers are more active in checking mirrors and the road. Over an hour, these behaviours reduce.	[27] [17]
Video based	Face position	Determining face pose involves analysing the orientation and position of the face, and head movements. Normal face orientation while driving should be forward-facing, deviations suggest distraction.	[20]
Video based	Drivers' Interaction with Car Interior	Quantifying the number, the type, and the conditions of events and interactions between car and driver	[19]
Video based	Emotions detection	Facial expression analysis in drivers assesses their emotional	[28][29][30]

	state, which can significantly impact driving behaviour and potentially lead to accidents.

Physiological parameter. Detecting distraction in drivers relies on a set of physiological parameters that provide insights into their cognitive and physical states.

These indicators include variations in heart rate, which can escalate during periods of distraction [17][18]. Skin conductance gauges changes in the electrical conductance of the skin, reflecting alterations in the autonomic nervous system that occur during distraction [19].

In addition, the monitoring of electroencephalography (EEG) signals can provide valuable information about brain activity. Sudden shifts in beta waves, for example, can suggest drowsiness and distraction [16]. Moreover, alterations in skin temperature can indicate changes in the autonomic nervous system, revealing physiological responses to distraction or drowsiness. The temperature fluctuates based on blood flow and sympathetic nervous system activity. Increased distraction can result in heightened sympathetic nervous system activity and vasoconstriction, causing a drop in skin temperature [19].

Ocular dynamics parameters. Parameters related to ocular dynamics can provide a more comprehensive understanding of the driver's alertness and focus. Measuring blink rate and amplitude provides valuable insights into cognitive workload and attention levels [20]. Tracking gaze direction is another essential parameter, as shifts in focus away from the road suggest distraction [20][21]. The analysis of saccades, which are rapid eye movements between fixations, can help assess the ability to shift attention efficiently. Pupil diameter is also a sensitive indicator, with dilation often associated with increased cognitive load. Variations in pupil size can reveal fluctuations in alertness, helping to identify moments of distraction or drowsiness [20]. In addition, evaluating eye movement patterns, including smooth pursuit, can determine the driver's capacity to track moving objects. Jerky or irregular movements may indicate impaired concentration [20][21]. PERCLOS, which stands for Percentage of Eye Closure, measures the percentage of time a person's eyes are closed over a specific duration, usually expressed as a percentage.

Vehicle dynamics parameters. Vehicle dynamics parameters include factors that contribute to understanding the state of the vehicle and the driver's attentiveness. These factors include steering wheel movement, acceleration, deceleration, and lateral movements, which provide valuable insights into the driver's engagement with the driving task [23].

To examine steering behaviour, it is necessary to analyse the frequency, amplitude, and smoothness of steering inputs. Abrupt or erratic steering changes may indicate distraction or drowsiness [23]. Additionally, acceleration and deceleration patterns are critical, as sudden or inconsistent changes in speed may suggest a lack of focus

[23][26]. Lateral movements, such as lane deviations or drifting, can also provide further clues about the driver's state [23][26].

Furthermore, the analysis of vehicle dynamics encompasses parameters such as yaw rate, which measures the vehicle's rotation around its vertical axis. Abnormal yaw behaviour, such as excessive swaying or instability, can be associated with impaired driving attention [23] [26].

Moreover, examining brake usage, gas pedal, and other control inputs contributes to a comprehensive evaluation of driver vigilance [19].

Video-based parameters. Video-based parameters involve the analysis of various visual cues within the captured video feed. The driver's posture is assessed for any unusual slouching or body movements that deviate from the norm [27]. Additionally, head movements can be scrutinised, with a focus on the head orientation that may indicate a lack of attentiveness [20] [23][27]. Another kind of analysis examines the driver's overall spatial awareness and responsiveness to the surrounding environment [19]. Another interesting aspect that has only recently begun to interest the DMS sector is driver emotion recognition. The correlation between drivers' emotional states and attention levels has been extensively studied, with research indicating that emotions significantly influence driving behaviour and cognition. Emotions like happiness and anger not only impact attentional demands but can also lead to aggressive driving behaviours [28]. Recognizing the importance of real-time detection of drivers' emotional states, the DMS field has developed systems to mitigate potential driving risks associated with negative emotions. For instance, Wu et al. [29] designed a DMS that uses a deep convolutional neural network for facial emotion recognition and audio resources to alleviate negative emotions, thereby reducing driving risks. The system has shown promising results in accuracy and reliability. Building on this, [6] proposed an emotion-aware vehicle architecture that adjusts the car's dynamics in response to the driver's emotions, linking negative emotional states to driving performance.

In summary, while modern DMS proposals in the literature highlight the significance of a multimodal approach, they struggle to introduce a system that truly evaluates all the aspects discernible from driver analysis through modern sensors. The aspiration for such a comprehensive system is to not only align with regulatory mandates but to significantly enhance road safety by mitigating risks associated with driver distraction and fatigue.

The literature also reveals a reliance on control unit data in DMS development. While these data provide valuable insights into vehicle dynamics, it does not provide a complete picture as it does not capture the physiological and behavioural nuances of the driver's state. Taken together, these studies highlight the importance of a multimodal, comprehensive approach that takes into account the myriad aspects of the driving experience, from the physical to the behavioural and from the individual to the universal, but also highlight various limitations, such as the difficulty in accurately detecting and differentiating between types of distraction and levels of fatigue. Most existing systems rely heavily on physical indicators, often neglecting cognitive aspects. There is also the ongoing challenge of ensuring accuracy and reliability in diverse and dynamic driving environments. Another limitation is the lack of integration between different types of indicators. Current systems often operate in silos, analysing specific data sets (such as visual cues or driving performance metrics) independently. In addition, many of these systems struggle to process data in real-time, which is essential for timely intervention and accident prevention.

In summary, while current research in DMS has laid a solid foundation, our work aims to overcome the existing limitations by creating a more integrated, adaptive, and real-time system. This advancement not only contributes significantly to the field of driver safety but also paves the way for the development of more sophisticated ADAS and autonomous driving technologies.

2 Material and methods

Our research aims to address the limitations and gaps identified in these studies by defining a comprehensive multimodal approach to interpreting and understanding data on dynamic vehicle behaviour during driving, and data collected through the use of cameras and analysis software based on computer vision and artificial intelligence. In this way, the potential of DMS and ADAS systems, in general, will be explored, where it is possible to use data fusion approaches to integrate different types of data with the common goal of detecting distraction while driving.

The study entails a systematic approach encompassing several key stages:

1. **Parameters analysis**: First, in-depth research is conducted to identify relevant parameters from existing literature and studies related to distraction video analysis and signal correlation during distracted driving.
2. **Selection of parameters**: Parameters essential for the assessment of distracted driving have been selected based on the literature review and the specific objectives of the study.
3. **Data collection**: Experimental studies have been carried out using a driving simulator to record participants' actions in various driving scenarios. The data collected included video footage processed using computer vision and deep learning software developed to analyse facial expressions, and degree of head rotation.
4. **Manual video analysis**: This stage involves observing and annotating driver behaviour during periods of distraction in the video footage. Key visual cues such as head direction have been examined for qualitative insights.
5. **Data pre-processing**: Raw data from the experimental phase are pre-processed, including cleaning, filtering, and organising, to prepare them for subsequent analysis.
6. **Algorithmic parameters calculation**: Algorithms are used to calculate specific parameters previously identified. This automated approach aimed to extract new information from the data for a more objective and consistent analysis.
7. **Correlation analysis**: The final phase focuses on exploring correlations between different types of signals during instances of driver distraction. Statistical methods were used to identify relationships between facial expressions, and vehicle dynamics parameters.

In this study we have chosen to experiment with a select few parameters identified from the state of the art: Time to Line Crossing (TLC), Steering wheel reversal rate (SWRR), Steering wheel rapid movement (SWRM), Engagement from emotional analysis, and head orientation measures such as yaw. This selection was made by referring to the most representative parameters in the state of the art to assess the driver's level of distraction:

1. Time to Line Crossing (TLC).
2. Steering wheel reversal rate (SWRR).
3. Steering wheel rapid movement (SWRM)
4. Engagement from the Emotional Analysis component assesses the “excitement” associated with the driver's affective state, which can influence driving behaviour.
5. Head orientation Significant deviations in yaw and pitch values from the normative range can indicate distractions, as the driver may be looking away from the road.

By selecting these parameters, we aim to create a multi-dimensional profile of the driver's state that is accurate and comprehensive, ensuring that our experimental setup is focused, manageable, and aligned with the key variables that will be empirically validated to affect driving performance. The inclusion of emotional engagement and attentional cues in the integration of these parameters offers a novel approach to DMS development that is sensitive to both the cognitive and affective dimensions of the driving task.

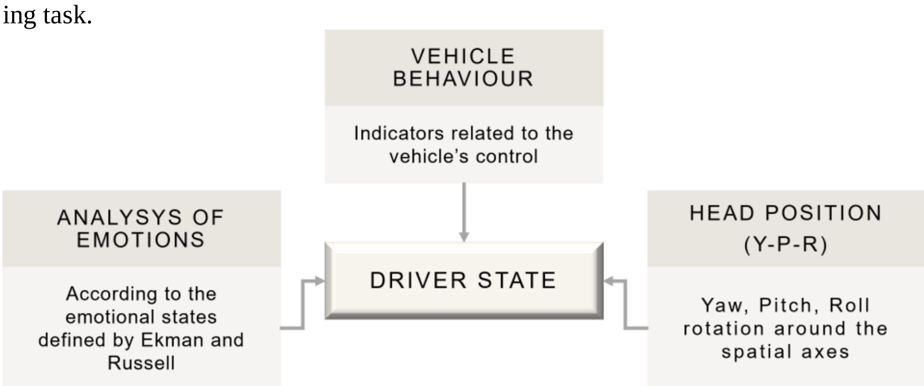

Fig. 1. Considered categories of data for the proposed multimodal approach

2.1 Computer vision and Deep Learning software

This work incorporates computer vision and deep learning algorithms capable of processing large datasets in real-time. This enables the system to adapt to individual driver characteristics, potentially providing personalised assessments and warnings:33

- Emotion Recognition Module: This module, based on a Convolutional Neural Network (CNN) tested and described in [30], was trained using a merged dataset with both lab-based and real-world data. The CNN was trained with public datasets like CK+ [31] and FER+ [32], and the “in the wild” dataset provided by Affectnet [33].

The final dataset composed of 250k photos was split using an 80-20 proportion for training and validation phases. In the final layer of the CNN architecture, a softmax function computes the scores for Ekman's emotions from each video frame captured by the camera, with the results normalised to a sum of 100 percent. Building on Russel's Circumplex model [34], which categorises Ekman's emotions based on valence and arousal, we calculate engagement from the emotion percentages. Engagement scores span from 0 for completely neutral expressions to 100 for highly engaging ones.

Engagement = Happiness(%) + Surprise(%) + Anger(%) + Fear(%) - Sadness (%) (1)

- Head Direction Module: This module, tested in [35], detects driver attentiveness from the estimation of head orientation with respect to the camera, using an approach similar to the one described in [36] to estimate yaw, pitch, and roll parameters. Eye Aspect Ratio (EAR) is also utilised, calculated based on the eye blinking detection method outlined in [37]. Originally designed for eye closure detection, the EAR has also proven to be useful in identifying frontal head tilting by users. This module returns yaw, pitch and roll values related to the degree of head rotation relative to the camera view.

3 Experimental Procedure

While this research endeavours to explore the correlations between signals typically associated with driver distraction, it is important to note that the experimental phase did not simulate drowsiness states. This omission was a deliberate methodological choice, primarily due to the challenges inherent in inducing a state of drowsiness in a controlled environment. Simulating drowsiness can be complex and requires careful monitoring to ensure the safety and well-being of participants, which may not always be feasible within the constraints of an experimental setting.

The explored case study was conducted using a driving simulator operating on the SCANeR Studio 1.7 platform, which is adept at simulating sensors and automated driving functionalities, essential for a realistic and controlled testing environment. The static driving simulator is equipped with an engine simulation and authentic car controls, such as a driver's seat, pedals, and gearbox. Additionally, it features a SensoDrive steering wheel with haptic force feedback to emulate real driving conditions. The simulation environment is enhanced with visual aids, including a video projector and a 15.6" display behind the steering wheel, presenting a full digital Human-Machine Interface. Synchronisation is accomplished by using the simulator machine timestamp, which enables a distributed system architecture that is necessary for testing and validating the AI algorithms involved. The vehicle's various parameters could be monitored and controlled by drivers through a tablet located on the dashboard. This interface served as both a control mechanism and a means of introducing distraction events, simulating the impact of external factors on the driver's attention.

Fig. 2. The driving simulator

The driver was captured from multiple angles using strategically positioned HD cameras. One of the cameras was connected to an Intel NUC11PAHi7, which ran the software for emotion analysis and head analysis. Other cameras provided additional cabin perspectives. This enabled a comprehensive analysis of the driver's reactions and engagement.
The key hardware integrated into the simulator included:

- Hikvision Digital Technology DS-2CE16H0T-ITF HD cameras for high-definition recording
- A Samsung Galaxy Tab A8 used for vehicle interaction and distraction induction.
- Onboard dedicated processing modules, integrated with the simulator, to monitor vehicle dynamics in real time, supporting a lifelike driving simulation and enabling crucial data collection.

The AVSimulation SCANeR software has been used to create two environmental scenarios for the driving simulation: highway and urban. The highway scenario incorporated random behaviours and overtaking capabilities to simulate real traffic conditions. The urban setting was characterised by a dense network of traffic signals, roads, and pedestrian crossings, with diverse road users including cars, motorcycles, bicycles, and pedestrians. These environments were designed to simulate varying traffic and weather conditions, challenging drivers to adapt their behaviour to the complexities of the road.

Within these simulations, three specific contexts were crafted to elicit different emotional and cognitive states in drivers. The first context mimicked the urgency of a parent late to collect their child from school. The second scenario envisioned the driver setting off on a much-anticipated vacation, aiming to induce a sense of relaxation and joy. The third scenario involved the driver being called into work on their day off for an urgent meeting, intended to provoke feelings of frustration and stress.

Fig. 3. An example of a simulated scenario

To further engage the drivers and introduce elements of distraction, five secondary tasks were integrated into each driving session. These tasks required drivers to interact with various objects within the vehicle and execute specific actions such as interacting with a simulated Human-Machine Interface (HMI) of the car's dashboard, in this case a tablet. These interactions involved completing various tasks to simulate potential distraction activities in real-world scenarios, including:

- Answering the phone
- Sending messages on the phone
- Talking on the phone while holding it to the ear
- Drinking

These were employed to mimic realistic distractions and observe the drivers' responses. Through a systematic combination of these distraction events and the sequence in which they were performed, paired with different driving conditions, five distinct scenarios emerged that were colour-coded for clarity. These scenarios, labelled with the colours blue, yellow, orange, green, and purple, provided a comprehensive list of distraction levels and driving conditions for analysis.

Nineteen participants, 9 males and 10 females aged between 21 to 53, were enlisted for the study. Each participated in three separate 10-minute sessions, each presenting one of the three contexts. Throughout these sessions, participants engaged in the prescribed tasks to simulate distraction, allowing for a comprehensive assessment of their reactions within the simulated driving environments.

4 Experimental Results

4.1 Qualitative analysis

The findings of our study are here analysed using visual data, highlighting the behaviour of certain indices across the five colour-coded test scenarios designed to study driver distraction under varying traffic conditions. The data encompassed telemetry metrics, facial expressions analysed through Ekman's framework, and head orientation movements measured in yaw, pitch, and roll.

We'll analyze how these data, retrieved during the distracting scenarios, compare with the data outside these events.

The aim of this qualitative analysis is, in fact, to spot particoular trends in these parameters, especially during moments of distraction.

First step regards the observations from the facial expression analysis software, which identified notable changes in engagement and valence among subjects during distractions, identifying patterns that suggest a relationship between certain expressions and decreased driving performance.

For example, subject 16 showed a significant shift in engagement during specific distraction tasks, such as responding to messages and using a tablet. This variability in emotional responses underscores the complexity of driver reactions to distractions and highlights the potential for emotional states to influence driving behaviour.

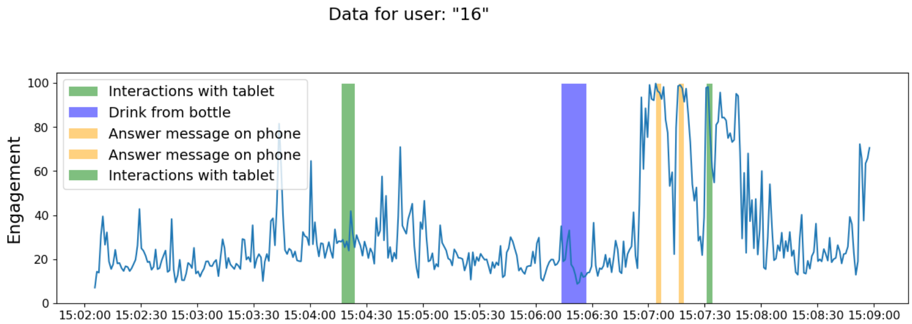

Fig. 4. Engagement and Valence data of subject coded "16" showing their trends across distractions events highlighted by the coloured vertical bars

Similarly, the analysis of data collected by the software to gauge driver attention based on head orientation reveals that significant yaw deviations correspond to decreased attention levels, as anticipated, since yaw rotation signifies turning around the vertical axis. Consequently, elevated values of this measurement suggest that the driver's head is turned away from the front. During the tests, most distraction activities required the user to turn right since the standard driving position in Europe is on the left of the car, resulting in the highest deviations observed either during these activities or at the beginning and end phases of the test when drivers typically glance around.

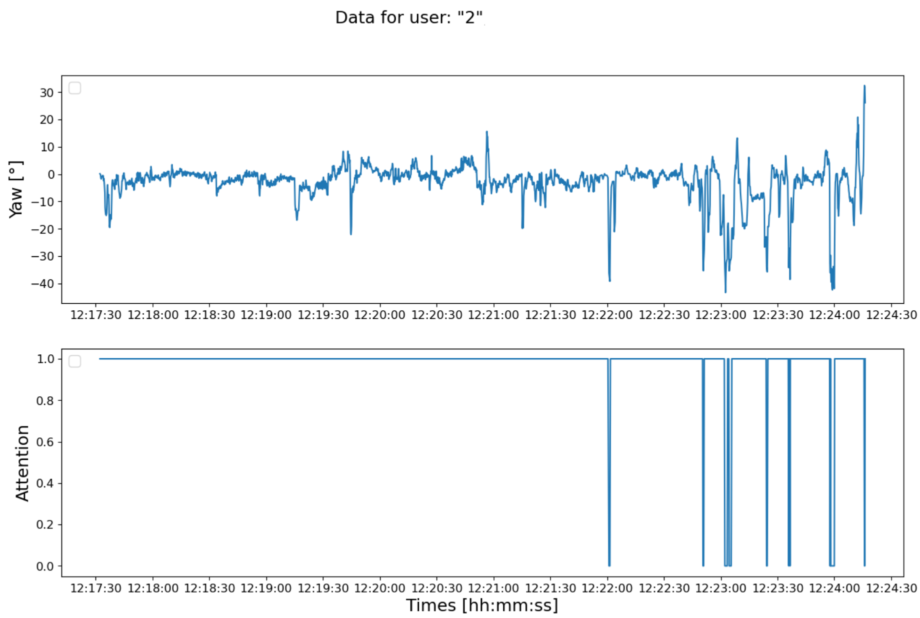

Fig. 5. Yaw and Attention data of subject coded “2” showing their trends across the driving simulation

Some subjects exhibited peaks in emotional engagement before the distraction event occurred, suggesting that anticipation or reaction to the instruction itself could be a distraction. However, for many subjects, no clear trends were identified, possibly due to inaccuracies in data collection, particularly with camera-based measurements. Factors such as inadequate lighting and improper camera angles could have compromised the data quality, affecting the reliability of emotion detection based on facial expressions.

Vehicle dynamics data further elucidated the impact of distractions on driving performance. The analysis of telemetry data shows a recurring trend highlighted in the literature. When drivers become distracted or anticipate distraction, they typically exhibit a decrease in vehicle speed. Such speed reduction is commonly associated with a protective instinct wherein drivers preemptively slow down upon recognizing the impending engagement in a secondary task that will divert their attention from driving. This behavior was consistent across subjects, with notable deviations in speed trends observed during distraction scenarios. Subject 2 exhibits this trend in the most clear and evident manner during the “Drink from bottle” distraction event, where cognitive engagement is required to locate and open the bottle.

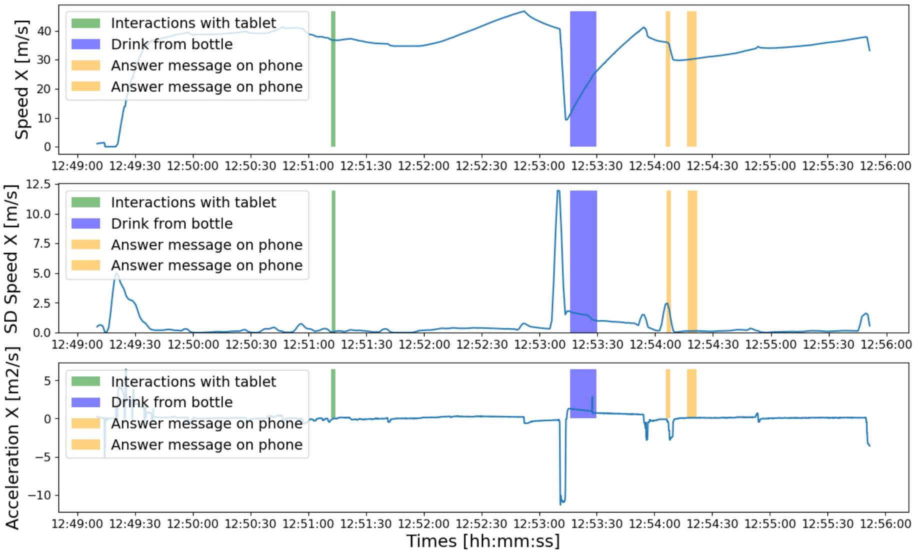

Fig. 6. Telemetry data of subject coded "2" showing their trends across distractions events highlighted by the coloured vertical bars

4.2 Quantitative analysis

A quantitative analysis has been then conducted to explore correlations between the available data and indicators, specifically focusing on steering wheel reversal rate (SWRR), steering wheel rapid movement (SWRM), and time to lane crossing (TLC). The aim was to determine whether data from non-distraction periods could predict outcomes during distraction events, offering insights into the predictive potential of these metrics across varying levels of driver distraction. Our analysis also involved a meticulous categorization of driving behaviour data, separating it into two groups: 'data with events,' captured during distraction periods, and 'data without events,' representing normal driving conditions. This segmentation facilitated a targeted examination of how distraction events influence driving behaviour. From these groups, we conducted additional subdivision based on thresholds exceeded by the specific indicators TLC, SWRR, and SWRM. The primary objective was to discern differences between the datasets with and without events across all combinations of indicators: this granular classification resulted in fourteen distinct subsets, seven corresponding to 'data with events' and seven to 'data without events'. This comparative analysis aimed to identify behavioral changes during distraction events, providing insights into how such events alter driving patterns. Through this approach, we sought to understand the specific impact of distraction events on driving behavior, independent of individual indicators, allowing for a comprehensive examination of the driving dynamics under varying conditions of attention. The boxplots in figure 7 serve to illustrate the dataset subdivision utilized as the foundation for subsequent analysis. At the top, we present the complete dataset without events on the left, alongside a subset extracted from it where all indicators exceeded their respective thresholds as indicated in the boxplot (e.g., TLC < 6.4 seconds, SWRM > 13 °/s, and SWRR > 6 °) on the right. The bottom row mirrors this setup, but now exclusively features data collected during distraction events. In particular, we show Speed Y data in the boxplots with the described subdivision.

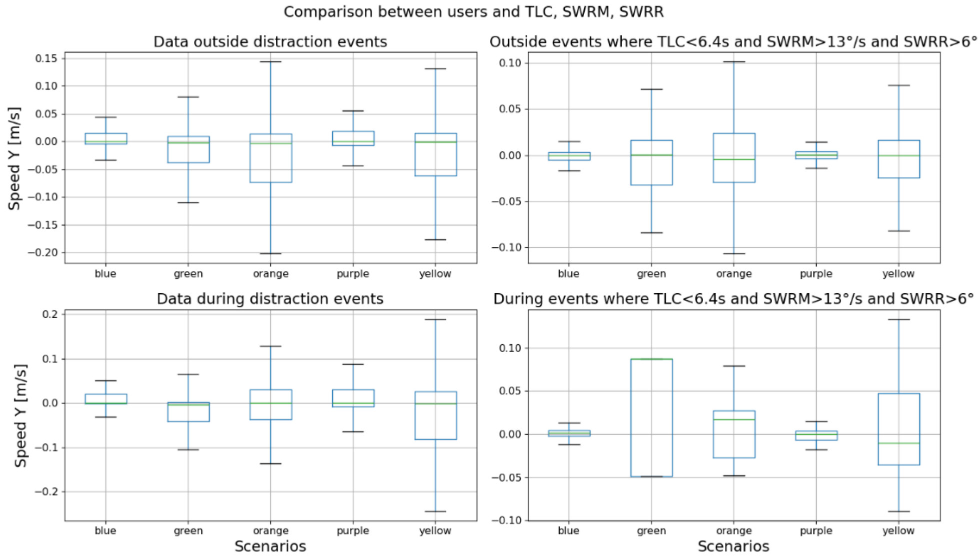

Fig. 7. Boxplots of ‘data without events’ (top) and ‘with events’ (bottom). To the right a subset from each dataset shown on the left is taken under the condition that the three indicators go over the thresholds indicated

This investigation sought to determine which combinations of indicators most closely mirrored the 'data without events', suggesting their potential utility in predicting key variables like Speed X or steering wheel angle during distraction events. This approach was grounded in statistical analyses, including Mann-Whitney, Kolmogorov-Smirnov, and t-Tests, to accommodate the non-normal distribution of our data, despite its large volume.

Variable	**N. KPI combinations**
acceleration_y	33
steeringWheelAngle	26
speed_y	23
speed_x	14
acceleration_x	11
tlc	10
roadInfo_laneGap	8
steeringWheelSpeed	1

These results illustrate the frequency of different variables appearing in combinations that could potentially predict distraction events based on various indicators. Acceleration in the y-direction appears most frequently, with 33 combinations across the several driving scenarios performed by the users, followed by steering wheel angle with 26 combinations and speed in the y-direction with 23 combinations. Other variables such as speed in the x-direction, acceleration in the x-direction, time to lane crossing (TLC), and road lane gap information also show notable occurrences. These findings suggest that certain variables may play a more prominent role in predicting distraction events, providing valuable insights for further analysis and mitigation strategies in enhancing driver safety. For instance, specific KPI combinations that yield y-axis acceleration values similar to those observed during distraction events indicate the occurrence of such events. This finding is instrumental in delineating the most effective combinations of indicators for predicting distraction, ranging from individual thresholds like TLC < 6.4 seconds, SWRM > 13 °/s, SWRR > 6 °, to composite thresholds involving multiple indicators.

In conclusion, the quantitative analysis underscores the significance of y-axis acceleration and steering wheel angle as key variables for detecting driver distraction. Through meticulous statistical testing and data segmentation, we've identified specific indicator combinations that offer a predictive lens for monitoring and potentially mitigating distraction-induced impairments in driving performance.

5 Discussion and Conclusion

This study aims to improve the understanding and development of DMS through a meticulous analysis of state-of-the-art parameters and processing techniques. Our primary objective was to investigate the essential components that underpin the effectiveness of DMS. To achieve this, we introduced a multimodal approach to data processing that seamlessly integrates information from the vehicle's control unit with vision-based data. This innovative fusion has significantly improved the accuracy and reliability of DMS. The findings from the experimental investigation provide a compelling narrative on the nature of driver distraction. The correlations observed among the diverse signals—ranging from steering patterns to emotional cues—demonstrate a significant alignment with the actual states of distraction as reported by participants, validating the ground truth data. These results underscore the potential of multi-signal analysis as a reliable indicator of distracted driving behaviors. Crucially, the study has shown that by integrating various data streams, such as the Standard Steering Wheel Reversal Rate (SWRR), Time to Line Crossing (TLC), Steering wheel rapid movement (SWRM), along with emotional engagement, and face orientation, we can achieve a robust detection system. The consistency of these signals with self-reported distraction instances strengthens the argument for a multimodal DMS capable of discerning subtle shifts in driver focus and engagement. This research contributes to the growing body of evidence that supports the use of complex data fusion in DMS. By leveraging a combination of behavioural, physiological, and emotional signals, we can enhance the predictive power of these systems, potentially leading to more responsive and adaptive safety mechanisms in vehicles. Moving forward, these conclusions not only pave the way for more advanced DMS development but also highlight the importance of continuous data validation against real-world scenarios. Future research can build on these foundations, exploring broader signal sets and refining detection algorithms to cater to the nuanced spectrum of driver behaviours. Although no states of drowsiness were simulated in the experiment, the findings related to distraction are still highly relevant, as the behaviours exhibited during distracted driving can share similarities with those observed in drowsy driving, such as delayed reaction times and decreased vehicle control. We acknowledge this limitation in our research design and propose that future studies could incorporate physiological measures, such as EEG or more accurate eye-tracking metrics that do not require the actual induction of drowsiness but can infer the state based on known biomarkers. Moreover, the development of advanced simulation technologies and safer experimental protocols could eventually allow for the inclusion of drowsiness simulations in a manner that is both ethical and effective. However, while the current study does not include drowsiness simulations, the insights gained from the analysis of distraction signals contribute valuable knowledge to the field of DMS and lay the groundwork for future research to build upon.

References

1. Antony M.M., Whenish R., 2021, Advanced Driver Assistance Systems (ADAS) in AKathiresh, M., Neelaveni, R. (eds) Automotive Embedded Systems. EAI/Springer Innovations in Communication and Computing. Springer, Cham. pp. 165-181
2. Owens J, Dingus T, Guo F, et al. Prevalence of Drowsy-Driving Crashes: Estimates from a Large-Scale Naturalistic Driving Study. Washington, DC: AAA Foundation for Traffic Safety; 2018.
3. Fitzharrris M., Liu S., Stephens A.N., Lenné M.G., 2017, the relative importance of real-time in-cab and external feedback in managing fatigue in real-world commercial transport operations, Traffic Injury Prevention, 18 (1), pp.71-78
4. Ortega J.D., Canas P.N., Nieto M., Otaegui O., Salgado L., 2022, Challenges of Large-Scale Multi-Camera Datasets for Driver Monitoring Systems, Sensors, Vol. 22 (7), pp. 2554
5. Hasenjäger, M., Heckmann, M., & Wersing, H. (2019). A survey of personalization for advanced driver assistance systems. IEEE Transactions on Intelligent Vehicles, 5(2), 335-344.
6. Ceccacci, S., Mengoni, M., Andrea, G., Giraldi, L., Carbonara, G., Castellano, A., & Montanari, R. (2020). A preliminary investigation towards the application of facial expression analysis to enable an emotion-aware car interface. In Universal Access in Human-Computer Interaction. Applications and Practice: 14th International Conference, UAHCI 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part II 22 (pp. 504-517). Springer International Publishing.
7. European Commission (2021) Road safety thematic report – Serious injuries. European Road Safety Observatory. Brussels, European Commission, Directorate General for Transport.
8. Gardony, A. L., Lindeman, R. W., & Brunyé, T. T. (2020, February). Eye-tracking for human-centered mixed reality: promises and challenges. In Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR) (Vol. 11310, pp. 230-247). SPIE.
9. Dehzangi, O., Sahu, V., Taherisadr, M., & Galster, S. (2018, November). Multi-modal system to detect on-the-road driver distraction. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC) (pp. 2191-2196). IEEE.

10. Muhammad, K., Ullah, A., Lloret, J., Del Ser, J., & de Albuquerque, V. H. C. (2020). Deep learning for safe autonomous driving: Current challenges and future directions. IEEE Transactions on Intelligent Transportation Systems, 22(7), 4316-4336.
11. Daza, I.G., Bergasa, L.M., Bronte, S., Yebes, J.J., Almazán, J., Arroyo, R. Fusion of Optimized Indicators from Advanced Driver Assistance Systems (ADAS) for Driver Drowsiness Detection. Sensors 2014, 14, 1106-1131.
12. Sandberg, D., Akerstedt, T., Anund, A., Kecklund, G., & Wahde, M. (2010). Detecting driver sleepiness using optimized nonlinear combinations of sleepiness indicators. IEEE Transactions on Intelligent Transportation Systems, 12(1), 97-108.
13. Zhang, H., Wu, C., Huang, Z., Yan, X., & Qiu, T. Z. (2016). Sensitivity of Lane Position and Steering Angle Measurements to Driver Fatigue. Transportation Research Record, 2585(1), 67-76.
14. Khandakar, A., Chowdhury, M. E., Ahmed, R., Dhib, A., Mohammed, M., Al-Emadi, N. A. M., & Michelson, D. (2019). Portable system for monitoring and controlling driver behavior and the use of a mobile phone while driving. Sensors, 19(7), 1563.
15. Nemcova, A., Svozilova, V., Bucsuházy, K., Smíšek, R., Mezl, M., Hesko, B., ... & Kolar, R. (2021). Multimodal features for detection of driver stress and fatigue. IEEE Transactions on Intelligent Transportation Systems, 22(6).
16. Dinges, DF: An Overview of Sleepiness and Accidents. Journal of Sleep Research 1995; 4(2):4–14.
17. Kircher, A., Uddman, M., & Sandin, J. (2002). Vehicle control and drowsiness. Statens väg- och transportforskningsinstitut.
18. Bunji, A: Evaluation of Mental Condition on Drivers by Analysis of Heart Rate Variability. Journal of the Society of Automotive Engineers of Japan No. 9437601, 1994.
19. Milardo, S., Rathore, P., Amorim, M., Fugiglando, U., Santi, P., & Ratti, C. (2021). Understanding drivers' stress and interactions with vehicle systems through naturalistic data analysis. IEEE Transactions on Intelligent Transportation Systems, 23(9), 14570-14581.
20. Bergasa, L. M., Nuevo, J., Sotelo, M. A., Barea, R., & Lopez, M. E. (2006). Real-time system for monitoring driver vigilance. IEEE Transactions on intelligent transportation systems, 7(1), 63-77.
21. Çetinkaya, M., & Acarman, T. (2023). Driver impairment detection using decision tree based feature selection and classification. Results in Engineering, 18, 101025.
22. Wylie, C. D., Shultz, T., Miller, J. C., Mitler, M. M., & Mackie, R. R. (1996). Commercial motor vehicle driver fatigue and alertness study: Technical summary.
23. Doudou, M., Bouabdallah, A. & Berge-Cherfaoui, V. Driver Drowsiness Measurement Technologies: Current Research, Market Solutions, and Challenges. Int. J. ITS Res. 18, 297–319 (2020).
24. Sagberg, F: Road Accidents Caused by Drivers Falling Asleep. Accident Analysis and Prevention 1999; 31: 639–649.
25. Verwey, WB & Zaidel, D: Predicting Drowsiness Accidents from Personal Attributes, Eye Blinks, and Ongoing Driving Behaviour. Report TM-97-B009. 1997; Soesterberg. TNO Human Factors Research Institute, The Netherlands.
26. Aksjonov, A., Nedoma, P., Vodovozov, V., Petlenkov, E., & Herrmann, M. (2018). Detection and evaluation of driver distraction using machine learning and fuzzy logic. IEEE Transactions on Intelligent Transportation Systems, 20(6), 2048-2059.
27. Bittner, R & Smrcka, P & Vysoky, P & Hana, K & Pousek, L & Schreib, P: Detecting of Fatigue States of a Car Driver. First International Symposium on Medical Data Analysis - ISMDA 2000, Frankfurt, Germany, September 2000, ISBN 3-540-41089-9, Springer-Verlag Berlin Heidelberg 2000, pp. 260–273.

28. Sârbescu, P. Aggressive driving in Romania: Psychometric properties of the driving anger expression inventory. Transportation research part F: traffic psychology and behaviour 2012, 15(5), 556-564.
29. Wu, Y. L., Tsai, H. Y., Huang, Y. C., & Chen, B. H. (2018, October). Accurate emotion recognition for driving risk prevention in driver monitoring system. In 2018 IEEE 7TH global conference on consumer electronics (GCCE) (pp. 796-797). IEEE.
30. A. Talipu, A. Generosi, M. Mengoni, L Giraldi, "Evaluation of Deep Convolutional Neural Network architectures for Emotion Recognition in the Wild," in 2019 IEEE 23rd International Symposium on Consumer Technologies, 25-27, 2019. IEEE https://doi.org/10.1109/ISCE.2019.8900994
31. P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, I. Matthews, "The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression," in 2010 ieee computer society conference on computer vision and pattern recognition-workshops, 2010. https://doi.org/10.1109/CVPRW.2010.5543262
32. E. Barsoum, C. Zhang, C. C. Ferrer, and Z. Zhang, "Training Deep Networks for Facial Expression Recognition with Crowd-Sourced Label Distribution" 2016. https://doi.org/10.1145/2993148.2993165
33. Mollahosseini, A., Hasani, B., & Mahoor, M. H. (2017). Affectnet: A database for facial expression, valence, and arousal computing in the wild. IEEE Transactions on Affective Computing, 10(1), 18-31.
34. Feldman Barrett, L., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. Journal of personality and social psychology, 74(4), 967.
35. Ceccacci, S., Generosi, A., Cimini, G., Faggiano, S., Giraldi, L., & Mengoni, M. (2021, January). Facial coding as a mean to enable continuous monitoring of student's behavior in e-Learning. In teleXbe.
36. Mallick, S. (2021, May 4). Head Pose Estimation using OpenCV and Dlib | LearnOpenCV #. LearnOpenCV – Learn OpenCV, PyTorch, Keras, Tensorflow With Examples and Tutorials. https://learnopencv.com/head-pose-estimation-using-opencv-and-dlib/
37. Soukupova, Tereza, and Jan Cech. (2016) "Eye Blink Detection Using Facial Landmarks." 21st Computer Vision Winter Workshop, Rimske Toplice, Slovenia