Supplementary Material

A. Specifications for simulator system and sensor suite

Fig. 5 provides a visual representation of the research simulator components. Tab. 5 presents a detailed list of integrated components in the MiX Telematic research simulator, including their specifications.

	Hardware specification
Main RGB camera	Varifocal bullet camera High-definition (HD) with a 1/2.8 SONY Sensor Resolution: Adjustable from 2MP to 5MP at 30 frames per second (Hikvision, China)
3D camera	ZED 2 Stereo full-color 3D camera Resolution: 1080P and 30 frames per second (StereoLabs, France)
Rear-view RGB camera	IPC full-color camera Resolution: 720P and 15 frames per second (Streamax, Australia)
IR camera	DSM camera with infrared LED illumination Resolution: 960P and 30 frames per second (Streamax, Australia)
Alcohol breathalyser	Autowatch 720 Tethered Alcohol Breathalyser (TAB) Sensor: Fuel Cell (Electrochemical) Measurement range: 0.005 to 0.400 BAC (grams/210 liters) Accuracy: ±0.02mg/L (PFK Electronics, South Africa)
Card reader	Mifare-type card reader operating at 13.56MHz Sensor: Fuel Cell (Electrochemical) To read unique identification codes (UIDs) from the pre-programmed Mifare Fudan (1K) cards provided to participants (DriveMate, UK)

Table 5. Simulator hardware specifications.

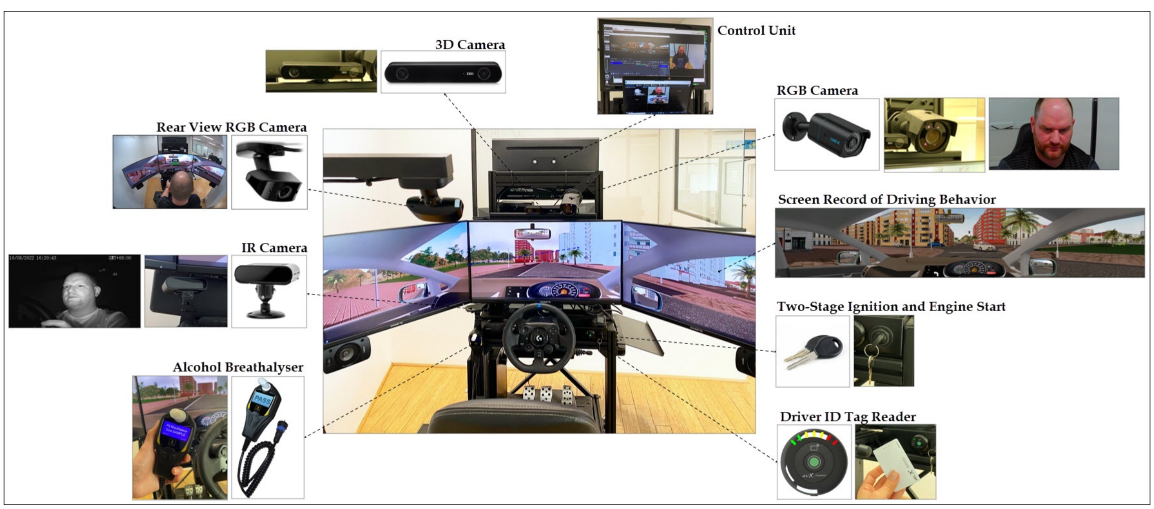

Figure 5. Integrated components of the research simulator.

Supplementary Material

B. Data collection procedure overview and BAC observations

The process of alcohol administration and subsequent driving sessions is illustrated in Fig. 6.

Tab. 6 shows the recorded Blood Alcohol Concentration (BAC) measurements of participants at targeted intoxication levels during our data collection.

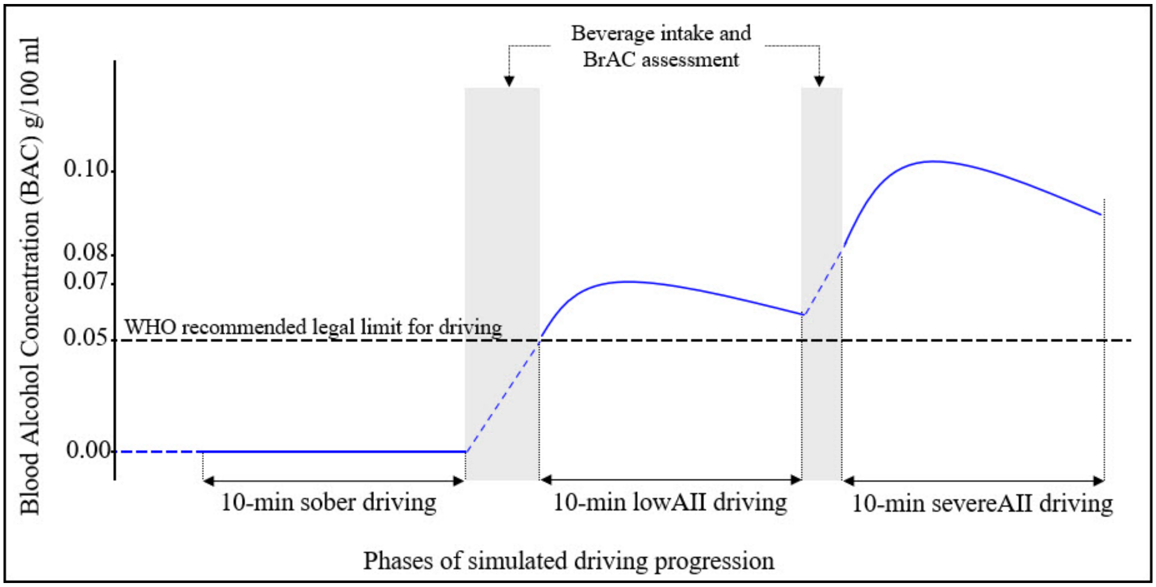

Figure 6. Overview of data collection procedure (Alcohol administration process and subsequent driving sessions).

Target intoxication level	Recorded BAC value	g/100ml
Sober: 0.00 g/100ml	Min	0.000
	Max	0.000
	Mean±SD	0.000 ± 0.000
LowAII: between 0.05 and 0.07 g/100ml	Min	0.051
	Max	0.069
	Mean±SD	0.058 ± 0.005
SevereAII: greater than 0.08 g/100ml	Min	0.081
	Max	0.165
	Mean±SD	0.104 ± 0.019

Table 6. BAC Observations at targeted intoxication levels.

Supplementary Material

C. Comparison with other works

When comparing our study with existing solutions, several key factors highlight the superiority and novelty of our proposed method in alcohol intoxication detection, which are summarized in Tab. 7

	Other works	Our work
Classification Method	Binary classification of sober vs. drunk [8, 11, 18, 25, 28, 39]	Multi-level classification system (sober, low alcohol impairment and severe alcohol impairment)
Sensor Usage	Heavily depend on additional sensors: Vehicle operation sensors (e.g., speed, pedal pressure, and position/orientation sensor [8, 11, 18, 25, 27, 28, 39] Physiological measurement devices (e.g., heart rate sensor, seat mat pressure sensor, and EMS, EDA, and PPG biometric sensors [39] [25] [8] Advanced monitoring sensors including eye and gaze tracking sensors [23, 25]	Exclusively utilizes RGB video data
Safety	Contingent on driving performance. Requires the possibly impaired driver to continue driving until intoxication is detected [8, 11, 18, 25, 27, 28, 39]	Not contingent on driving performance. It can identify signs of intoxication from the driver's face even before any attempt to drive

Table 7. Comparison with other works.

In summary, while our reported accuracy might be slightly lower in direct comparison, the unique combination of a detailed 3-class classification system, minimal sensor requirements, and avoidance of prolonged impaired driving positions our proposed method as a pioneering and promising solution for alcohol intoxication detection.

D. Detailed limitation of the current work and roadmap to real-world implementation

To successfully implement alcohol intoxication detection in real-world scenarios, the ideal system should operate in real time, delivering immediate and precise results while seamlessly adapting to varying environments, including diverse lighting conditions and driving behavior. However, one significant challenge in alcohol intoxication detection for drivers (as discussed in section 5.2) is the necessity of conducting our research using a driving simulator due to legal constraints. While efforts have been made to replicate real-world driving experiences within the simulator, it cannot fully replicate all variables and events from actual driving scenarios, which presents challenges in adapting to diverse lighting conditions and other real-world factors.

Besides inherent limitations in simulator-based data collection, and those detailed in section 5.2, our current work has identified several other limitations that indicate areas for improvement and avenues for future research in alcohol intoxication detection for drivers. Age-related factors are one consideration in this context, and our data collection included participants of diverse ages to account for age-related changes in facial features and behaviors that may impact detection accuracy. For future studies, investigating age-related variations in physiological changes associated with alcohol intoxication can provide valuable insights. If significant differences are found, developing age-specific detection models customized for distinct facial and physiological traits within specific age groups can enhance accuracy.

Another concern lies in achieving dependable and accurate alcohol intoxication detection in real-world implementation, solely relying on facial features. This challenge arises from the variability in facial expressions, the vulnerability of facial feature analysis to external factors like lighting, and the potential for subjective interpretation, all of which may increase the risk of false positives. One approach to addressing this challenge, while retaining an RGB-based solution without relying on additional sensors, involves integrating various data sources such as body movement patterns and driving performance. The integration of this data with contextual information such as the driver's history, time of day, location, and driving behavior within a multimodal detection system offers the potential to enhance the accuracy of intoxication detection. Additionally, implementing a real-time feedback loop that allows the detection algorithm to continuously learn and adapt based on user feedback and new data, can ensure the system's enduring relevance and precision.

Furthermore, employing a multimodal detection system could potentially mitigate the challenges posed by the diminished performance of RGB cameras in low-light conditions. A thorough investigation is required to determine whether integrating patterns related to intoxicated behavior in RGB video footage and applying image enhancement techniques like histogram equalization or adaptive contrast enhancement can improve the system's effectiveness. Additionally, it is worth noting that investing in Low-Light RGB Cameras tailored for dim conditions or utilizing infrared (IR) lights to illuminate areas lacking visible light can enhance the quality of RGB video captured in low-light or complete darkness.