



**Nova**  
NOVA SCHOOL OF  
SCIENCE & TECHNOLOGY

DEPARTMENT  
OF PHYSICS

**TELMA SOFIA CALDEIRA ESTEVES**

Bachelor of Science in Biomedical Engineering

**SLEEPY DRIVERS:  
EYE-STATE AND DROWSINESS MONITORING  
USING FACE VIDEO**

MASTER IN BIOMEDICAL ENGINEERING

NOVA University Lisbon  
December, 2023



# SLEEPY DRIVERS: EYE-STATE AND DROWSINESS MONITORING USING FACE VIDEO

**TELMA SOFIA CALDEIRA ESTEVES**

Bachelor of Science in Biomedical Engineering

**Adviser:** André Ribeiro Lourenço

*Invited Professor, Instituto Superior de Engenharia de Lisboa - Instituto Politécnico de Lisboa  
CEO, CardioID Technologies*

**Co-adviser:** Ricardo Nuno Pereira Verga e Afonso Vigário

*Associate Professor, NOVA School of Science and Technology - NOVA University Lisbon*

## **Examination Committee:**

**Chair:** Hugo Filipe Silveira Gamboa

*Associate Professor with Habilitation, NOVA School of Science and Technology - NOVA University Lisbon*

**Rapporteur:** Hugo Humberto Plácido da Silva

*Senior Researcher, Instituto de Telecomunicações, Instituto Superior Técnico*

**Adviser:** André Ribeiro Lourenço

*Invited Professor, Instituto Superior de Engenharia de Lisboa - Instituto Politécnico de Lisboa*

## **Sleepy Drivers: Eye-state and Drowsiness Monitoring Using Face Video**

Copyright © Telma Sofia Caldeira Esteves, NOVA School of Science and Technology, NOVA University Lisbon.

The NOVA School of Science and Technology and the NOVA University Lisbon have the right, perpetual and without geographical boundaries, to file and publish this dissertation through printed copies reproduced on paper or on digital form, or by any other means known or that may be invented, and to disseminate through scientific repositories and admit its copying and distribution for non-commercial, educational or research purposes, as long as credit is given to the author and editor.

*Para a Zoca e para os meus avós.*

## ACKNOWLEDGEMENTS

Seria inconcebível terminar esta dissertação sem a expressão sincera dos meus agradecimentos a todas as pessoas que contribuíram para que este trabalho tomasse forma e passasse, finalmente, para o papel.

Ao Professor Doutor André Lourenço (CardioID), um grande *role model*, agradeço do fundo do coração. Ter a tua orientação, a vários níveis, constituiu e constitui uma oportunidade que valorizo acima de tudo. A confiança que depositaste em mim e neste projeto e a tranquilidade e competência características da tua liderança foram contributos inestimáveis para o meu bem-estar e para a execução final deste trabalho. A Professora Doutora Ana Rebelo (INESC TEC) e o Doutor (*parabéns novamente!*) João Ribeiro Pinto (INESC TEC) foram também dois orientadores de excelência que, devido à existência de um limite no tamanho das equipas de orientação, não puderam ser classificados como tal na proposta desta dissertação. Não obstante, Ana e João, a vossa disponibilidade para as reuniões periódicas, o vosso apoio, a vossa passagem de conhecimento, e a vossa compreensão foram, no mínimo, fundamentais. Agradeço também ao Professor Ricardo Vigário (FCT NOVA) por ter permitido que este projeto avançasse, ao aceitar ser o elemento de ligação com a Universidade, e por ter uma postura sempre encorajadora. Em suma, um grande obrigada aos meus orientadores. Foram verdadeiramente espetaculares.

Neste projeto, tive a imensa sorte de ser acolhida não por uma, mas por duas equipas excecionais: o INESC TEC, onde iniciei os trabalhos, e a CardioID, onde os continuei e finalizei. Senti-me integrada em ambas, até nos tempos de pandemia.

No INESC TEC, fui imediatamente recebida com workshops de acolhimento aos *Master Students* e, no Verão, organizaram ainda a *summer school* VISUM. São uma equipa incansável que proporciona todas as oportunidades para aprendermos. Além dos meus orientadores, deixo um profundo agradecimento ao VCMI e ao *Biometrics Subgroup*, em particular à Doutora Ana Filipa Sequeira e ao Professor Doutor Jaime Cardoso, cujas reuniões alargaram os meus horizontes. Um agradecimento especial ainda à Inês Antunes, pelos grandes momentos de empatia que partilhámos antes, durante, e após a nossa breve passagem pelo *subgroup*.

Na CardioID, encontrei um ambiente que se tornou como uma família para mim (estou consciente do cliché). Além do André, todos me deixaram uma marca indelével e contribuíram significativamente para este trabalho. Carlos, David, Eduardo, Lourenço, Roberto, Zé, obrigada pelo companheirismo, por me fazerem sentir engenheira, obrigada

---

por momentos como o do transporte do nosso simulador.

A todos os professores "por quem passei", desde o início, o meu profundo reconhecimento. Obrigada por partilharem o vosso conhecimento e por moldarem o crescimento de tantas pessoas. Obrigada às minhas escolas: Externato de Penafirme, Liceu Pedro Nunes e, claro, FCT NOVA.

Obrigada à grande família Biomédica e Bioacoustics, que me proporcionou os melhores colegas de curso. Um carinho especial também à ANEEB, In-Nova, e AEFCT, e a todos os colegas do associativismo que viraram amigos. É um orgulho ver crescer estas associações, agora de fora, mas nunca a grande distância. Obrigada aos colegas da dança, que entraram na minha vida mais recentemente e logo marcaram a diferença, desafiando-me e lembrando-me de que quem dança é mais feliz.

À Mawecas, um agradecimento pela amizade pura, sem fronteiras e condições. Independentemente de para onde voes (sempre mais alto), sei que nos voltaremos sempre a encontrar no ManelBar, com conforto e mocaccinos. Gandini e William, nunca antecipei que nos juntaríamos como Vaticanum e não poderia desejar melhores colegas de casa. Todos os momentos que partilhámos e partilharemos serão sempre lembrados com carinho. Acrescentando a Vera, agradeço-vos por me motivarem sempre e por estarem ao meu lado nestes últimos anos, especialmente nos momentos mais difíceis. Correia, Afonso, Sara, Diogo, obrigada por serem pessoas incríveis. Sou uma sortuda pelos amigos que tenho.

À minha família, o meu agradecimento mais profundo. Aos meus avós, à Titizocas, aos meus pais, pelo incansável esforço que fizeram para nos dar o melhor possível e por serem compreensivos nesta etapa, que envolveu tantos momentos em que precisei de mimo, quanto momentos em que precisei de espaço. Ao meu irmão, Guizocos, pelo orgulho que me dás ao ver-te crescer e tornar num jovem cada vez mais bondoso e por me inspirares a ser melhor. E aos meus queridos gatinhos, Mia, Zoquinha e Simba, e cãozinho Buddyzecos, pela fonte inesgotável de amor e carinho que sempre serão, onde quer que estejam.

Ao António, um obrigada do tamanho do mundo. Seria impossível chegar a este ponto sem o teu apoio incansável, carinho sem medida, e a fé que depositas em mim. *May we continue being dorky together.*

Muito obrigada a todos por fazerem parte desta jornada, por acreditarem em mim, e por tornarem esta realização possível.

**Espero que todas as pessoas aqui mencionadas (e todas as pessoas, na verdade) façam uma pausa de 10/20 minutos e bebam 1 café se ficarem com sono enquanto conduzem.**

Com profunda gratidão,

Telma

*“I love deadlines. I love the whooshing noise they make as they go by.” (Douglas Adams)*

## ABSTRACT

Driver drowsiness has been identified as a major contributor to road accidents, accounting for up to 20% of all incidents, often resulting in fatalities and severe injuries. Even though detecting and mitigating it is crucial to improving road safety, existing datasets for drowsiness detection are limited in complexity. This dissertation addresses this critical need by undertaking two distinct processes.

In the primary development process, machine learning techniques are employed to analyze face video data, with a particular focus on the eye region as an indicator of drowsiness. The secondary process centers on the proposal of a comprehensive data acquisition protocol, capturing multimodal data streams coupled with personal and idiosyncratic factors.

The eye-state classifier achieved an accuracy of 73.75%, and a manual eye-state labeling approach led to a binary classification model of drowsiness with 80.00% accuracy and 100.00% recall for the drowsy class.

In conclusion, this work intended to contribute to future studies on effective drowsy driving alert systems based on eye-state. To further enhance this research, increasing the dataset size and subject variability is key. As to new research directions, expanding into interpretability and causality studies based on the proposed data acquisition protocol holds promise for advancing road safety initiatives.

**Keywords:** Drowsiness, Drowsiness detection, Machine learning, Eye-state classification, Road safety.

## RESUMO

A sonolência dos condutores foi identificada como um dos principais factores que contribuem para acidentes rodoviários, representando até 20% de todos os incidentes, resultando frequentemente em mortes e ferimentos graves. Embora a sua deteção e atenuação sejam cruciais para melhorar a segurança rodoviária, os dados existentes para a deteção da sonolência são limitados em termos de complexidade. Esta dissertação aborda esta necessidade crítica através de dois processos distintos.

No processo de desenvolvimento primário, são utilizadas técnicas de aprendizagem automática para analisar dados de vídeo de rostos, com especial incidência na região dos olhos como indicador de sonolência. O processo secundário centra-se na proposta de um protocolo de aquisição de dados, capturando fluxos de dados multimodais associados a factores pessoais e idiossincráticos.

O classificador do estado ocular alcançou uma precisão de 73,75% e uma abordagem manual de classificação do estado ocular levou a uma classificação binária de sonolência com 80,00% de exatidão e 100,00% de revocação para a classe sonolenta.

Em conclusão, este trabalho pretende contribuir para estudos futuros sobre sistemas eficazes de alerta de condução sonolenta baseados no estado ocular. Para melhorar esta investigação, é fundamental aumentar o tamanho do conjunto de dados e a variabilidade dos sujeitos. Quanto a próximos passos, a expansão para estudos de interpretabilidade e causalidade baseados no protocolo de aquisição de dados proposto é promissor para o avanço das iniciativas de segurança rodoviária.

**Palavras-chave:** Sonolência, Deteção de sonolência, *Machine learning*, Classificação de estado ocular, Segurança rodoviária.

---

The research work described in this dissertation was carried out in accordance with the norms established in the ethics code of Universidade Nova de Lisboa. The work described and the material presented in this dissertation, with the exceptions clearly indicated, constitute original work carried out by the author.

# CONTENTS

<b>List of Figures</b>	<b>xiv</b>
<b>List of Tables</b>	<b>xvi</b>
<b>Abbreviations</b>	<b>xvii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Context and Motivation . . . . .	1
1.2 Objectives . . . . .	3
1.3 Contributions . . . . .	4
1.4 Dissertation Outline . . . . .	4
<b>2 Literature Review and Theoretical Framework</b>	<b>6</b>
2.1 Drowsiness, the State . . . . .	6
2.2 Evaluating Drowsiness . . . . .	7
2.3 Sensoring Drowsiness . . . . .	8
2.4 Detecting Drowsiness . . . . .	13
2.4.1 Machine Learning . . . . .	13
2.4.2 Prior Art Approaches . . . . .	13
2.4.3 Solutions on the Market . . . . .	13
2.4.4 Challenges . . . . .	14
2.5 Methods Overview . . . . .	15
2.6 Performance Assessment . . . . .	16
2.6.1 Accuracy . . . . .	16
2.6.2 Sensitivity (Recall) . . . . .	16
2.6.3 Precision . . . . .	16
<b>3 Data</b>	<b>17</b>
3.1 Driving Data Essencials . . . . .	17
3.2 Available Databases . . . . .	18
3.3 The <i>DROZY Database</i> . . . . .	19
3.3.1 Database Description . . . . .	19
3.3.2 Database Preparation . . . . .	20

## CONTENTS

---

3.4	The <i>Closed Eyes in the Wild (CEW) Dataset</i> . . . . .	21
3.5	Experience Design . . . . .	22
<b>4</b>	<b>Developing a Face-Based Driver Drowsiness Monitoring System</b>	<b>26</b>
4.1	Face-landmark Study . . . . .	26
4.2	Eye-Region Extraction . . . . .	28
4.3	Feature Engineering for Eye-State Classification . . . . .	28
4.4	Eye-State Classification . . . . .	29
4.5	Feature Engineering for Drowsiness-Level Classification . . . . .	31
4.6	Drowsiness Classification . . . . .	32
<b>5</b>	<b>Results and Discussion</b>	<b>34</b>
5.1	Face-landmark Study and Eye ROI Extraction . . . . .	34
5.2	Eye-state classification . . . . .	38
5.2.1	The Challenge of Ambiguous Frames and Human-like Dilemma	39
5.3	Drowsiness classification . . . . .	41
<b>6</b>	<b>Conclusion</b>	<b>43</b>
6.1	Conclusions . . . . .	43
6.2	Future Work . . . . .	44
	<b>Bibliography</b>	<b>46</b>
	<b>Appendices</b>	
<b>A</b>	<b>The <i>AUTOMOTIVE Project</i></b>	<b>59</b>
A.1	Introduction . . . . .	59
A.1.1	CardioID Technologies Lda . . . . .	60
A.1.2	INESC TEC's VCMi Group . . . . .	61
A.2	Data Acquisition . . . . .	61
A.2.1	Sensors . . . . .	62
A.2.2	<i>AUTOMOTIVE DD Simulator</i> . . . . .	63
A.2.3	<i>AUTOMOTIVE IHC Simulator</i> . . . . .	64
A.3	Supervised Internships . . . . .	65
A.3.1	Development of a Web-Application for Real-Time Drowsiness Monitoring . . . . .	66
A.3.2	A DBN Approach for Driver Drowsiness Detection . . . . .	66
<b>B</b>	<b><i>AUTOMOTIVE Acquisition Mapping</i></b>	<b>70</b>
<b>C</b>	<b><i>AUTOMOTIVE Acquisition Briefing</i></b>	<b>72</b>
<b>D</b>	<b>Prior Art Approaches</b>	<b>79</b>

---

<b>E</b>	<b><i>DROZY Database Data Study</i></b>	<b>84</b>
<b>F</b>	<b>Landmark error table</b>	<b>86</b>
<b>G</b>	<b>More on the electrocardiogram (ECG)</b>	<b>88</b>
<b>Annexes</b>		
<b>I</b>	<b>Questionnaires for the AUTOMOTIVE Acquisition Protocol</b>	<b>91</b>
I.1	Circadian Rythm . . . . .	92
I.2	Epworth Sleepiness Scale . . . . .	98
I.3	Immersive Tendencies Questionnaire . . . . .	99
I.4	MSSQ-Short . . . . .	103
I.5	Subject Characterization . . . . .	105
I.6	System Usability Scale . . . . .	108

## LIST OF FIGURES

2.1	Different sensors for driver drowsiness detection . . . . .	8
2.2	Summary of measurement approaches used for driver drowsiness detection	9
2.3	Example of a 5-landmark placement. . . . .	15
2.4	Example of a 68-landmark placement. . . . .	15
2.5	Example of a 98-landmark placement. . . . .	15
2.6	Landmark placement examples. . . . .	15
5.1	Histogram of error distribution for the 4 algorithm combinations tested against manually labeled landmarks. . . . .	35
5.2	FaceBoxes-MobileNet. . . . .	36
5.3	RetinaFace-MobileNet. . . . .	36
5.4	FaceBoxes-MobileFaceNet. . . . .	36
5.5	RetinaFace-MobileFaceNet. . . . .	36
5.6	Error representation in an example frame for the 4 algorithm combinations tested. . . . .	36
5.7	MTCNN face detection in a sample frame. . . . .	37
5.8	Open and closed eye images before and after preprocessing, with the respective LBP image transformation and feature histograms. . . . .	38
5.9	Histogram of accuracy distribution for the two algorithms, SVC and KNN, under a grid search. . . . .	39
5.10	An example of a "dubious"eye that gathered indecision among humans. . .	40
5.11	Gini feature importance analysis of the binary drowsiness classification of each used featured. . . . .	42
A.1	A holistic view of the AUTOMOTIVE Project . . . . .	60
A.2	Running simulation in the AUTOMOTIVE DD Simulator. . . . .	64
A.3	Experimental setup of the AUTOMOTIVE DD Simulator during use. . . . .	64
A.4	Running simulation on different scenarios in the AUTOMOTIVE IHC Simulator. . . . .	67
A.5	Design of the experimental setup of AUTOMOTIVE IHC Simulator. . . . .	68
A.6	Extraction of eye aspect ratio (EAR) and mouth landmarks performed. . . . .	68
A.7	An example of incorrect placement of landmarks in real-time, both in a frame retrieved from the DROZY database and in a frame retrieved from the webcam. . . . .	69

G.1	The heart . . . . .	89
G.2	Electrical conduction events in the heart . . . . .	89
G.3	Schematic representation of the placement of eletrodes in an invasive methodology as well as the resulting leads . . . . .	90

## LIST OF TABLES

1.1 Levels of driving automation with respective driving support features. . . .	3
2.1 References for self-assessing your score in the KSS. . . . .	8
2.2 Frequency bands of heart rate variability (HRV) . . . . .	11
2.3 Summary of the open issues to be considered in drowsiness detection techniques . . . . .	14
3.1 Main databases currently available for research in biometrics and drowsiness monitoring. . . . .	18
3.2 DROZY Database Contents. . . . .	20
3.3 Relative frequency of Karolinska Sleepiness Scale (KSS) scores and KSS groups.	22
5.1 Algorithm processing speed of each combination of face detection and landmark prediction . . . . .	37
5.2 Confusion matrix of the eye-state classification. . . . .	40
5.3 Performance metrics of the eye-state classification. . . . .	40
5.4 Confusion matrix of the 3-class drowsiness classification. . . . .	41
5.5 Performance metrics of the 3-class drowsiness classification. . . . .	41
5.6 Confusion matrix of the binary drowsiness classification. . . . .	41
5.7 Performance metrics of the binary drowsiness classification. . . . .	42
E.1 Summary information of each sample. . . . .	85
F.1 Minimum, mean, and maximum distribution of error (by Euclidian distance of pixels), by algorithms applied, over all samples. . . . .	86
F.2 Minimum, mean, and maximum distribution of error (by Euclidian distance of pixels), by facial subsection and algorithms applied, over all samples. . .	87

## ABBREVIATIONS

<b>ADAS</b>	advanced driver-assistance systems 60, 62
<b>AECS</b>	average eye closure speed 11
<b>AMQP</b>	advanced message queuing protocol 62
<b>ANN</b>	artificial neural network 13, 81
<b>ANS</b>	autonomous nervous system 10
<b>API</b>	application programming interface 62
<b>AUTOMOTIVE</b>	AUTOMatic multiMODal drowsiness detecTION for smart VEhicles 1, 2, 3, 4, 22, 24, 59, 60, 61, 62, 64
<b>avg</b>	average 81
<b>AVR</b>	amplitude-velocity ratio 7
<b>BAC</b>	blood alcohol concentration 2
<b>CardioID</b>	CardioID Technologies Lda 2, 59, 60, 61
<b>CEW</b>	Closed Eyes in the Wild xii, 21, 29
<b>CTM</b>	Centre for Telecommunications and Multimedia 61, 65
<b>DD</b>	Desk-Driver 61, 64
<b>DDAW</b>	driver drowsiness and attention warning 2
<b>EAR</b>	eye aspect ratio xiv, 11, 66, 68, 80
<b>ECG</b>	electrocardiogram xiii, 9, 10, 11, 13, 18, 20, 44, 60, 62, 88, 89, 90
<b>EDA</b>	electrodermal Activity 9, 13
<b>EEG</b>	electroencephalography 7, 9, 20
<b>EMG</b>	electromyography 9, 20
<b>EOG</b>	electrooculography 9, 20
<b>ERDF</b>	the European Regional Development Fund 59
<b>ESS</b>	Epworth Sleepiness Scale 7, 23, 24, 91
<b>FCT</b>	Fundação para a Ciência e a Tecnologia 59

## ABBREVIATIONS

---

<b>HF</b>	high frequency 11
<b>HRV</b>	heart rate variability xvi, 10, 11
<b>IHC</b>	In-Half-Car 22, 61, 64, 67, 68
<b>INESC TEC</b>	the Institute for Systems and Computer Engineering, Technology and Science 2, 4, 59, 60, 61
<b>ISEL</b>	the Instituto Superior de Engenharia de Lisboa 2, 59
<b>ITQ</b>	Immersive Tendencies Questionnaire 23, 24, 91
<b>JDS</b>	Johns Drowsiness Scale 7
<b>KNN</b>	K-Nearest Neighbors 30, 39
<b>KSS</b>	Karolinska Sleepiness Scale xvi, 7, 8, 19, 20, 21, 22, 23, 84, 85
<b>LBP</b>	local binary pattern 28, 29, 38, 39, 43, 82, 83
<b>LF</b>	low frequency 11
<b>MEQ</b>	Morningness-Eveningness Questionnaire 23, 91
<b>MQTT</b>	Message Queuing Telemetry Transport 62
<b>MSSQ</b>	Motion Sickness Susceptibility Questionnaire 22, 24, 91
<b>MTCNN</b>	Multi-task Cascaded Convolutional Neural Network 15, 27, 28
<b>NHTSA</b>	National Highway Traffic Safety Administration 2
<b>NIR</b>	near-infrared 20
<b>NPC</b>	non-playable character 63
<b>OSS</b>	Objective Sleepiness Scale 7
<b>PERCLOS</b>	percentage of eye closure 7, 11, 12, 13, 32, 43, 81, 82
<b>PPG</b>	photoplethysmogram 62
<b>PSG</b>	polysomnography 20
<b>PVT</b>	psychomotor vigilance test 19, 20, 21
<b>REM</b>	rapid eye movement 6
<b>ROI</b>	region of interest 43
<b>SAE</b>	Society of Automotive Engineers 2
<b>SDK</b>	software development kit 62
<b>SDLP</b>	standard deviation of lane position 8, 9
<b>SEM</b>	slow eye moments 11

<b>SSS</b>	Stanford Sleepiness Scale 7
<b>std</b>	standard deviation 81
<b>SUS</b>	System Usability Scale 23, 24, 91
<b>SVM</b>	Support Vector Machine 30, 39, 40, 43, 82, 83
<b>SWA</b>	Steering Wheel Rotation Angle 8, 14, 44
<b>ULF</b>	Ultra Low Frequency 11
<b>ULHT</b>	the Lusophone University of Humanities and Technologies 2, 59, 64, 65
<b>VAS</b>	Visual Analog Scale 7
<b>VCMI</b>	Visual Computing and Machine Intelligence 60, 61
<b>VLf</b>	Very Low Frequency 11
<b>WHO</b>	World Health Organization 1



## INTRODUCTION

This dissertation starts with an exposition of the context and motivation behind it. Then, this introductory chapter goes on to unveil this work's core objectives, breaking them down into more manageable sub-objectives, and their alignment with the *AUTOMATIC multiModal drowsiness detecTion for smart VEhicles (AUTOMOTIVE) Project* - during which this work started. Afterward, the structure of this dissertation is outlined, providing a roadmap for what lies ahead, and this chapter concludes by presenting the contributions made by this work.

### 1.1 Context and Motivation

Drowsiness, further described in chapter 2, essentially translates into an intermediate state between wakefulness and sleep [1]. This condition can range from inopportune to dangerous and life-threatening if the person suffering it simultaneously engages in an activity requiring alertness, such as driving.

In 2018, the World Health Organization (WHO) reported that yearly "deaths from road traffic crashes [had] increased to 1.35 million" worldwide, stating road traffic injury as the leading cause of death for people with ages ranging from 5 to 29 years old [2]. Evidence suggests that driver fatigue and sleepiness can contribute to up to 20% of all road accidents [3], while statistics show that these crashes result in death or severe injuries about 50% more times than the average road traffic crash [4], possibly because they tend to be high-speed impacts which a drowsy driver cannot suitably refrain [5].

In light of this, drowsy driving can be claimed as one of the most common preventable causes of death on the roads [6], [7]. Hence, it becomes relevant to develop methods that can continuously track the driver for the entire driving session, looking for drowsiness indicators. Tools such as these can be particularly useful when applied to truck

drivers, shift workers, and, among others, anyone driving a long-distance route [5], even if healthy. In fact, it is documented that healthy but sleep-deprived individuals represent a higher threat to traffic safety than drivers with sleep disorders, as it takes just 17 hours of prolonged wakefulness to decrease the driving performance of a healthy person in an identical way to the effect of a 0.05% (0.5 g/l) blood alcohol concentration (BAC), the legal limit to drive in numerous jurisdictions [6].

With increasing advances in automotive technology, the world has been introduced to more intelligent and complex vehicles. This development must cover the safety topic at hand by progressively incorporating drowsiness detection and monitoring methods in the industry's products. In fact, Regulation (EU) 2019/2144 of the European Parliament and Council requires automobiles (from July 6, 2022 for new types, and July 7, 2023 for all new vehicles) to be equipped with driver drowsiness and attention warning (DDAW) systems.

Considering that human error is responsible for at least 90% of road accidents [8], there is hope that security will be increased with the eventual transition to fully-automated vehicles. However, to understand when such advances will become widely available, it is vital to comprehend the stages leading to it. The Society of Automotive Engineers (SAE)'s J3016 standard defines six levels for driving automation, hereby compiled in Table 1.1, ranging from level 0 to level 5, the maximum predicted evolutionary state. Of these, only in level 5 can the vehicle's occupants be entirely passengers, as it does not need a driver to take over command in any situation [9]. This means that only at this stage in autonomous driving can drowsy driving cease to be a serious safety hazard. Since the United States Department of Transportation's National Highway Traffic Safety Administration (NHTSA) believes such advances to only become available after 2025 [10], it is reasonable to expect that there is still a long road ahead before fully autonomous vehicles become accessible worldwide. At least until then, the driver will play a critical part in road safety, and works such as the present and the globality of the *AUTOMOTIVE Project*, in which the present work was integrated, hope to contribute to the study and development of driver monitoring solutions. The *AUTOMOTIVE Project*, further presented in Appendix A, constituted a research partnership between the Institute for Systems and Computer Engineering, Technology and Science (INESC TEC), CardioID Technologies Lda (CardioID), the Instituto Superior de Engenharia de Lisboa (ISEL), and the Lusophone University of Humanities and Technologies (ULHT), and its main aim was to investigate computer vision and deep learning methodologies for the automatic detection of fatigue states in drivers using multiple sensory information. To this endeavour, having an approach for drowsiness detection based on the driver's face video was important, as it had not been yet done in the context of the project, and as it is, as clarified in the following chapter, one of the main data sources used in the Literature to address the driver drowsiness problem.

Table 1.1: Levels of driving automation with respective driving support features, adapted from [9]. Level 5 fully-autonomous vehicles are only expected to get in the market after 2025.

	User's behavior while in the driver's seat	Driving Support Features	Feature Examples
Level 0	Its always considered that the user is driving, whenever the support features are engaged. Moreover, constant supervision of these support features is also required for safe driving.	Providing warnings and momentary assistance	Automatic emergency braking; Blind spot warning; Lane departure warning
Level 1		Providing steering OR brake/acceleration support to the driver	Lane centering OR adaptive cruise control
Level 2		Providing steering AND brake/acceleration support to the driver	Lane centering AND adaptive cruise control at the same time
Level 3	The user is not actually driving when these features are engaged. Nonetheless, when the features do request, the user must drive.	Can drive the vehicle under limited conditions and will not operate unless some required conditions are met	Traffic jam chauffeur Local driverless taxi; Pedals/steering wheel may or may not be installed
Level 4	The user is not actually driving when these features are engaged, and these automated features will not require the user to take over driving.		
Level 5		The features can drive the vehicle under all conditions.	

## 1.2 Objectives

The present work was carried out with **two central objectives** at heart:

1. to propose a functional method to detect and monitor drowsiness in automobile vehicle drivers using a seemingly non-intrusive setup;
2. to address the current data gap by proposing an experimental design and setup for data acquisition in a simulated driving experience.

**Objective number 1** was decomposed in the following steps:

1. Reviewing the Literature regarding the field of driver drowsiness;
2. Implementing a baseline working approach focused on real-time face analysis;
3. Choosing and extracting relevant features for a machine learning, primarily face-based drowsiness classification solution;
4. Implementing the algorithm, evaluating its performance, and performing refinements.

**Objective number 2** was decomposed in the following steps:

1. Investigating the available databases and identifying points of improvement;
2. Understanding the *AUTOMOTIVE Project's* envisioned set-up of an immersive car simulator for drowsiness studies; the workings of the Unity simulation and data transmission flows already developed; and the available sensors for acquisition;

3. Exploring other ways to enrich and complement the data to be retrieved from sensors;
4. Designing, mapping, and describing the acquisition experience from start to finish.

### 1.3 Contributions

The present dissertation continued the works conducted in the *AUTOMOTIVE Project*. Whilst the previous works [11]–[14] focused on different combinations of signals, analysis, datasets, and overall purposes, the current work delivered contributions on the driver’s face video analysis, providing a different and valuable framework to monitor driver drowsiness in this project and overall.

Since one of the *AUTOMOTIVE Project*’s objectives was to develop an environment to simulate a real driving experience, during this work, participation in the installation of the simulation vehicle set-up also took place.

Since real driving data availability is scarce (especially labeled and accounting for added conditionalism to driving - such as drowsiness or health issues), by manually labeling data and encompassing the development and description of a data acquisition protocol tailored to the *AUTOMOTIVE Project*’s objectives and conditions, featuring a multimodal approach considering physiological signals, driver behaviour, vehicle behaviour, and individual characteristics, this work also contributed to the driver drowsiness data acquisition in this project and overall.

Specifically in the INESC TEC’s environment, this work also involved participation in the supervision of two summer internships.

This work also resulted in an article accepted in the *IEEE Access* journal [15], contributing to open access research and exposing a comprehensive view of the research and development work conducted for the *AUTOMOTIVE Project*.

### 1.4 Dissertation Outline

The purpose of the current chapter was to explain the context in which the current work unfolds, as well as its core objectives. Following it, four chapters constitute the main body of this dissertation:

- Chapter 2, "Literature Review and Theoretical Framework", will further expand on the context of this work by explaining concepts that are pre-required to understand the remaining dissertation and give an overview of the more current developments made in the scientific literature in the interest of this work.
- Chapter 3, "Data", explains the context around using and accessing data for driving research, the databases used in the current work, and the endeavor made to promote more conscious and insightful research through a new data acquisition protocol.

- Chapter 4, "Developing a Face-Based Driver Drowsiness Monitoring System", presents the computer vision and machine learning techniques explored in this work, as well as their fitting to the DROZY database, described in the previous chapter.
- Chapter 5, "Results and Discussion", presents and discusses the results retrieved over the course of the current work.

After the main body, chapter 6, "Conclusion", will summarize the conclusions of the current work and point suggestions for future lines of research around driver drowsiness detection, as well as possible expansions of the practiced methodology.

In the end, appendices and annexes are provided as support material.

## LITERATURE REVIEW AND THEORETICAL FRAMEWORK

This chapter serves as a robust knowledge base, delving into the theoretical concepts necessary to address this research problem. It commences by exploring key concepts pertaining to drowsiness, subsequently presenting the realm of sensor technologies and detection methods. This chapter also underscores the current challenges that shape this critical problem.

### 2.1 Drowsiness, the State

Normal human sleep is composed of a rapid eye movement (REM) stage and four non-REM stages. These stages have well-defined characteristics and alternate cyclically, being a standard human adult cycle of approximately 90 minutes, commencing in a non-REM stage and finishing in REM sleep [16], [17]. According to Carskadon and Dement [18], “sleep is a reversible behavioral state of perceptual disengagement from and unresponsiveness to the environment”.

The term “drowsiness”, in Literature often used interchangeably with the terms “fatigue” and “sleepiness”, relates to a physiological need to sleep, as it is the intermediate state between wakefulness and sleep [1]. This state comes with the impairment of visual perception, of higher cognitive functions, the inability to maintain visually focused attention, among other undesirable consequences for safe driving [19].

Since physiological needs are biological requirements that cannot be continuously avoided, a drowsy state is beholden right before a sleeping episode occurs [1]. This is a state that, if met, means that the body is already fighting to remain awake, possibly indulging in microsleeps in the process [20]. Microsleeps have been associated with

poor driving performance [21] and, during these episodes, attention gaps can impair the driver's ability to respond to events [22].

"Fatigue", on the other hand, relates to a state where the subject feels tired, and this often comes from the prolonged engagement in physical activity and can be relieved once the subject ceases said activity [1].

"Sleepiness", unlike "drowsiness" and "fatigue", can be considered not a state but a probability of falling asleep at a given time [1].

This work specifically aims to detect and monitor drowsiness.

## 2.2 Evaluating Drowsiness

Studies reported great inter-variability in how drowsiness affects drivers' performance [23], being that even for a given self-declared drowsiness level, indicators such as eye blink duration vary considerably [24]. Hence, objectively and non-intrusively measuring drowsiness has been a constant challenge in driver drowsiness research.

On one hand, there are approaches to more objectively measure drowsiness, such as the Objective Sleepiness Scale (OSS), which combines features of electroencephalography (EEG) and eye movements, such as percentage of eye closure (PERCLOS), to characterize four sleep scoring stages [25]. This approach still presents some disadvantages such as an insufficient number of stages [25] and, mainly, its intrusiveness. The Johns Drowsiness Scale (JDS) stands on the continuous tracking of the amplitude-velocity ratio (AVR) of blinks through the company Optalert's<sup>1</sup> drowsiness detection products, such as the Optalert glasses, a wearable that measures the subject's eyelid velocity by infrared oculography, using a LED built into the glasses frame [26]. The AVRs are very stable between people, which means they do not have to be adjusted for individual subjects [27]. These continuous measurements are then translated into the patented 10-point JDS, which can hence form the basis of a new objective method for continuously monitoring the driver's drowsiness state, through the company's products.

Nevertheless, some subjective scales are also used to score drowsiness, by residing in the subjects' response concerning standardized sleep symptoms. Some examples include the Epworth Sleepiness Scale (ESS) [28], the Stanford Sleepiness Scale (SSS) [29], the Visual Analog Scale (VAS) [30], and the Karolinska Sleepiness Scale (KSS) [31].

The KSS, used in the current work, is a 9-point scale that defines verbal reference for each stage, as shown in Table 2.1 [31]. The ground-truth for the drowsiness score is assessed by having the subject periodically report their perceived score. Besides not being an objective measurement, this scale can cause the driver's state to be influenced since they are periodically disturbed [24], [25]. However, positive correlation can be seen between this subjective measurement, and more objective measurements [32], which

---

<sup>1</sup><https://www.optalert.com/>

presents the KSS as a very easy, free-to-use, and non-intrusive way to score drowsiness in research, as long as the researches are aware of its limitations and properly disclose them.

Table 2.1: References for self-assessing your score in the KSS.

KSS score	Verbal reference
1	Extremely alert.
2	Very alert.
3	Alert.
4	Rather alert.
5	Neither alert nor sleepy.
6	Some signs of sleepiness.
7	Sleepy, but no effort to keep awake.
8	Sleepy, some effort to keep awake.
9	Very sleepy, great effort to keep awake, fighting sleep.

### 2.3 Sensoring Drowsiness

Motivated by the severeness of the driver drowsiness problem, nowadays there is a substantial amount of research in detection techniques. The most commonly used in-vehicle measurement sensors for the detection of driver drowsiness can be grouped, according to [33], into vehicle behavioral sensors, driver behavioral sensors, and physiological signals sensors. Figure 2.1 illustrates these three forms of sensing, and Figure 2.2 graphically summarizes the information acquired in each.

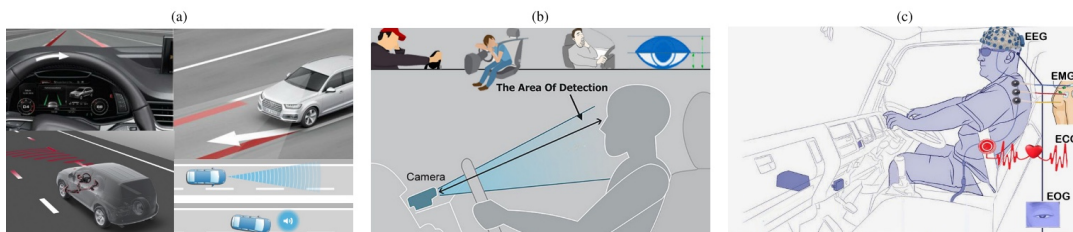


Figure 2.1: Different sensors for driver drowsiness detection. Illustrations, from [33], of (a) vehicle behavioural sensors, (b) driver behavioural sensors, and (c) physiological signals sensors

The most commonly adopted vehicle behavioural sensors in the industry are used to obtain the Steering Wheel Rotation Angle (SWA) and the standard deviation of lane position (SDLP) [33]. The SWA is obtained through an angle sensor mounted on the steering axis

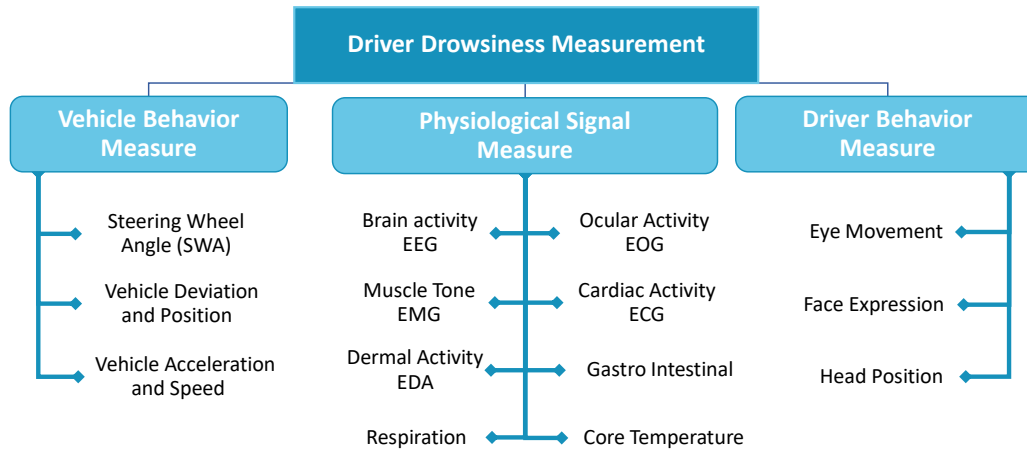


Figure 2.2: Summary of measurement approaches used for driver drowsiness detection. Adapted of [25]

and can measure abnormal driving behavior through metrics like the number of steering wheel reversals [34]. The SDLP measures the vehicle's position when compared to the road's middle lane and is one of the metrics used in the position and deviation of the vehicle approach [13], [35], others being the deviation from the lateral lane [36], the yaw deviation [37] and the heading difference [38]. Vehicle speed, acceleration, and braking are other variables used to assess the driver's alertness through their driving performance [33].

The driver behavior, on the other hand, is mainly monitored through face video [33], from which we can retrieve a lot of information regarding the eye behavior, but also general face expression and posture. These vision-based methods have the advantage of being totally non-invasive, although the idea of having one's face being recorded might induce some worry among users concerned with their privacy.

Physiological signal's sensors include sensors to obtain the electrocardiogram (ECG) and electrooculography (EOG), electromyography (EMG), electrodermal Activity (EDA), and EEG electrophysiological signals. EOG measures the ocular activity, EMG the muscle tone, and EEG the brain activity. The ECG, usually a graph of micro voltage over time, displays the electrical activity of the heart, by measuring small electric changes created by the heart muscle tissue when its chambers are contracting and relaxing, during each cardiac cycle. Other physiological measures that can be obtained are respiration related metrics, gastrointestinal parameters, and temperature [33].

Overall, it can be concluded that each source presents its unique advantages and challenges, warranting careful consideration in the quest for an optimal solution:

#### Physiological signals:

- *Advantage:* Physiological signals, such as EEG, offer unparalleled objectivity in

assessing the driver's state, particularly in terms of alertness and sleep patterns. They are robust to most environmental conditions.

- *Challenge:* These signals are typically acquired intrusively, often requiring visible wires or sensors on the driver. This poses logistical challenges for experiments and discomfort issues for long-term deployment.

**Driver behavior:**

- *Advantage:* Observing driver behavior through cameras provides a non-intrusive means of data collection.
- *Challenge:* However, this method raises concerns about user comfort, especially over extended periods or in market-ready solutions where privacy and constant surveillance may be unsettling to users. Besides, environmental conditions such as low lighting (as in night driving, when a subject can be more susceptible to drowsiness) or the use of glasses can have a great impact.

**Vehicle behavior:**

- *Advantage:* Monitoring vehicle behavior is non-intrusive and doesn't impact the user's comfort.
- *Challenge:* This approach primarily detects impairment after it has occurred, potentially limiting its capacity for anticipatory interventions. Additionally, environmental conditions such as road dampness and adverse weather can impact the driving performance and, hence, the metrics retrieved.

Recognizing the limitations of individual data streams, a multimodal approach emerges as an ideal solution, whether for experimental purposes or long-term deployment: combining physiological signals, driver behavior, and vehicle behavior data acquired non-intrusively may offer a holistic view of driver drowsiness, addressing both comfort and data comprehensiveness concerns.

While the algorithm development part of the current work primarily focuses on the information obtained from face video, for the data acquisition design ECG and vehicle behavioral measurements are also employed, comprising data from the 3 forms of sensing presented (vehicle behavior, physiological signal, and driver behavior). The specific, non-intrusive sensing systems proposed to acquire and access the data in this experience design are presented in Appendix A.

As for the ECG, this physiological signal can be a valuable indicator of drowsiness since, during this state, the heart rate becomes slower, more irregular, and the blood pressure drops [39]. This can be measured by calculating the heart rate variability (HRV). The HRV is obtained through an R-peak detection algorithm of the ECG. Then, we can obtain information regarding the autonomous nervous system (ANS) activity of the subject,

which is known to present alterations during drowsiness episodes [40] - wakefulness is characterized by an increase in sympathetic activity and/or decrease in parasympathetic activity, while extreme relaxation states (as in non-REM sleep) are characterized by a decrease in sympathetic activity and/or increase in parasympathetic activity [39]–[42]. A low frequency (LF) band (0.04-0.15 Hz), is a reliable measurement of sympathetic activity, while a high frequency (HF) band (0.15-0.4 Hz) provides information about the parasympathetic activity. Hence, by calculating the ratio between lower and higher frequencies (LF/HF), information regarding the state of the individual can be obtained [43]. A more discriminate division of frequency bands is presented in Table 2.2. For the sake of conciseness, and as, although interesting and used in the experimental protocol design, it is not the main source of data in this work, more information on the ECG was arranged in Appendix G.

Table 2.2: Frequency bands of HRV

Component	Frequency Range (Hz)
Ultra Low Frequency (ULF)	$\leq 0.003$
Very Low Frequency (VLF)	0.003 - 0.04
LF	0.04 - 0.15
HF	0.15 - 0.4

Following the aforementioned driver behavior monitoring, having face recordings as the main data source for drowsiness detection can prove advantageous due to non-invasiveness and the ability to predict drowsiness before driving impairment occurs [44].

In the face, symptoms of drowsiness can be visible in terms of eye movement (in the rate of eye blinking [45], eye aspect ratio (EAR) [46], [47], slow eye moments (SEM) [48], the PERCLOS [49], and the average eye closure speed (AECS) [50]), face expression (for example in yawning [51]), and head position (for example head nodding [52]) [33].

Regarding eye features, insights from various studies on the features more consistently encountered in Literature that are also possible to assess in the current work are synthesized below:

**"Blink rate"** (also denoted as **"Blink frequency"**), typically represented as number of eye blinks per minute:

- According to Monster [53], multiple factors can influence the blink rate, resulting in an average range of 2 to 50 blinks per minute.
- Factors affecting the blink rate include:
  - Task focus and cognitive processes (decreased blink rate when the subject is more focused and processing information) [54], [55];
  - Dryness of the eyes, whether due to health conditions or climatic factors (increased blink rate) [56];

- Conditions such as Parkinson’s disease (decreased blink rate) [57];
- Gender discrepancies - women demonstrated a higher blink rate than men while reading [58].
- According to Johns et al. [59], alert individuals exhibit a blink rate of about 17.2, **increasing** to 27.3 when drowsy. However, according to Pasaribu et al. [47], normal individuals exhibit an average blink rate of 10, **decreasing** to below 10 when sleepy. Islam et al. [46] agree, adding that the number is 4-6 for a drowsy person.

**"Blink duration"**, consistently defined as seconds or milliseconds that a blink lasts (duration from the initiation of the eyelid closing to the point of maximum speed during its reopening [60]):

- Blink duration is one of the most reliable features for drowsiness detection [61], [62].
- The Harvard Database of Useful Biological Numbers [63], [64] identifies 0.1-0.4 seconds as the average duration for a single blink. Interestingly, Schellini et al. [65] report a slightly different duration, ranging from 0.25 to 0.75 seconds.
- In the event of drowsiness, Galley’s experiment [60] reported an **increase** in blink duration associated with lower alertness, indicating a duration of 0.175 seconds for drowsiness.

**PERCLOS**, denoting the "PERcentage of eye CLOSure", introduced in 1999 and defined as the percentage of time that the eyelids are closed (more than 80% shut) [66]:

- PERCLOS has undergone thorough validation as an indicator for identifying drowsiness across various scenarios, encompassing naturalistic driving contexts, simulated environments, and laboratory conditions [67]–[70].
- PERCLOS consistently demonstrates an increase in response to drowsiness [67], [71]–[73].

Notably, accurate assessment of eye features through video relies on the use of high-quality cameras capturing a large number of frames per minute [74].

Other interesting eye features found in the Literature include the eyelid opening and closure speeds, the blink amplitude in terms of eyelid closure, and the delay between eyelid closing and opening. Nevertheless, these features would need a higher frame rate to be extracted through video analysis than the one available for use in the current work and, as such, are not detailed in this section.

## 2.4 Detecting Drowsiness

In the domain of drowsiness detection, a plethora of methodologies has been explored and can be seen in the Literature, leveraging the diverse sensing possibilities elucidated in the preceding section. However, a pivotal turning point in this pursuit seems to have emerged with the advent of computer vision technologies (especially regarding face-based drowsiness detection) and machine learning.

### 2.4.1 Machine Learning

Machine learning is the subfield of computer science that, according to Arthur Samuel in 1959 [75], gives "computers the ability to learn without being explicitly programmed." Deep learning, a subset of machine learning [76], is "the study of artificial neural networks (ANNs) and related machine learning algorithms that contain more than one hidden layer" [77].

### 2.4.2 Prior Art Approaches

The table presented in Appendix D summarizes the literature consulted where the authors approached the drowsiness detection problem based on measures that are possible to obtain through the acquisition of the driver's face video. The methods used and their performance are displayed.

### 2.4.3 Solutions on the Market

Further concurring with the severity of the context in which the present thesis unwinds, automotive manufacturers and other companies have developed numerous solutions, targeting not only the detection of drowsiness but also general attention-lacking. In this section, some of these will be briefly presented (for a more complete present-day review see [25]).

- **CardioWheel, by CardioID Technologies** [78], continuously acquires the driver's ECG through a cover equipped with conductive elements, and through it detects drowsiness and cardiac pathologies and performs the driver's biometric identification.
- **Steer Device, by STEER, Inc.** [79], is a bracelet that detects the level of drowsiness through heart rate and EDA analysis.
- **IR-LED, by Siemens** [80], is a camera system equipped with an infrared light-emitting diode and an infrared light sensor, which help to detect microsleeep events.
- **FaceLAB Driver Safety, by Seeing Machines** [25], detects driver drowsiness in real-time through blink and PERCLOS analysis, recording face video for face, head, eye, eyelid, and gaze tracking.

- **Driver Attention Warning, by Saab** [81], uses an infrared camera to monitor eye blinks and track gaze and head orientation.
- **Driver Alert Control, by Volvo** [82], [83], uses lane departure as a detection feature, monitoring the car movements with respect to the road markings.
- **Active Safety, by Volkswagen** [84], uses the SWA and lane departure as detection features, and continuously evaluates traffic signals.
- **Attention Assist, by Mercedes-Benz** [85], uses the SWA as a detection feature (it learns the driver steering pattern at the beginning of a driving session), as well as braking and acceleration events, the duration of the session, and road conditions.
- **Driver Attention Monitor, by Lexus** [86], tracks the eye and head movements to detect whether the driver is looking forward or not, and also has an obstacle detection feature.
- **Seeing Machines** [87], uses camera-based driver and operator monitoring safety technology to prevent automotive accidents.

#### 2.4.4 Challenges

After conducting a study on both current research on driver drowsiness detection systems and marketed solutions, Doudou et al. [25] compiled the open issues in this field, regarding the three types of information that can be acquired from the driver, presented in Section A.2.1. This study also encompassed the assessment of each issue's significance to the detection accuracy, which was displayed on a 3-point scale, with 1 being the highest significance and 3 the lowest. The result is presented in Table 2.3.

Table 2.3: Summary of the open issues to be considered in drowsiness detection techniques regarding the three types of information that can be acquired from the driver, presented in Section A.2.1, and their associated level of significance, displayed on a 3-point scale, being 1 the highest significance and 3 the lowest. Adapted from [25].

Technology Issues	Vehicle Behaviour	Driver Behaviour	Physiological Signal	Significance
Use of standard objective measures	X	X	X	1
Environment conditions Influence	X	X		3
Difficulty to extract symptoms		X		2
Circadian and wake phases	X	X	X	1
Accuracy with fewer sensors			X	2
Accuracy with fewer sensors			X	1
Impact of artefacts			X	2
Real-time detection support		X	X	1

The current research addresses the significant issues outlined in the previous table mainly through the proposed data acquisition protocol. This comprehensive protocol is designed to provide a holistic perspective by encompassing the acquisition of physiological signals, thereby addressing the potential influence of environmental conditions. Furthermore, the integration of multimodal data adds an extra layer of robustness to the research approach, ensuring a comprehensive examination of driver drowsiness and its multifaceted determinants. Additionally, circadian phases are considered through the administration of questionnaires, allowing for a nuanced understanding of individual variations. As a first drowsiness detection approach, a very light method (in terms of processing power and speed) will be implemented, fit for real-time detection.

## 2.5 Methods Overview

Following are introduced some elements that will be mentioned in the development phase of the current work, in order to streamline it and promote a more concise and comprehensible exposition.

Face detection and landmark placement involve various methods to identify and locate facial features within an image or video frame. These methods can yield differing results in terms of both the number of landmarks identified (the most commonly found in Literature were 5, 68, and 98) and their specific placements, as shown in Figure 2.6. Typically, facial landmarks encompass key points like the eyes, nose, mouth, and other notable features.

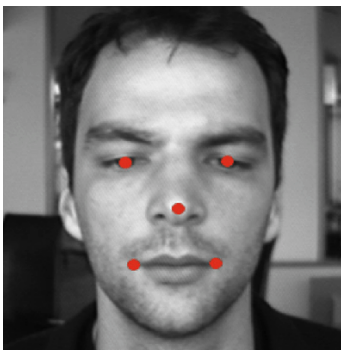


Figure 2.3: Example of a 5-landmark placement. Image retrieved from [88].

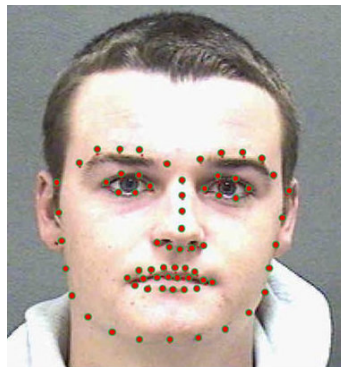


Figure 2.4: Example of a 68-landmark placement. Image retrieved from [89].

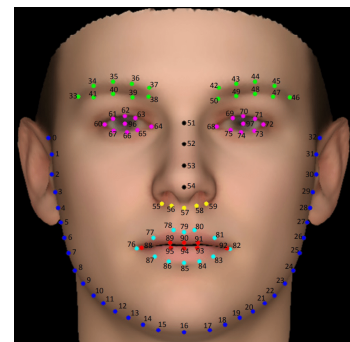


Figure 2.5: Example of a 98-landmark placement. Image retrieved from [90].

Figure 2.6: Landmark placement examples.

In this work, the Multi-task Cascaded Convolutional Neural Network (MTCNN), introduced in 2016 by Zhang [91], was employed for face detection and landmark placement. RetinaFace, introduced in 2019 by Deng [92]; FaceBoxes, introduced in 2017 by

Zhang [93]; MobileNet (V2), introduced in 2018 by Sandler [94]; and MobileFaceNet, introduced in 2018 by Chen [95], were also employed for the same purpose.

## 2.6 Performance Assessment

In machine learning, various performance metrics are used to evaluate the quality of a classification model. In this work, the commonly used metrics "accuracy", "sensitivity (recall)", and "precision", were employed.

### 2.6.1 Accuracy

Accuracy measures the overall correctness of a classification model. It is defined as the ratio of correctly predicted instances to the total number of instances:

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Instances}} \quad (2.1)$$

### 2.6.2 Sensitivity (Recall)

Sensitivity, also known as recall, quantifies the ability of a model to correctly identify positive instances (e.g., the true positives). It is defined as:

$$\text{Sensitivity (Recall)} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (2.2)$$

### 2.6.3 Precision

Precision measures the accuracy of positive predictions made by the model. It is defined as:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (2.3)$$

This chapter delves into the driving data essentials and introduces the available databases, as well as the choice for the DROZY database (a pivotal resource in this work), highlighting its attributes and subsequent treatment. The proposed data acquisition protocol is also presented here, following the AUTOMOTIVE Project’s set-up, presented in Appendix A.

### 3.1 Driving Data Essentials

Driving research is a field with numerous projects typically carried out either in simulated or naturalistic environments. In naturalistic environments, subjects get on the road to drive, following a heavy line of permits (sometimes at a government and municipal level), depending on the research objectives. This makes naturalistic data hard to find. Virtual driving environments, on the other hand, stand out due to the minimal risk to the volunteers and general population during the experiments, and the possibility of fully controlling vehicle physics and traffic complexity. As such, whether by collecting large-scale data and training deep neural networks without any logistic restraints or by emulating exact traffic circumstances to validate trained models in extreme scenarios, the use of simulators has been a key contributor to driving research, particularly autonomous driving.

Regarding the data that can be acquired in driving research, either through simulated or naturalistic environments, the most commonly used in-vehicle measurement sensors for detecting driver drowsiness can be categorized into vehicle behavior, driver behavior, and physiological signals (according to [33] and as presented in Section 2.3).

For this kind of drowsiness detection research, an ideal dataset would encompass data streams from these 3 sensing sources. However, the pursuit of an ideal dataset faces

significant hurdles, especially when it comes to acquiring a trustworthy ground truth for drowsiness, which, often (for the practicality of experience and participants' comfort), should be obtained non-intrusively.

Moreover, the complexities of human behavior and individual variability come into play. Circadian typology, for instance, has been shown to influence an individual's susceptibility to drowsiness at a certain time of day [96], introducing an additional layer of complexity. Furthermore, it is reasonable to hypothesize that personal factors such as anxiety levels, sleep quality, driving frequency, caffeine intake, familiarity with manual or automatic transmission, prior experience with the specific vehicle or road conditions, and more may all impact a person's alertness during driving. Yet, to the extent of this author's knowledge, there is currently no database that seamlessly integrates sensor data, objective ground truth information, and the intricate personal and idiosyncratic data that may contribute to a more nuanced understanding of driver drowsiness.

### 3.2 Available Databases

Following, an eye-level summary of the available databases is presented in Table 3.1. This data is either publicly available or can be provided upon request for research purposes.

Table 3.1: Main databases currently available for research in biometrics and drowsiness monitoring with face video and ECG.

Database	Data Type	Access
DROZY, The ULG Multimodality Drowsiness Database[97]	KSS, ECG, Face Video	Public
UTA Real-Life Drowsiness Dataset[98]	KSS, Face Video	Upon Request
<i>Swedish National Road and Transport Research Institute</i> [99]	<i>KSS, ECG</i>	<i>Upon Request</i>
Stress Recognition in Automobile Drivers[100]	ECG	Public
<i>E-HOL</i> [101]	<i>ECG</i>	<i>Public</i>
<i>Youtube Faces Database</i> [102]	<i>Face Video</i>	<i>Public</i>
NTHU Driver Drowsiness Detection Dataset[103]	Face Video	Upon Request

### 3.3 The DROZY Database

The DROZY (ULg Multimodality Drowsiness Database) database<sup>1</sup> [97] offers the essential elements for this study. It provides access to face video data and a well-established, although subjective, ground truth measure, the Karolinska Sleepiness Scale (KSS) (previously presented in Section 2.2. Additionally, the database’s public availability enables broader access in dissemination and avoids access waiting times.

While it is noteworthy that the videos in the DROZY database were not recorded in a conventional driving setting, they were acquired during psychomotor vigilance tests (PVTs) while the subjects generally reported increasing levels of sleepiness. This distinction is, of course, significant, but the focus here is on the shared behaviors associated with drowsiness, which can be extrapolated to the driving context. PVTs require subjects to quickly react to visual cues (which they do not know when will be activated) and, as the reaction is timed, if performed while the subjects are getting increasingly sleepy, they can provide a measure of sleepiness-related effects in alertness [104].

Given the aforementioned absence of an ideal database for driver drowsiness detection research, and that the AUTOMOTIVE database was not completely ready, this choice served as a pragmatic approach that had to be made in order to advance the research.

Please refer to reference [97] for the research paper introducing this database, which was instrumental to this work. The data was collected by the Laboratory for Signal and Image Exploitation (INTELSIG), which is part of the Department of Electrical Engineering and Computer Science of the University of Liège (ULg), Liège, Belgium.

#### 3.3.1 Database Description

Within DROZY, data was meticulously collected from a cohort of 14 youthful and healthy individuals, comprising three male and eleven female subjects. This data collection unfolded over the course of three consecutive 10-minute (approximately) sessions of PVTs. This experience induced sleep deprivation through prolonged waking, thereby creating a fairly controlled environment for the examination of drowsiness.

Each subject’s journey through these PVTs has yielded an array of synchronized data streams, summarized in Table 3.2. In this table, the data used in the work currently reported is highlighted in bold.

For data organization, sessions are identified by the subject number (ranging from 1 to 14) and whether it’s their first, second, or third PVT, as such: [#subject]-[#PVT].

Regarding the encoded videos, the database authors state that, because of backup incidents, some of the data are lost (tests 9-1, 10-2, 12-2, 12-3, 13-3), and test 7-1 did not happen. Furthermore, some frames were found to be missing during the development part of this work, including consecutive frames of tests 5-2 and 5-3, which deemed them untrustworthy samples for retrieving metrics related to eye-state over time, such as

<sup>1</sup>Available in: <http://www.drozy.ulg.ac.be/>

Table 3.2: DROZY Database Contents. Contents used in this work are highlighted in bold.

Data Type	Description
KSS Scores	<b>Subjective self-evaluated drowsiness levels according to the Karolinska Sleepiness Scale, ranging from 1 (Extremely alert) to 9 (Very sleepy, great effort to keep awake, fighting sleep), assessed at the beginning of each PVT.</b>
PVT Data	Stimulus times and corresponding reaction times, measured in milliseconds.
polysomnography (PSG) Signals	Includes 5 electroencephalography (EEG) channels (Fz, Cz, Pz, C3, and C4, all referenced to A1), 2 electrooculography (EOG) channels (horizontal and vertical), eletromyography (EMG), and the ECG, all sampled at 512 Hz, available in EDF format.
<b>Kinect v2 Sensor Videos</b>	<b>near-infrared (NIR) intensity videos, with one video available per PVT session, sized at 512x424 pixels, and grayscale values in 8 bits, preserved in .mp4 format. 20 videos were recorded at 30 FPS, while 16 at 15 FPS.</b>
Face Landmarks (Manual)	2D (in pixels) and 3D (in millimeters) annotations of 68 facial landmarks that were manually extracted for 720 hand-selected frames (approximately 51 frames per subject).
Face Landmarks (Automatic)	2D and 3D annotations of 68 facial landmarks for all frames, obtained through automatic processing using subject-specific constrained local models (CLMs).

blinking frequency. For this reason, these samples were not considered in these metrics' retrieval.

### 3.3.2 Database Preparation

To facilitate a comprehensive analysis of the video dataset, a structured approach was undertaken. The dataset comprises 36 videos, each corresponding to a distinct testing session. Key video attributes, including the sample rate (in frames per second, FPS), test duration, and frame count, were meticulously compiled and organized into a summary table for reference, that can be found in Appendix E.

In total, the dataset encompasses 5 hours, 54 minutes, and 11 seconds of footage, comprising 497 609 frames. The average duration per video test was approximately 9 minutes and 50 seconds. Notably, 16 videos were recorded at a sample rate of 15 FPS, while 20 videos were captured at 30 FPS.

Since the KSS scores were obtained at the outset of each test session, their evolution between consecutive videos was examined. It was assumed that KSS scores progressed

linearly over time, and thus, the minimum frame count required for a one-point increase in KSS was determined. This calculation yielded an approximate threshold of 1318 frames at a frame rate of 15 FPS, equivalent to approximately 1.5 minutes. To account for potential variability, an uncertainty margin of 25% was introduced, resulting in 65.9 seconds (988 frames at 15 FPS or 1976 frames at 30 FPS).

Subsequently, two distinct segments of each video were selected for analysis: the initial approximately 65 seconds, labeled with the KSS score from the beginning of the video, and the final approximately 65 seconds, labeled with the KSS score from the beginning of the following video (if available). In total, 57 samples were obtained. This process ensured that only data segments with a high degree of confidence in KSS labeling were included in the analysis further performed in the current work. This step was further deemed as important due to the already existent inter-variability in a self-assessed score. As a result, video data not encompassed within these carefully defined segments was excluded from further analysis, aligning with the imperative to retain as much data integrity as possible throughout this study.

Since the KSS comprises nine points and invites individuals to self-assess their drowsiness levels based on recollections of specific descriptors provided during its introduction, an approach was adopted to mitigate the subjectivity and get more interpretability out of the KSS scores. Three primary classes were defined, each encapsulating a range of three KSS points, thus offering a more generalized assessment of drowsiness levels. These classes were identified as "Alert", "Moderately Drowsy", and "Very Drowsy". According to Wang [105], who also employs the KSS, two levels of drowsiness may not provide sufficient support for alerting the driver before the risk of a crash becomes critical. Notably, however, that study also affirms that detection accuracy is associated with a lower number of drowsiness levels, with accuracies showing improvement when transitioning from three to two-class problems. To illustrate the distribution of the dataset across these newly established classes, a summary is presented in Table 3.3.

As the research progressed, the need arose to incorporate a metric that encapsulated the notion of "driving time" within the dataset. To achieve this, the video durations were used to create the metric "time on task" (i.e., time on the PVTs), which was extrapolated into "driving time", and paired with the KSS scores.

### 3.4 The *Closed Eyes in the Wild (CEW) Dataset*

The CEW [106] is a publicly available dataset, composed of 1192 samples of closed eyes (collected from the Internet) and 1231 samples of open eyes (selected from the Labeled Face in the Wild (LFW) database [107]).

The online repository<sup>2</sup> presents three forms of the dataset: face images with background; face images warped; and eye images only. The latter was downloaded and no

<sup>2</sup><http://parnec.nuaa.edu.cn/upload/tp1/02/db/731/template731/pages/xtan/ClosedEyeDatabases.html>

Table 3.3: Relative frequency of KSS scores and KSS groups ("classes").

KSS score	Relative frequency	KSS group	Relative frequency
1	0.00	1 (Alert)	0.29
2	0.09		
3	0.20		
4	0.06	2 (Moderately Drowsy)	0.22
5	0.02		
6	0.14		
7	0.27	3 (Very Drowsy)	0.49
8	0.17		
9	0.05		

preparation was needed before using this dataset, as its structure and labeling are straightforward - it is organized in 2 folders, one with samples of closed eyes, and the other with open eyes.

This dataset's availability, simplicity, and similarity in the images with the DROZY eye frames, were the key factors for the decision to use this dataset as training data for the eye-state classification purpose.

### 3.5 Experience Design

The data acquisition protocol for the current thesis work and, consecutively, for the *AUTOMATIC multiMODal drowsiness detectIOn for smart VEHICLES (AUTOMOTIVE) Project*, was influenced by the work of Naurois *et al.* [24], and prepared with the AUTOMOTIVE In-Half-Car (IHC) Simulator and sensors in mind - please find the description of this work enclosed in Appendix A.

The experience design addresses the need for robust and complete data. Hence, it was expected to collect data regarding vehicle movement (steering wheel angle, vehicle speed, acceleration, and braking), the driver's behavior (facial expressions and features), physiology (their electrocardiogram), and their sleepiness state, as well as additional information that can be useful to understand the subjects and their driving experience further.

Appendix B shows the entire flowchart representing the process, while Appendix C presents the briefing document produced to ensure a consistent explanation of the protocol to participants in the experiment.

Before the acquisition, the volunteer would be checked for a valid driver's license. They must also not be diagnosed with epilepsy nor be susceptible to kinocytosis (the short form of the Motion Sickness Susceptibility Questionnaire (MSSQ) [108] is to be

applied for this purpose). Furthermore, their susceptibility to sleepiness is to be assessed by applying the Epworth Sleepiness Scale (ESS)<sup>©</sup> [28], and the Morningness-Eveningness Questionnaire (MEQ) is to be filled by the subject, to determine their circadian rhythm's relation to the different times of the day. Other information is to be collected, including the subject's age, sex, self-assessed sleep quality, coffee consumption, comfort with the driving task, yearly driving frequency, general daily schedule, diagnosed sleep conditions, and use of medication or devices that may influence their natural heart rate.

For the day of the acquisition, the subject is to be asked not to drink caffeinated or alcoholic beverages and to sleep six to nine hours the previous night. If possible, the acquisition is to be scheduled after lunchtime since the probability of falling asleep in this period increases three-fold [109]. Just before the start of the data collection, the volunteer is to have a short test drive for as long as needed, to acclimate to the particular driving experience in the simulator. This initial preparation period aims to minimize the influence of the adaptation to the simulator in the collected data.

After this initial period, the subject is to drive a car out of an urban area and into a highway, taking approximately 2 minutes. They should drive on the aforementioned highway for approximately sixty minutes before arriving at their destination, another urban area where the drive ends. During the drive, the subject is to be periodically asked (every 15 minutes) for their self-assessed score on the KSS, on a range from 1 (the highest level of alertness) to 9 (the highest level of sleepiness where the subject is combating falling asleep). After the acquisition, the general well-being of the subject is to be assessed (looking for signs of kinocytosis and other side effects), and they should fill out two final questionnaires: the System Usability Scale (SUS) and the Immersive Tendencies Questionnaire (ITQ).

The protocol described above complies with the data regulations law in Portugal.

### 3.5.0.1 Questionnaires

All the questionnaires referred in 3.5 are included in annex I, being that only the one referring to the characterization of the subject was developed by the current work's author. The other questionnaires included were either, to the best of the author's knowledge, free to use, or specifically authorized for use in the present thesis.

The questionnaires used are following described briefly:

As questionnaires that aim to collect information for the study:

- The **Horne and Ostberg's MEQ** - as included in annex I.1 - is a 16-question questionnaire in which multiple-choice questions and visual-analog scales are used to assess a subject's circadian typology. The MEQ is the most widely used measure for morningness [110], and its original version was developed by Horne and Ostberg in 1976 [109]. The version intended for use in the current study was translated into European Portuguese by Silvério, Silva, and Macedo in 1998 [111].

- The **ESS**<sup>®</sup> - as included in annex I.2 - is an 8-question questionnaire in which the subject's answers to multiple-choice questions allow the extraction of their position in this sleepiness scale. The ESS questionnaire is a tool mainly used to detect excessive daytime sleepiness [28]. The version intended for use in the current study was a European Portuguese copy of the original questionnaire, obtained through the official Epworth Sleepiness Scale platform <sup>3</sup> and licensed for use in the present thesis and the AUTOMOTIVE project.
- The **ITQ** - as included in annex I.3 - is a widely used tool for evaluating individuals' engagement with mediated environments and their susceptibility to becoming fully immersed in such experiences. This questionnaire, introduced by Witmer and Singer in 1998 [112], consists of 18 questions, each answered on a 7-point Likert scale. The ITQ aims to capture an individual's capacity for immersion, which is characterized by the ability to focus on an event while disregarding any distracting or disruptive external factors in their surroundings [113].
- The **Subject Characterization Questionnaire** - as included in annex I.5 - is designed for the present study as a way to gather additional information useful for the driver drowsiness problem. It includes 10 questions, varying in structure, where participants are asked to provide information about their age, gender, self-assessed sleep quality, caffeine consumption habits, personal experience with driving a car, sleep and heart pathologies, and details about their typical daily schedule. This questionnaire serves as a fundamental tool for capturing essential demographic and lifestyle information, ensuring an even more comprehensive understanding of the subjects.

As questionnaires that aim to exempt/safeguard some participants from participating in the study:

- The **MSSQ-short** - as included in annex I.4 - is the short-form of the MSSQ questionnaire. This version has two blocks of 9 questions each, one to be answered with childhood experiences in mind, and the other referring to adulthood. This questionnaire aims to determine a subject's susceptibility to experiencing kinetosis symptoms as a consequence of motion. Kinetosis symptoms can take the form of nausea and headaches, which are all undesired outcomes for any participant in a practical study.

As a questionnaire that aims to assess the acquisition setup's usability:

- The **SUS** - as included in annex I.6 - is a widely used questionnaire designed by John Brooke in 1995 [114] to assess the usability of a system or product. It consists of a set of ten questions, each answered on a 5-point Likert scale. It aims to evaluate

---

<sup>3</sup><http://epworthsleepinessscale.com/>

a user's subjective perception of a system's usability, covering aspects such as ease of use, learnability, and overall user satisfaction. It provides insights into the user experience and constitutes an essential tool for assessing the usability of systems and interfaces.

## DEVELOPING A FACE-BASED DRIVER DROWSINESS MONITORING SYSTEM

Initially, a method that served as an introduction to the problem, focusing on the real-time computation of eye-aspect ratios — an approach motivated by various instances in the Literature, was undertaken. This method, quick to implement and execute, served as the stepping stone in this journey and formed the basis for a supervised internship (noted in Appendix A). As this method did not end up a part of the main pipeline of this work, it is presented and reported in Appendix A's section A.3.1.

This chapter then delves into the methodology adopted to tackle the primary development process of this dissertation, following a logical progression. In section 4.1, "Face-landmark Study", algorithms for face detection and landmark placement are explored, enabling the eye region to be localized in the face frame, which is important for the next step. Then comes the extraction of the left eye in frames (explained in section 4.2, "Eye-region extraction") and subsequent process of deriving features from these images (explained in section 4.3, "Feature engineering for eye-state classification") in order to then perform the "Eye-state classification"(section 4.4). Moving forward, from the eye-state classification throughout the video samples, features that enable the detection of driver drowsiness are retrieved (explained in section 4.5, "Feature engineering for drowsiness-level classification"), laying the foundation for the final step, "Drowsiness classification"(section 4.6), where the methodologies culminate to detect driver drowsiness.

### 4.1 Face-landmark Study

This part's focus is on the approach that will allow the extracting of the eye region from video frames, involving an evaluation of algorithms for face detection and landmark

placement. The face detection phase serves the purpose of identifying the face’s bounding box within an image, while the landmark prediction/inference phase positions the landmarks with respect to the top-left corner of this bounding box.

Given the wealth of available content and implementations on face detection algorithms online, the current work capitalized on existing resources, mainly on the PFL library [115] resulting in four combinations of algorithms that employ 68-point face landmark predictions:

- FaceBoxes + MobileNet
- FaceBoxes + MobileFaceNet
- RetinaFace + MobileNet
- RetinaFace + MobileFaceNet

It’s worth noting that the positioning-scheme of these 68 landmarks is depicted in Figure 2.6, for reference.

As mentioned in the previous chapter, the DROZY database includes 270 samples annotated with 68 manually placed landmarks. These manual annotations serve as a ground truth for evaluating the performance of the four combinations, by comparing the predicted landmarks generated by each combination with the ground truth.

To quantify the predictions’ quality, an error calculation was carried out for each landmark, by computing the Euclidean distance between the predicted landmark position (in terms of  $(x,y)$  coordinates) and the corresponding ground truth (true landmark position). This distance measurement is expressed in pixels and is mathematically represented in equation 4.1.

$$Error = \sqrt{(x_{\text{predicted}} - x_{\text{true}})^2 + (y_{\text{predicted}} - y_{\text{true}})^2} \quad (4.1)$$

For instance, if an algorithm combination places a given landmark at coordinates  $(50,70)$ , and the true landmark’s position is  $(59,81)$ , the resulting error would be calculated as follows:

$$Error = \sqrt{(50 - 59)^2 + (70 - 81)^2} = \sqrt{202} \approx 14.21 \text{ pixels}$$

The Multi-task Cascaded Convolutional Neural Network (MTCNN) application was also performed and this algorithm quickly delivers, in addition to face bounding box coordinates and 5 landmarks locations (as pictured in Figure 2.3), a probability score associated with each face detection, which allowed the exclusion of face detections with a lower likelihood of representing an actual face (sometimes, in a single frame, 2 or more faces were detected, erroneously).

## 4.2 Eye-Region Extraction

This part of the work managed to be executed only on having, for each considered video frame, the face bounding box coordinates and a fair indication of the left eye position.

To advance to further analysis, the dimensions of the detected faces were computed in pixels - specifically, the width and height of the face bounding box were determined, based on the coordinates retrieved from the previous step using MTCNN.

For the purpose of isolating the left eye region, the left eye landmark - represented by coordinates  $(x_{\text{leftEye}}, y_{\text{leftEye}})$  and retrieved from the previous step using MTCNN - was chosen.

Then, for a width reference, the observation that the width of one eye typically occupies around 20% of the face width [116] was used. Subsequently, the height of the eye, after trials, was consistently found to be accommodated within 15% of the face height.

The landmark and size references were then used to calculate the coordinates of the top-left,  $P_0$ , and bottom-right,  $P_1$ , corners of the eye bounding box, as described in equations 4.2 and 4.3, respectively.

$$P_0(x, y) = (x_{\text{leftEye}} - 0.1 \times \text{FaceWidth}, y_{\text{leftEye}} + 0.075 \times \text{FaceHeight}) \quad (4.2)$$

$$P_1(x, y) = (x_{\text{leftEye}} + 0.1 \times \text{FaceWidth}, y_{\text{leftEye}} - 0.075 \times \text{FaceHeight}) \quad (4.3)$$

These computed eye bounding box coordinates were then employed to extract the corresponding eye frames for the next step.

## 4.3 Feature Engineering for Eye-State Classification

Following the approach in articles [117] and [118], the histograms of the local binary patterns (LBPs) transformation images were selected as features for the eye-state classification.

The LBP algorithm evaluates the texture characteristics of an image by considering the collection of pixels around each pixel. It assigns a score that ultimately reflects the presence or absence of contrast in the examined region.

To derive these informative LBP-based features, a series of sequential transformations were applied to both the CEW dataset, which later served as the training data for eye-state classification, and the DROZY dataset, later used for testing in the eye-state classification phase:

1. **Image preprocessing:** Given that the cropped eye frames had limited definition, these images were initially pre-processed to enhance the quality of the dataset. For that, it was first applied a gaussian smooth filter, regulated by a sigma parameter, and then a histogram equalization.

2. **Computing the LBPs:** The initial step involved the computation of LBPs for the images, yielding a texture image for each. The LBP algorithm employs two critical parameters: radius and the number of sampling points. Experimentation with various parameter values was performed to identify the better configuration from the pool of trials.
3. **Generating the histogram (H):** Subsequently, histograms were constructed based on the texture images. These histograms provided a compact representation of the LBP textures.
4. **Generating the normalized histogram (NH):** This step was necessary due to variations in the pixel dimensions of eye images across datasets and frames (within the DROZY dataset, the dimensions depended on the displayed size of each subject's face in the frame, which, in turn, was related to their head size and proximity to the recording, which varied slightly over time). As such, a normalization process was implemented to express the histogram intensities as percentages, so a relative frequency was presented instead of an absolute one. The normalization operation performed for each image's histogram went as represented in equation 4.4:

$$NH_{intensities} = \frac{H_{intensities}}{P_x} \times 100 \quad (4.4)$$

Where:

- $P_x$  represents the total number of pixels in the image,
- $H_{intensities}$  denotes the original histogram intensities  $(h_0, h_1, \dots, h_n)$ , and
- $NH_{intensities}$  denotes the normalized histogram intensities  $(nh_0, nh_1, \dots, nh_n)$ .

The resulting normalized histograms were employed as inputs for the subsequent step.

## 4.4 Eye-State Classification

In the pursuit of eye-state classification using a machine learning approach, data division strategy was firstly employed to ensure robust model training and evaluation. Last step's resultant histograms were employed as follows:

- **Training dataset** - For training, the ones originated from the Closed Eyes in the Wild (CEW) database, which were labeled as "eye open" or "eye closed", were used.
- **Validation dataset** - To gauge the classification performance, a validation dataset was established using 50 histograms from the DROZY database, equally divided between open and closed eyes and manually labeled accordingly. This validation dataset allowed for the assessment of classifier performance and the estimation of confidence in the classification results obtained during the testing phase.

- **Testing dataset** - In accordance with the validation dataset, the testing dataset was composed also of histograms from the DROZY database - this time, of all the samples (naturally, unlabeled) admitted to the study.

Before proceeding, arbitrary samples of open and closed eyes were selected from the 3 datasets, to confirm that the texture images and their respective normalized histograms, retrieved from applying the methods described in section 4.3, manifested apparent differences.

Two different classifiers were evaluated, a Support Vector Machine (SVM) and a K-Nearest Neighbors (KNNs) classifier. SVM was expected to perform well, following the approaches in the Literature [117], [118], and since this is a machine learning method typically employed in binary classification problems, as the present (class "eye open" and "eye closed"). To fine-tune their performance, a grid search was conducted on specific hyperparameters. Using the training dataset and splitting them in 5-folds for a cross-validation where each batch persevered the same percentage of classes.

For the SVM classifier, the key parameters taken under consideration, as well as their accepted values, and short description, were:

- **Kernel:** "linear", "rbf"  
Kernel type of the algorithm;
- **C:** 0.1, 1, 10, 100  
Strength of the parameter regularization;
- **Gamma:** 0.1, 0.01, 0.001  
In case kernel = "rbf", gamma is an indicator for the shape of the decision boundary.

For the KNNs classifier, the key parameters taken under consideration, as well as their accepted values, and short description, were:

- **Number of neighbors:** 2, 3, 4, 5, 6, 7;
- **P:** 1, 2  
1: the metric used is the Manhattan distance; 2: the metric used is the Euclidean distance;
- **Weights:** "uniform", "distance"  
"uniform": all points are weighted equally; "distance": points are weighted by the inverse of their distance, giving a greater influence to neighbor samples that are closer;
- **Algorithm:** "ball\_tree", "kd\_tree"  
Type of algorithm used.

By selecting the model with the highest accuracy on the validation set, the best-performing classifier was retained. Subsequently, this classifier was applied to the testing data, resulting in an eye-state classification for each frame.

To enhance the input data for the subsequent steps of the process, manual labeling of the eye-state for the entire testing dataset was also conducted. This approach aimed to lay a trustworthy foundation for the subsequent stages and analysis of drowsiness classification, independent of this stage's performance.

## 4.5 Feature Engineering for Drowsiness-Level Classification

At the start of this step, the eye-state throughout the video samples is available data. Before extracting features from the eye-state progression, the adequacy of the frame rate in relation to determining human blinks was assessed. The Nyquist theorem, as shown in Equation 4.5, was applied to evaluate the appropriateness of the frame rate.

$$f_s > 2f_m \quad (4.5)$$

Here,  $f_s$  represents the sampling frequency (frame rate), and  $f_m$  represents the frequency of interest (blink rate). The DROZY database has samples with a frame rate of 15 frames per second (FPS) and others with one of 30 FPS. Meanwhile, as previously mentioned in section 2.3 the duration of a blink tends to vary between 0.1-0.4 seconds [63], [64]. This means that, in the interest of this work:

$$1/0.4 \leq f_m \leq 1/0.1$$

$$2,5 \leq f_m \leq 10 \text{ (Hz)}$$

As such, when applying the Nyquist theorem following equation 4.5, we get:

$$f_s \geq 10 \times 2 \qquad f_s \geq 2,5 \times 2$$

$$f_s \geq 20 \qquad f_s \geq 5$$

For a sample with a sampling frequency of 30 FPS, its frame rate will not constitute a problem, according to Nyquist's theorem. However, for a sample with a sampling frequency of 15 FPS, it might be - if the duration of a blink is especially short, there is the risk of it not being properly registered in the stream of data. This factor was also taken in caution, and a sample (sample number 28) was removed from consideration since its sampling frequency was 15 FPS but it was "Alert" and, as such, the blink instances could be shorter according to the Literature presented in chapter 2.

Given the eye-state progression over time, several important features, matching the information provided on section 2.3, were extracted. In order to retrieve them, the first concepts to denote were the frequency and duration of blinking:

- a frequency sample list was considered as  $F = [f_1, f_2, \dots, f_n]$ , where  $f = 1/\lambda$ , and  $\lambda$  represents the time distance between 2 consecutive blink events.
- a duration sample list was considered as  $D = [d_1, d_2, \dots, d_n]$ , where  $d$  represents the time duration of each blink event.

The features were, hence, extracted as follows:

- **Blink Rate**  
Number of blink events of the entire sample, divided by the sample time;
- **Mean Blink Frequency**  
Mean value of  $F$ ;
- **Standard Deviation of Blink Frequency**  
Standard deviation value of  $F$ ;
- **Mean Blink Duration**  
Mean value of  $D$ ;
- **Standard Deviation of Blink Duration**  
Standard deviation value of  $D$ ;
- **percentage of eye closure (PERCLOS)**  
Amount of time the eyes were closed divided by the sample's total time.

The **Driving Time** was also used, following an extrapolation from the "time on task", due to its impact on cognitive fatigue processes, as previously detailed in chapter 3.

Although yawn counting was contemplated as an additional feature, it was ultimately not included in the analysis due to the limited occurrence of only four yawns (spaced throughout 3 video samples) in the entire dataset. As such, the data size was deemed insufficient for meaningful classification.

## 4.6 Drowsiness Classification

For the drowsiness classification task, the features extracted in the previous section were employed. To promote the robustness of the classification model, the dataset was manually partitioned, with the sample allocation being disclosed in Appendix E. This partitioning aimed to meet two specific criteria:

1. **Person Inclusivity** - A manual data split was executed to ensure that each individual in the test dataset had at least one distinct sample available in the training dataset.

2. **Diversity of Subjects** - Given the relatively small size of the database, it was critical to ensure that the test dataset incorporated samples from various individuals, since a random attribution could result in a testing dataset with multiple samples from the same people. This diversity was necessary to avoid overfitting to individual characteristics and promote the model's ability to generalize.

For the actual drowsiness classification, the Random Forest algorithm, with number estimators = 100 (default value in *Scikit Learn* library), was employed, both for a 3-class problem (alert, moderately drowsy, and very drowsy) and binary problem (alert, very drowsy). The relative frequency of these classes was presented in Table 3.3. Random Forest is a powerful ensemble learning method that combines multiple decision trees to make accurate predictions. Upon drowsiness classification, an analysis was performed to determine which features had the most significant impact on the model's predictive performance. Gini feature importance was used as it is a common method to get this information particularly in decision tree-based models like the Random Forest. This step was crucial for better comprehending and discussing the classification.

## RESULTS AND DISCUSSION

This chapter unfolds in the order set forth in the preceding chapter, showcasing and discussing the previously presented methods' findings. In this chapter, figures, tables, and graphics constituted a preferred form for the representation of results.

### 5.1 Face-landmark Study and Eye ROI Extraction

The analysis of the application of face detectors and 68-landmark predictors following the methods explained in Section 4.1 resulted, primarily, in a summary table with the computed errors in relation to the true coordinates. This table is presented in Appendix F. Two graphical approaches were taken from this table to visualize the table's data and significance better.

First, given a frame sample, the mean error in position (in pixels) throughout all the 68 landmarks predicted in it was computed. This procedure was performed for all the frame samples in analysis in this step, and a histogram with the distribution of the mean error found in these samples was generated. A normalized Gaussian distribution approximation was calculated using the mean and standard deviation values over the histogram. This procedure was performed for the four combinations, and the result is depicted in Figure 5.1.

The histograms on both combinations using MobileNet are similar, and these combinations perform slightly worse than the ones using MobileFaceNet. The combinations using MobileFaceNet fit better to the Gaussian distribution approximation. Between the two combinations using MobileFaceNet, the combination with RetinaFace obtains the overall better performance in this analysis framework where we consider the mean error, calculated as in equation 4.1, as the performance indicator.

Then, we still had to see how these mean errors ranging up to 5 pixels materialized

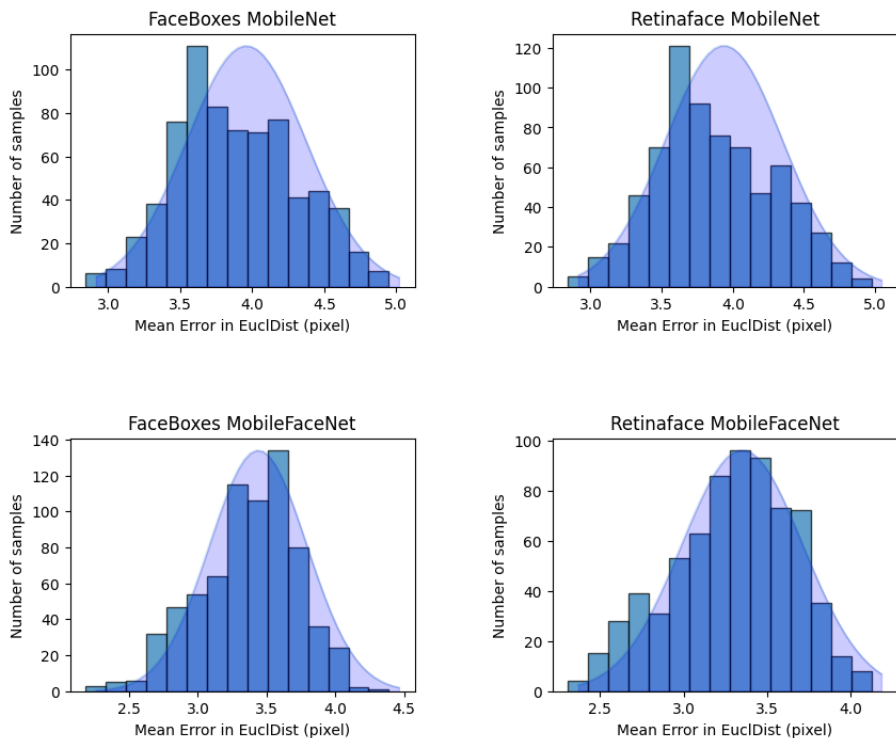


Figure 5.1: Histogram of error distribution for the 4 algorithm combinations tested against the manually labeled landmarks available on the DROZY database. "N" represents the number of samples; "Mean error in Euclidean Distance" represents the average error, in pixels, of each sample image. With a normalized Gaussian distribution approximation over the histogram, calculated from the mean and standard deviation value of each histogram. Representation of the data included in Appendix F.

in a picture frame, mainly how they were distributed throughout the face and, primarily, the eye region. Given a landmark from the 68 points (for example, the landmark indicating the most-left point of the left eyebrow), the mean error in its position (in pixels) throughout all the frame samples in analysis in this step was computed. This mean error took the form of a radius centered on the true landmark position in a random frame, giving a sense of an area of probability for the placement of that landmark. This procedure was performed for all 68 landmarks of each of the four combinations, and the result is depicted in Figure 5.6.

Although MobileFaceNet usage performed better when looking at the histograms, in the region of interest of the current work - the eye - the performance was reasonably steady across all four combinations. The region consisting of the chin and face limits was where MobileFaceNet exhibited a more apparent benefit. As can be observed in the figure, the error radius, especially considering the region of interest, is substantial. The landmark prediction in the left eye, for a reason unbeknownst, performs more accurately than in the right eye. It should be taken into consideration that this is a visualization based on the mean error, and therefore the error may be even worse on occasion, and also

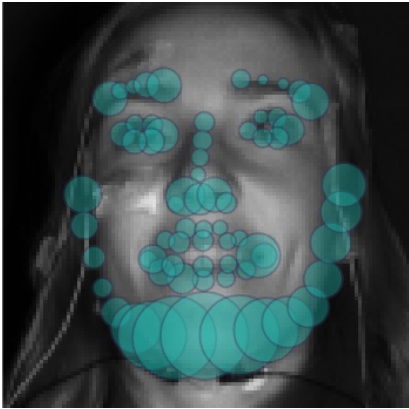


Figure 5.2: FaceBoxes-MobileNet.

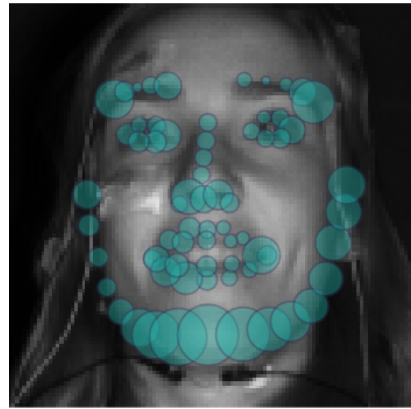


Figure 5.4: FaceBoxes-MobileFaceNet.

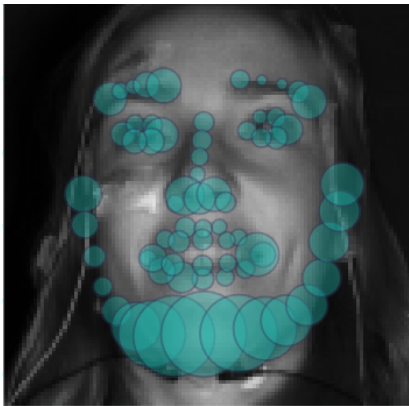


Figure 5.3: RetinaFace-MobileNet.

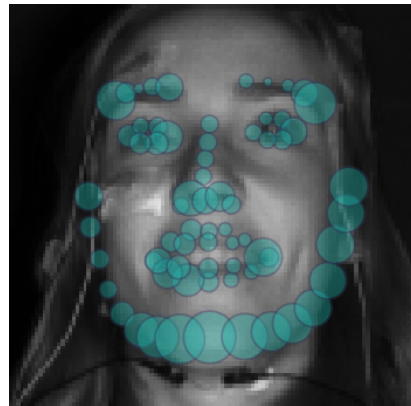


Figure 5.5: RetinaFace-MobileFaceNet.

Figure 5.6: Error representation in an example frame for the 4 algorithm combinations tested. Each circle's radius corresponds to the mean error of each landmark point throughout all the samples. Representation of the data included in Appendix F.

that the radius circumferences for the points in the upper and bottom eyelids are almost touching in the left eye (in the right eye, some overlap can even be observed in all four combinations). As such, an eye-state detector based on the analysis and variation of the placement of the upper and bottom eyelids' landmarks over time was deemed to have too much uncertainty to be a viable option for a step in such a starting point of this work, as it would introduce too much error.

On the other hand, detection of the eye-state based not on upper and bottom eyelid landmark placement monitoring over time but on a machine learning approach, where we could take advantage of the vast number of available images of open and closed eyes to train an algorithm to detect if DROZY's eyes were open or closed, seemed more promising, and also coherent with the Literature. For this endeavor, only a fair idea of the eye's location was needed to provide the region of interest to the algorithm, similar to the cropped eye photographs found in eye databases, instead of the whole face.

Algorithm	Processing speed (FPS)
RetinaFace + MobileFaceNet	1,41
RetinaFace + MobileNet	9,75
FaceBoxes + MobileNet	15,82
FaceBoxes + MobileFaceNet	1,45
MTCNN	79,29

Table 5.1: Algorithm processing speed of each combination of face detection and landmark prediction

Using an MTCNN for prediction resulted well in this task as, of the five landmarks provided, one corresponds to a fair center of the left eye - an example of detection can be seen in Figure 5.7. Additionally, as sometimes face detectors would incorrectly pick up more than one face in the same frame (which we know to be impossible considering DROZY's acquisition settings), using MTCNN for face detection as well was advantageous as it yielded a probability score, allowing the exclusion of detection instances with a lower likelihood of representing an actual face.

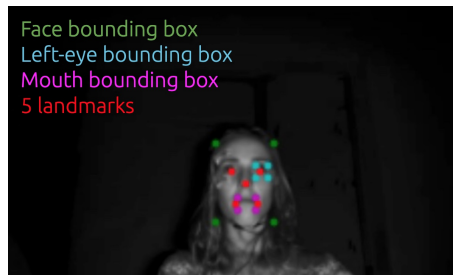


Figure 5.7: MTCNN face detection (represented by the green dots) and landmark prediction (represented by the red dots) in a sample frame.

Regarding the processing speed of the frames, the five methods (for face detection and 68 and 5 landmark prediction) are compared in Table 5.1.

Considering the speed, the production of a probability score for face detection, and the effectively needed data for the next step, the MTCNN was selected for detecting the face and the left eye landmark.

Given the left eye landmark, the eye bounding box was drawn according to the logic in equations 4.2 and 4.3, and cropped frames were extracted. The eye width was considered to be 20% of the face width, as found in the Literature, and the eye width was successfully and consistently encompassed in the cropped eye frame. Initially, as it was not yet figured out that the eye height was accommodated within 15% of the face height, some frames turned out improperly cropped (with the eye not being fully inside the image) until trial and error resulted in the 15% value - consequently perfecting equations 4.2 and 4.3 and the cropping of frames.

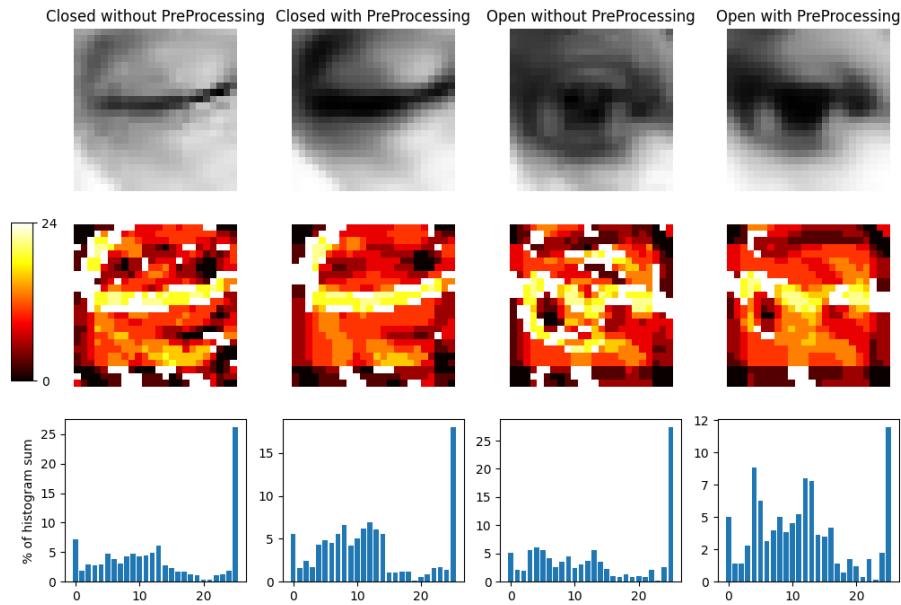


Figure 5.8: Open and closed eye images before and after preprocessing (first row), with the respective LBP image transformation (second row) and feature histograms (third row).

## 5.2 Eye-state classification

Following the logic of the machine learning approach justified in the previous section and the data division strategy presented in Section 4.4, the training, validation, and test datasets went through the four sequential steps presented in Section 4.3: image preprocessing (1); computing the local binary patterns (LBPs) (2); generating the histogram (3); and generating the normalized histogram (4). After experimentation, the parameters with the best outcome were the following:

1. Preprocessing
  - Sigma parameter (suavization factor): 1
2. Computing the LBPs:
  - Radius: 3
  - Number of points: 24

Representative results from two arbitrary samples (one closed-eye and one open-eye) from steps 1, 2, and 4 are showcased in Figure 5.8 for discussion.

Looking at the four histograms in Figure 5.8, we can see a difference in the preprocessing application and between the open and closed eyes' histograms (whether the preprocessing was applied or not).

Zooming in on the first line of each figure, we can observe that the suavization process reduced the noise in the images, and this, together with the local equalization, made the

histograms retrieved from the preprocessed images between the different eye-states more distinct than the ones retrieved from the non-preprocessed images. As such, applying this preprocessing was a helpful step in enhancing the binary classification performance.

Furthermore, throughout the eye images, the normalized LBP histogram layout was steady across an eye-state, concurring with the claims that it constitutes a good feature for this problem.

As for the classification step, the results from the grid search conducted on the hyperparameters explained in Section 4.4 are summarized in Figure 5.9, where the tested classifiers are compared. These results favored using the Support Vector Machine (SVM) over the K-Nearest Neighbors (KNN) since the average accuracy is greater for the former.

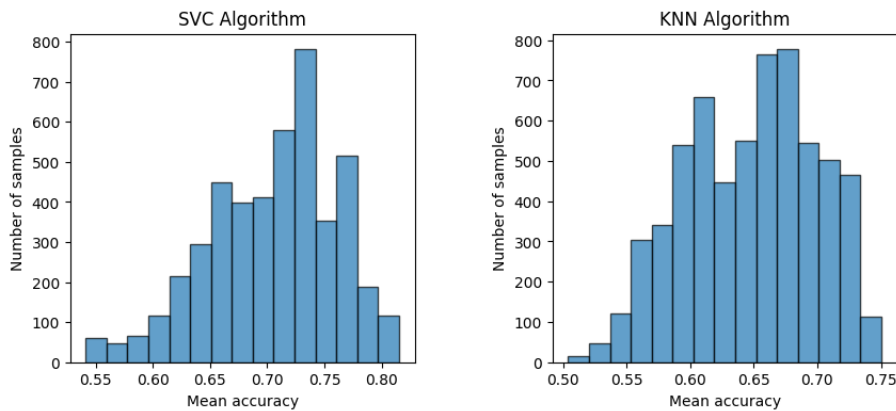


Figure 5.9: Histogram of accuracy distribution for the two algorithms, SVC and KNN, under a grid search.

After obtaining these results, the parameters were fine-tuned with minor changes to obtain the fittest option for this problem, considering a SVM-based solution. In this step, the choice of the parameters was once again based on the evolution of the accuracy score of the algorithm when applying a 5-fold cross-validation over the validation dataset.

The final best-performing model was a SVM with the following hyperparameters:

- Kernel: "rbf"
- C: 30
- Gamma: 0.01

This SVM resulted in the confusion matrix depicted in Table 5.2, accomplishing the performance depicted in Table 5.3.

### 5.2.1 The Challenge of Ambiguous Frames and Human-like Dilemma

Given the accuracy provided by the best model, the validation dataset was examined to showcase the frames that were being incorrectly classified. This process ended up revealing a challenge in classifying eye frames that could be characterized as "dubious- these

Table 5.2: Confusion matrix of the eye-state classification.

		Actual Class	
		Closed	Open
Predicted Class	Closed	70	235
	Open	17	638

Table 5.3: Performance metrics of the eye-state classification.

Performance Metrics	Value
Accuracy	0,74
Recall of Closed eyes	0,23
Precision of Closed eyes	0,80
Recall of Open eyes	0,97
Precision of Open eyes	0,73

were frames that even human observers, when isolated from time context information (previous and following frames), found challenging to categorize as either open or closed eyes. To explore this further, a collection of such "dubious" frames was presented to friends and family for their input, resulting in opinions diverging, and a consensus remaining elusive. An example of such a frame is illustrated in Figure 5.10. Subsequently, a light experiment was conducted by feeding additional "dubious" eye frames, unlabeled, into the SVM model, only to see that it mirrored human uncertainty in classifying these frames (it did not appear to favor any class by default).

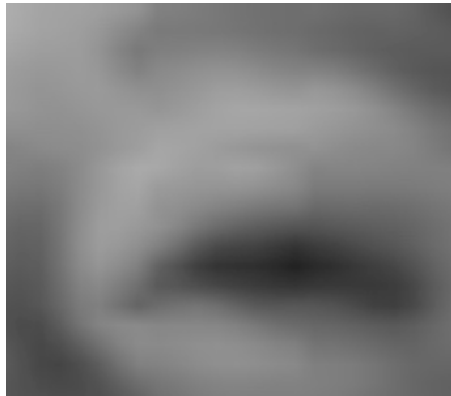


Figure 5.10: An example of a "dubious" eye that gathered indecision among humans.

In light of these observations, and recognizing that the testing dataset, which was being labeled automatically by this classifier, naturally comprised frames with varying degrees of ambiguity, manual labeling of eye-state (open/closed) for all frames within the DROZY dataset was undertaken to provide a more trustful basis for the drowsiness classification. However, it is important to acknowledge that manual labeling introduced a potential labeling bias, as the labeler's judgment was the deciding factor when dealing with "dubious" eye frames. A way to mitigate it would be to have more than one person

performing the labeling of the same samples but, as these amounted to 88920, it was not feasible. This manual labeling of the frames also relied on the examination of the video as a continuum, instead of looking at frames in isolation.

### 5.3 Drowsiness classification

Regarding the 3-class classification, the resulting random forest resulted in the confusion matrix depicted in Table 5.4, and corresponding performance metrics, displayed in Table 5.5. These results reveal a less-than-ideal classification, especially evident in the intermediate class (corresponding to the moderately drowsy).

Table 5.4: Confusion matrix of the 3-class drowsiness classification.

		Actual Class		
		Alert	Moderately Drowsy	Very Drowsy
Predicted Class	Alert	1	0	0
	Moderately Drowsy	0	1	1
	Very Drowsy	2	1	6

Table 5.5: Performance metrics of the 3-class drowsiness classification.

Performance Metrics	Value
Accuracy	0,66
Precision of Alert State	1,00
Recall of Alert State	0,33
Precision of Drowsy State	0,66
Recall of Drowsy State	0,86

These can be expected since we are dealing with a small dataset and a 3-class classification problem with imbalanced class distribution (i.e., significantly more samples in the "very drowsy" class). As it happens, the three classes have significant overlaps in feature space, so, understandably, it is more challenging for the Random Forest to distinguish them. However, no Very Drowsy sample was classified as Alert, constituting a positive outcome.

Regarding the binary classification, the resulting random forest resulted in the confusion matrix depicted in Table 5.6, and corresponding performance metrics, displayed in Table 5.7.

Table 5.6: Confusion matrix of the binary drowsiness classification.

		Actual Class	
		Alert	Very Drowsy
Predicted Class	Alert	1	0
	Very Drowsy	2	7

Table 5.7: Performance metrics of the binary drowsiness classification.

Performance Metrics	Value
Accuracy	0,80
Precision of Alert State	1,00
Recall of Alert State	0,33
Precision of Drowsy State	0,77
Recall of Drowsy State	1,00

An evolution in the retrieved performance metrics, in comparison to the 3-class classification, is evident. Still, the imbalanced class distribution might also be at play here: the "very drowsy" class has a higher frequency, so the classifier might be biased toward this class. The recall for this class being 100% also strengthens this possibility.

A smaller dataset can often lead to issues when performing multi-class classification compared to binary classification. The problem itself is also more complex in this 3-class approach than in the binary one since, in addition to there being more possible outcomes, there are also overlaps in feature space. As such, dealing with only two classes was plausibly easier for the Random Forest to differentiate.

Regarding feature importance, a higher Gini feature importance indicates that a feature contributes more to the model's predictive performance. Figure 5.11 illustrates the distribution of Gini feature importances across the selected features. From it, we can conclude that driving time was, actually, the more relevant feature in this classification, but the eye analysis still served as a relevant indicator of drowsiness, with PERCLOS and blink duration still displaying significant importance. These conclusions are also consistent with the Literature - examples mentioned in Section 2.3.

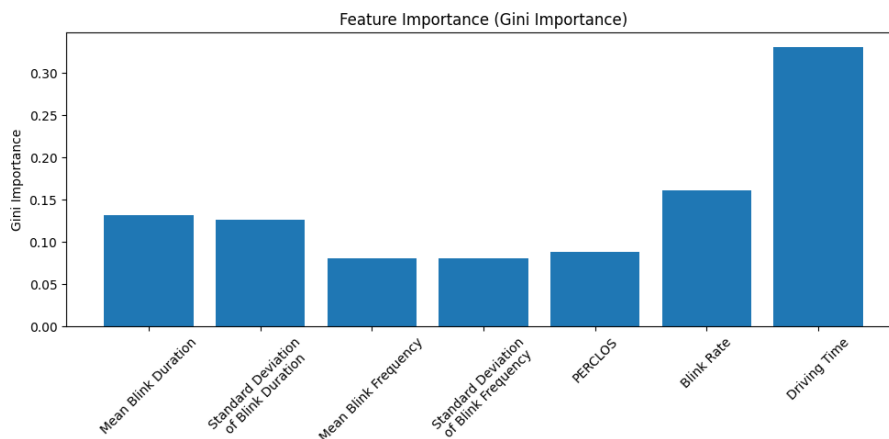


Figure 5.11: Gini feature importance analysis of the binary drowsiness classification of each used featured.

## CONCLUSION

This chapter first delves into this work's conclusions, presented in Section 6.1, as well as points out some directions for future developments, in Section 6.2.

### 6.1 Conclusions

In this dissertation, a comprehensive exploration of driver drowsiness detection was undertaken.

Regarding the core objective number 1 of this dissertation, *i.e.*, "to propose a functional method to detect and monitor drowsiness in automobile vehicle drivers using a seemingly non-intrusive setup", a first baseline approach, focused on real-time face analysis, was achieved, even though the main line of research followed a more robust approach afterward. The investigation turned to eye-state classification, performing a study of methods to extract the eye region of interest (ROI) where the accuracy and speed were considered; selecting local binary patterns (LBPs)'s histograms as features; employing a Support Vector Machine (SVM); and navigating the ambiguity sometimes present in eye frames. The investigation then proceeded to drowsiness-level classification, considering features such as blink frequency, blink duration, and percentage of eye closure (PERCLOS), leveraging a random forest model, and scrutinizing feature relevance.

Ultimately, a procedure was created that takes the face video of a person and the time they have spent driving, which are both non-intrusive measurements, and outputs a drowsiness level prediction. As such, this dissertation's core objective number 1 was addressed and completed. Nevertheless, ultimately, the drowsiness-classification performance analysis was only registered for the manually labeled eye-state data pipeline, and the performance achieved for a 3-class classification, of 66,66% accuracy and 66,66% recall for the drowsy state, is not deemed by this author as sufficient considering the

severity involved in a road safety issue. The binary classification performed better, with 80% accuracy and 100% recall for the drowsy state, an improvement that might have resulted from eliminating the intermediate (moderate) class. Challenges emerged that might have impacted these results, namely the small number of samples in the dataset, the imbalanced class distribution, the fact that some frames were missing, the low sample rate in which some of the dataset's videos were acquired (adequate sample rate is necessary to allow accurate assessment of eye features through video), the ambiguity in some eye frames, and the potential for bias from the manual labeling.

This work also addressed the need for comprehensive datasets in the context at hand. It contributed to this vital area by encompassing the development and description of a data acquisition protocol tailored to the AUTOMOTIVE Project's objectives and conditions, featuring a multimodal approach considering physiological signals, driver behaviour, vehicle behaviour, and individual characteristics. As such, the core objective number 2 of this dissertation, *i.e.*, "to address the current data gap by proposing an experimental design and setup for data acquisition in a simulated driving experience", was done thoroughly.

This dissertation strove to contribute to the overarching goal of equipping vehicles with the capability to monitor driver drowsiness continuously. While meaningful results were achieved, it also serves as a call to action, emphasizing the critical need for expanding datasets, refining classification models, and mitigating potential bias.

## 6.2 Future Work

The groundwork laid in this research may provide a foundation for future studies.

To use this research as a baseline to later enhance it, concerning its core objective 1, by replicating the same or similar methods described in this dissertation, increasing the dataset size and subject variability (ideally, including different ethnic backgrounds) is key.

Multimodal drowsiness detection would then be a great direction to pursue, focusing on different combinations of data sources and methods (possibly exploring deep learning approaches). Taking the AUTOMOTIVE Project as a possible stage for this work, a combination of the current work (on face video and driving time as data sources) with that of Rodrigues' [14] (on the electrocardiogram (ECG)), and Fernandes' [13] (on the Steering Wheel Rotation Angle (SWA)) would be an opportunity for performing this exciting analysis, combining driver behaviour, physiological signals, and vehicle behaviour, and analyzing each input's importance in different settings (*i.e.*, low-lighting). Allied with optimization for real-time computing, this kind of development would, in turn, give the foundation for further studies assessing and implementing an effective alert system.

Additionally, the presence of individual characteristics in driver drowsiness datasets constitutes a step towards our understanding of the human factors involved in safe driving and promotes new lines of research, such as interpretability and causality in artificial

intelligence, which might originate new projects and dissertation topics.

It is this author's hope that this dissertation acts as a catalyst, inspiring further investigation in this critical safety domain.

## BIBLIOGRAPHY

- [1] M. W. Johns, “A sleep physiologist’s view of the drowsy driver”, *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 3, no. 4, pp. 241–249, 2000, ISSN: 13698478. DOI: 10.1016/S1369-8478(01)00008-0 (cit. on pp. 1, 6, 7).
- [2] World Health Organization, “Global Status Report on Road Safety 2018”, Tech. Rep., 2018, pp. 1–8. [Online]. Available: <https://www.who.int/publications/i/item/9789241565684> (visited on 02/12/2021) (cit. on p. 1).
- [3] P. Jackson, C. Hilditch, A. Holmes, N. Reed, N. Merat, and L. Smith, “Fatigue and Road Safety: A Critical Analysis of Recent Evidence”, *Road Safety Web Publication No. 21*, no. 21, pp. 1–88, 2011. [Online]. Available: [www.dft.gov.uk/pgr/roadsafety/research/rsrr%7B%5C%7D0A%7B%5C%7D](http://www.dft.gov.uk/pgr/roadsafety/research/rsrr%7B%5C%7D0A%7B%5C%7D) (visited on 02/12/2021) (cit. on p. 1).
- [4] D. Flatley, L. a. Reyner, and J. a. Horne, “Sleep-Related Crashes on Sections of Different Road Types in the UK”, no. 52, 2004 (cit. on p. 1).
- [5] The Royal Society for the Prevention of Accidents, “Road Safety Factsheet”, Tech. Rep., 2020. [Online]. Available: <https://www.rospa.com/media/documents/road-safety/driver-fatigue-factsheet.pdf> (visited on 02/12/2021) (cit. on pp. 1, 2).
- [6] A. W. MacLean, D. R. Davies, and K. Thiele, “The hazards and prevention of driving while sleepy”, *Sleep Medicine Reviews*, vol. 7, no. 6, pp. 507–521, 2003, ISSN: 10870792. DOI: 10.1016/S1087-0792(03)90004-9 (cit. on pp. 1, 2).
- [7] K. Dobbie, “Fatigue-related crashes : An analysis of fatigue-related crashes on Australian roads using an operational definition of fatigue”, *Australian Transport Safety Bureau*, no. May, 2002, ISSN: 0158-3077. 1-877071-05-6 (cit. on p. 1).
- [8] J. Treat, N. Tumbas, S. McDonald, *et al.*, *Tri-Level Study of the Causes of Traffic Accidents*, 1979. DOI: 10.1136/bmj.1.4975.1098 (cit. on p. 2).
- [9] Society of Automotive Engineers, “Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles”, SAE International, Tech. Rep., 2018, p. 35. DOI: [https://doi.org/10.4271/J3016\\_201806](https://doi.org/10.4271/J3016_201806) (cit. on pp. 2, 3).

- [10] N. Administration, *Automated Vehicles for Safety*. [Online]. Available: <https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety> (visited on 02/14/2021) (cit. on p. 2).
- [11] C. S. Silveira, “Driver’s Fatigue State Monitoring using Physiological Signals”, M.S. thesis, Faculdade de Engenharia da Universidade do Porto, Porto, Portugal, 2017 (cit. on p. 4).
- [12] L. de Oliveira, “Driver drowsiness detection using non-intrusive signal acquisition”, M.S. thesis, Faculdade de Engenharia da Universidade do Porto, Porto, Portugal, 2018 (cit. on p. 4).
- [13] M. Fernandes, “Driver drowsiness detection using non-intrusive electrocardiogram and steering wheel angle signals”, M.S. thesis, Faculdade de Engenharia da Universidade do Porto, Porto, Portugal, 2019 (cit. on pp. 4, 9, 44).
- [14] L. A. Rodrigues, “Driver Drowsiness Detection with Peripheral Cardiac Signals”, M.S. thesis, Instituto Superior Técnico, Lisbon, Portugal, 2021 (cit. on pp. 4, 44).
- [15] T. Esteves, J. R. Pinto, P. M. Ferreira, *et al.*, “Automotive: A case study on automatic multimodal drowsiness detection for smart vehicles”, *IEEE Access*, vol. 9, pp. 153 678–153 700, 2021, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2021.3128 016 (cit. on pp. 4, 60, 62).
- [16] American Sleep Association / Reviewers and Writers, *What is Sleep and Why is It Important?*, 2016. [Online]. Available: <https://www.sleepassociation.org/about-sleep/what-is-sleep/> (cit. on p. 6).
- [17] T. Hori, Y. Sugita, E. Koga, *et al.*, “Proposed supplements and amendments to ‘A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects’, the Rechtschaffen & Kales (1968) standard.”, *eng, Psychiatry and clinical neurosciences*, vol. 55, no. 3, pp. 305–310, Jun. 2001, ISSN: 1323-1316 (Print). DOI: 10.1046/j.1440-1819.2001.00810.x (cit. on p. 6).
- [18] S. A. Keenan, M. Hirshkowitz, and H. Casseres, “Monitoring and staging human sleep”, in *Principles and practice of sleep medicine*, M. Kryger, T. Roth, and W. Dement, Eds., 5th ed., St. Louis: Elsevier Saunders, 2011, ch. Normal Hum, pp. 16–26, ISBN: 9780123786111. DOI: 10.1016/B978-0-12-378610-4.00138-8 (cit. on p. 6).
- [19] N. Lamond and D. Dawson, “Quantifying the performance impairment associated with fatigue”, *Journal of Sleep Research*, vol. 8, no. 4, pp. 255–262, 1999, ISSN: 09621105. DOI: 10.1046/j.1365-2869.1999.00167.x (cit. on p. 6).
- [20] Y. Liang, W. J. Horrey, M. E. Howard, *et al.*, “Prediction of drowsiness events in night shift workers during morning driving”, *Accident Analysis and Prevention*, vol. 126, no. November 2017, pp. 105–114, 2019, ISSN: 00014575. DOI: 10.1016/j.aap.2017.11.004 (cit. on p. 6).

- [21] H. J. Moller, L. Kayumov, E. L. Bulmash, J. Nhan, and C. M. Shapiro, "Simulator performance, microsleep episodes, and subjective sleepiness: normative data using convergent methodologies to assess driver drowsiness", *Journal of Psychosomatic Research*, vol. 61, no. 3, pp. 335–342, 2006, ISSN: 00223999. DOI: 10.1016/j.jpsychores.2006.04.007 (cit. on p. 7).
- [22] L. N. Boyle, J. Tippin, A. Paul, and M. Rizzo, "Driver performance in the moments surrounding a microsleep", *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 11, no. 2, pp. 126–136, 2008, ISSN: 13698478. DOI: 10.1016/j.trf.2007.08.001 (cit. on p. 7).
- [23] C. C. Liu, S. G. Hosking, and M. G. Lenné, "Predicting driver drowsiness using vehicle measures: Recent insights and future challenges", *Journal of Safety Research*, vol. 40, no. 4, pp. 239–245, 2009, ISSN: 00224375. DOI: 10.1016/j.jsr.2009.04.005 (cit. on p. 7).
- [24] C. Jacobé de Naurois, C. Bourdin, A. Stratulat, E. Diaz, and J. L. Vercher, "Detection and prediction of driver drowsiness using artificial neural network models", *Accident Analysis and Prevention*, vol. 126, pp. 95–104, 2019, ISSN: 00014575. DOI: 10.1016/j.aap.2017.11.038 (cit. on pp. 7, 22, 81).
- [25] M. Doudou, A. Bouabdallah, and V. Berge-Cherfaoui, "Driver Drowsiness Measurement Technologies: Current Research, Market Solutions, and Challenges", *International Journal of Intelligent Transportation Systems Research*, vol. 18, no. 2, pp. 297–319, 2020, ISSN: 18688659. DOI: 10.1007/s13177-019-00199-w (cit. on pp. 7, 9, 13, 14).
- [26] M. W. Johns, A. Tucker, R. Chapman, K. Crowley, and N. Michael, "Monitoring eye and eyelid movements by infrared reflectance oculography to measure drowsiness in drivers", *Somnologie*, vol. 11, no. 4, pp. 234–242, 2007, ISSN: 14329123. DOI: 10.1007/s11818-007-0311-y (cit. on p. 7).
- [27] M. Johns, A. Tucker, R. Chapman, N. Michael, and C. Beale, *Johns-et-al-APSS-2006-JDS.pdf*, 2006 (cit. on p. 7).
- [28] M. W. Johns, "A new method for measuring daytime sleepiness: the Epworth sleepiness scale.", eng, *Sleep*, vol. 14, no. 6, pp. 540–545, Dec. 1991, ISSN: 0161-8105 (Print). DOI: 10.1093/sleep/14.6.540 (cit. on pp. 7, 23, 24).
- [29] E. Hoddes, V. Zarcone, H. Smythe, R. Phillips, and W. C. Dement, "Quantification of sleepiness: a new approach.", eng, *Psychophysiology*, vol. 10, no. 4, pp. 431–436, Jul. 1973, ISSN: 0048-5772 (Print). DOI: 10.1111/j.1469-8986.1973.tb00801.x (cit. on p. 7).
- [30] T. H. Monk, "A Visual Analogue Scale technique to measure global vigor and affect.", eng, *Psychiatry research*, vol. 27, no. 1, pp. 89–99, Jan. 1989, ISSN: 0165-1781 (Print). DOI: 10.1016/0165-1781(89)90013-9 (cit. on p. 7).

- [31] K. Kaida, M. Takahashi, T. Åkerstedt, *et al.*, “Validation of the Karolinska sleepiness scale against performance and EEG variables”, *Clinical Neurophysiology*, vol. 117, no. 7, pp. 1574–1581, 2006, ISSN: 1388-2457. DOI: <https://doi.org/10.1016/j.clinph.2006.03.011> (cit. on p. 7).
- [32] R. Tremaine, J. Dorrian, L. Lack, *et al.*, “The relationship between subjective and objective sleepiness and performance during a simulated night-shift with a nap countermeasure”, *Applied Ergonomics*, vol. 42, pp. 52–61, Dec. 2010. DOI: 10.1016/j.apergo.2010.04.005 (cit. on p. 7).
- [33] M. Doudou, A. Bouabdallah, V. Cherfaoui, and V. Charfaoui, “A Light on Physiological Sensors for Efficient Driver Drowsiness Detection System”, *Sensors & Transducers Journal, International Frequency Sensor Association (IFSA)*, vol. 224, no. 8, pp. 39–50, 2019. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-02162758> (visited on 02/12/2021) (cit. on pp. 8, 9, 11, 17).
- [34] Y. Dong, Z. Hu, K. Uchimura, and N. Murayama, “Driver inattention monitoring system for intelligent vehicles: A review”, *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 2, pp. 596–614, 2011, ISSN: 15249050. DOI: 10.1109/TITS.2010.2092770 (cit. on p. 9).
- [35] M. Ingre, T. Åkerstedt, B. Peters, A. Anund, and G. Kecklund, “Subjective sleepiness, simulated driving performance and blink duration: Examining individual differences”, *Journal of Sleep Research*, vol. 15, no. 1, pp. 47–53, 2006, ISSN: 09621105. DOI: 10.1111/j.1365-2869.2006.00504.x (cit. on p. 9).
- [36] P. M. Forsman, B. J. Vila, R. A. Short, C. G. Mott, and H. P. Van Dongen, “Efficient driver drowsiness detection at moderate levels of drowsiness”, *Accident Analysis and Prevention*, vol. 50, pp. 341–350, 2013, ISSN: 00014575. DOI: 10.1016/j.aap.2012.05.005 (cit. on p. 9).
- [37] S. Furugori, N. Yoshizawa, C. Iname, and Y. Miura, “Estimation of driver fatigue by pressure distribution on seat in long term driving”, vol. 26, pp. 53–58, 2005 (cit. on p. 9).
- [38] D. M. Morris, J. J. Pilcher, and F. S. Switzer, “Lane heading difference: An innovative model for drowsy driving detection using retrospective analysis around curves”, *Accident Analysis and Prevention*, vol. 80, pp. 117–124, 2015, ISSN: 00014575. DOI: 10.1016/j.aap.2015.04.007 (cit. on p. 9).
- [39] J. Vicente, P. Laguna, A. Bartra, and R. Bailón, “Drowsiness detection using heart rate variability”, *Medical and Biological Engineering and Computing*, vol. 54, no. 6, pp. 927–937, 2016, ISSN: 17410444. DOI: 10.1007/s11517-015-1448-7 (cit. on pp. 10, 11).

- [40] E. Michail, A. Kokonozi, I. Chouvarda, and N. Maglaveras, "EEG and HRV markers of sleepiness and loss of control during car driving", *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS'08 - "Personalized Healthcare through Technology"*, pp. 2566–2569, 2008, ISSN: 1557-170X. DOI: 10.1109/iembs.2008.4649724 (cit. on p. 11).
- [41] G. D. Furman, A. Baharav, C. Cahan, and S. Akselrod, "Early detection of falling asleep at the wheel: A heart rate variability approach", *Computers in Cardiology*, vol. 35, no. July, pp. 1109–1112, 2008, ISSN: 02766574. DOI: 10.1109/CIC.2008.4749240 (cit. on p. 11).
- [42] A. Baharav, S. Kotagal, V. Gibbons, *et al.*, "Fluctuations in autonomic nervous activity during sleep displayed by power spectrum analysis of heart rate variability.", eng, *Neurology*, vol. 45, no. 6, pp. 1183–1187, Jun. 1995, ISSN: 0028-3878 (Print). DOI: 10.1212/wnl.45.6.1183 (cit. on p. 11).
- [43] Task Force of the European Society of Cardiology and The North American Society of Pacing Electrophysiology, "Heart Rate Variability", *Circulation*, vol. 93, no. 5, pp. 1043–1065, 1996. DOI: 10.1161/01.CIR.93.5.1043 (cit. on p. 11).
- [44] S. Wickramanayake, D. Fernando, S. Jayawardena, and S. Darshana, "Efficient perclos and gaze measurement methodologies to estimate driver attention in real time", Jan. 2014. DOI: 10.1109/ISMS.2014.56 (cit. on p. 11).
- [45] C. Papadelis, Z. Chen, C. Kourtidou-Papadeli, *et al.*, "Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents.", eng, *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, vol. 118, no. 9, pp. 1906–1922, Sep. 2007, ISSN: 1388-2457 (Print). DOI: 10.1016/j.clinph.2007.04.031 (cit. on p. 11).
- [46] A. T. Islam, N. Rahaman, and M. A. R. Ahad, "A study on tiredness assessment by using eye blink detection", *Jurnal Kejuruteraan*, 2019. [Online]. Available: <https://api.semanticscholar.org/CorpusID:236165897> (cit. on pp. 11, 12, 66, 80).
- [47] N. T. B. Pasaribu, A. Prijono, R. Ratnadewi, R. P. Adhie, and J. Felix, "Drowsiness detection according to the number of blinking eyes specified from eye aspect ratio value modification", in *Proceedings of the 1st International Conference on Life, Innovation, Change and Knowledge (ICLICK 2018)*, Atlantis Press, 2019/07, pp. 171–174, ISBN: 978-94-6252-762-1. DOI: 10.2991/iclick-18.2019.35. [Online]. Available: <https://doi.org/10.2991/iclick-18.2019.35> (cit. on pp. 11, 12, 66, 80).
- [48] H. S. Shin, S. J. Jung, J. J. Kim, and W. Y. Chung, "Real time car driver's condition monitoring system", *Proceedings of IEEE Sensors*, pp. 951–954, 2010. DOI: 10.1109/ICSENS.2010.5690904 (cit. on p. 11).

- [49] R. I. Hammoud, G. Witt, R. Dufour, A. Wilhelm, and T. Newman, "On driver eye closure recognition for commercial vehicles", *SAE International Journal of Commercial Vehicles*, vol. 1, no. 1, pp. 454–463, 2009, ISSN: 1946391X. DOI: 10.4271/2008-01-2691 (cit. on p. 11).
- [50] Q. Ji, Z. Zhu, and P. Lan, "Real-time nonintrusive monitoring and prediction of driver fatigue", *IEEE Transactions on Vehicular Technology*, vol. 53, no. 4, pp. 1052–1068, 2004, ISSN: 00189545. DOI: 10.1109/TVT.2004.830974 (cit. on p. 11).
- [51] S. Abtahi, B. Hariri, and S. Shirmohammadi, "Driver drowsiness monitoring based on yawning detection", in *2011 IEEE International Instrumentation and Measurement Technology Conference*, May 2011, pp. 1–4. DOI: 10.1109/IMTC.2011.5944101 (cit. on p. 11).
- [52] T. Brandt, R. Stemmer, and A. Rakotonirainy, "Affordable visual driver monitoring system for fatigue and monotony", in *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No.04CH37583)*, vol. 7, Oct. 2004, 6451–6456 vol.7. DOI: 10.1109/ICSMC.2004.1401415 (cit. on p. 11).
- [53] A. Monster, H. Chan, and D. O'Connor, "Long-term trends in human eye blink rate.", *Biotelemetry and patient monitoring*, vol. 5, no. 4, pp. 206–222, 1978 (cit. on p. 11).
- [54] A. Maffei and A. Angrilli, "Spontaneous eye blink rate: An index of dopaminergic component of sustained attention and fatigue", *International Journal of Psychophysiology*, vol. 123, pp. 58–63, 2018 (cit. on p. 11).
- [55] M. K. Holland and G. Tarlow, "Blinking and thinking", *Perceptual and motor skills*, vol. 41, no. 2, pp. 403–406, 1975 (cit. on p. 11).
- [56] K. Nakamori, M. Odawara, T. Nakajima, T. Mizutani, and K. Tsubota, "Blinking is controlled primarily by ocular surface conditions", *American journal of ophthalmology*, vol. 124, no. 1, pp. 24–30, 1997 (cit. on p. 11).
- [57] A. J. Zametkin, J. R. Stevens, and R. Pittman, "Ontogeny of spontaneous blinking and of habituation of the blink reflex", *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, vol. 5, no. 5, pp. 453–457, 1979 (cit. on p. 12).
- [58] A. Bentivoglio, S. Bressman, E. Cassetta, D. Carretta, P. Tonali, and A. Albanese, "Analysis of blink rate patterns in normal subjects", *Movement disorders : official journal of the Movement Disorder Society*, vol. 12, pp. 1028–34, Nov. 1997. DOI: 10.1002/mds.870120629 (cit. on p. 12).
- [59] M. Johns and C. Hocking, "The effects of unintentional drowsiness on the velocity of eyelid movements during spontaneous blinks", *Physiological Measurement*, 2020, ISSN: 0967-3334. DOI: 10.1088/1361-6579/abd5c3 (cit. on p. 12).

- [60] N. Galley, R. Schleicher, and L. Galley, "Blink parameters as indicators of driver's sleepiness—possibilities and limitations", *Vision in vehicles*, vol. 10, pp. 189–196, 2004 (cit. on p. 12).
- [61] T. Danisman, I. M. Bilasco, C. Djeraba, and N. Ihaddadene, "Drowsy driver detection system using eye blink patterns", in *2010 International Conference on Machine and Web Intelligence*, IEEE, 2010, pp. 230–233 (cit. on pp. 12, 82).
- [62] D. F. Dinges and R. Grace, "Perclos: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance", *US Department of Transportation, Federal Highway Administration, Publication Number FHWA-MCRT-98-006*, 1998 (cit. on p. 12).
- [63] R. Milo, P. Jorgensen, U. Moran, G. Weber, and M. Springer, "Bionumbers—the database of key numbers in molecular and cell biology", *Nucleic acids research*, vol. 38, no. suppl\_1, pp. D750–D753, 2010 (cit. on pp. 12, 31).
- [64] H. R. Schiffman, *Sensation and perception: An integrated approach*. John Wiley & Sons, 1990 (cit. on pp. 12, 31).
- [65] S. A. Schellini, A. A. Sampaio Jr, E. Hoyama, A. A. Cruz, and C. R. Padovani, "Spontaneous eye blink analysis in the normal individual", *Orbit*, vol. 24, no. 4, pp. 239–242, 2005 (cit. on p. 12).
- [66] W. Wierwille, "Historical perspective on slow eyelid closure: Whence perclos", in *Technical Proceedings of Ocular Measures of Driver Alertness Conference, Herndon, VA.(FHWA Technical Report No. MC-99-136)*. Washington, DC: Federal Highway Administration, Office of Motor Carrier and Highway Safety, 1999, pp. 31–53 (cit. on p. 12).
- [67] T. Abe, "Perclos-based technologies for detecting drowsiness: Current evidence and future directions", *Sleep Advances*, vol. 4, no. 1, zpad006, 2023 (cit. on p. 12).
- [68] A. R. Sparrow, C. M. LaJambe, and H. P. Van Dongen, "Drowsiness measures for commercial motor vehicle operations", *Accident Analysis & Prevention*, vol. 126, pp. 146–159, 2019 (cit. on p. 12).
- [69] J. M. Cori, C. Anderson, S. S. Soleimanloo, M. L. Jackson, and M. E. Howard, "Narrative review: Do spontaneous eye blink parameters provide a useful assessment of state drowsiness?", *Sleep medicine reviews*, vol. 45, pp. 95–104, 2019 (cit. on p. 12).
- [70] A. W. Cai, J. E. Manousakis, T. Y. Lo, J. A. Horne, M. E. Howard, and C. Anderson, "I think i'm sleepy, therefore i am—awareness of sleepiness while driving: A systematic review", *Sleep medicine reviews*, vol. 60, p. 101 533, 2021 (cit. on p. 12).
- [71] E. C.-P. Chua, S.-C. Yeo, I. T.-G. Lee, *et al.*, "Individual differences in physiologic measures are stable across repeated exposures to total sleep deprivation", *Physiological reports*, vol. 2, no. 9, e12129, 2014 (cit. on p. 12).

- [72] M. L. Jackson, G. A. Kennedy, C. Clarke, *et al.*, “The utility of automated measures of ocular metrics for detecting driver drowsiness during extended wakefulness”, *Accident Analysis & Prevention*, vol. 87, pp. 127–133, 2016 (cit. on p. 12).
- [73] A. Mortazavi, A. Eskandarian, and R. Sayed, “Effect of drowsiness on driving performance variables of commercial vehicle drivers”, *International Journal of Automotive Technology*, vol. 10, pp. 391–404, 2009 (cit. on p. 12).
- [74] Z. Haq and Z. Hasan, “Eye-blink rate detection for fatigue determination”, Aug. 2016, pp. 1–5. DOI: 10.1109/IICIP.2016.7975348 (cit. on p. 12).
- [75] A. L. Samuel, “Some Studies in Machine Learning Using the Game of Checkers”, *IBM J. Res. Dev.*, vol. 3, pp. 210–229, 1959 (cit. on p. 13).
- [76] N. Buduma and N. Locascio, *Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms*. O’Reilly Media, 2017, ISBN: 9781491925560. [Online]. Available: <https://books.google.pt/books?id=n0I1DwAAQBAJ> (cit. on p. 13).
- [77] P. Ongsulee, “Artificial intelligence, machine learning and deep learning”, *International Conference on ICT and Knowledge Engineering*, pp. 1–6, 2018, ISSN: 2157099X. DOI: 10.1109/ICTKE.2017.8259629 (cit. on p. 13).
- [78] *CardioID | CardioWheel*. [Online]. Available: <https://www.cardio-id.com/cardiowheel> (cit. on p. 13).
- [79] *STEER: Wearable Device That Will Not Let You Fall Asleep by Creative Mode — Kickstarter*. [Online]. Available: <https://www.kickstarter.com/projects/creativemode/steer-you-will-never-fall-asleep-while-driving?lang=fr> (cit. on p. 13).
- [80] *IR-LED Detects Drivers in Microsleep | Research & Technology | Oct 2010 | Photonics.com*. [Online]. Available: [https://www.photonics.com/Articles/IR-LED%7B%5C\\_%7DDetects%7B%5C\\_%7DDrivers%7B%5C\\_%7Din%7B%5C\\_%7DMicrosleep/a44727](https://www.photonics.com/Articles/IR-LED%7B%5C_%7DDetects%7B%5C_%7DDrivers%7B%5C_%7Din%7B%5C_%7DMicrosleep/a44727) (cit. on p. 13).
- [81] *Saab Driver Attention Warning System - Saabnet.com*. [Online]. Available: <https://www.saabnet.com/tsn/press/071102.html> (cit. on p. 14).
- [82] *Driver Alert Control (DAC)\* - operation | Driver Alert System | Driver support | S60 2014 | Volvo Support*. [Online]. Available: <https://www.volvocars.com/en-th/support/manuals/s60/2013w46/driver-support/driver-alert-system/driver-alert-control-dac---operation> (cit. on p. 14).
- [83] *Driver Alert Control (DAC)\* | Driver Alert System | Driver support | V40 2018 | Volvo Support*. [Online]. Available: <https://www.volvocars.com/en-th/support/manuals/v40/2017w17/driver-support/driver-alert-system/driver-alert-control-dac> (cit. on p. 14).

- [84] *Driver alert system* | Volkswagen UK. [Online]. Available: <https://www.volkswagen.co.uk/technology/car-safety/driver-alert-system> (cit. on p. 14).
- [85] M. Taylor, *No Doze: Mercedes E-Class alerts drowsy drivers*. [Online]. Available: <https://www.autoweek.com/news/a2032716/no-doze-mercedes-e-class-alerts-drowsy-drivers/> (cit. on p. 14).
- [86] *Lexus Safety Technology* | Lexus.com. [Online]. Available: <https://www.lexus.com/safety> (cit. on p. 14).
- [87] *Seeing Machines* | *Technology to get everyone home safe*. [Online]. Available: <https://seeingmachines.com/about/> (cit. on p. 14).
- [88] B. Johnston and P. Chazal, “A review of image-based automatic facial landmark identification techniques”, *EURASIP Journal on Image and Video Processing*, vol. 2018, p. 86, Sep. 2018. DOI: 10.1186/s13640-018-0324-4 (cit. on p. 15).
- [89] A. Elmahmudi and H. Ugail, “A framework for facial age progression and regression using exemplar face templates”, *The Visual Computer*, vol. 37, Jul. 2021. DOI: 10.1007/s00371-020-01960-z (cit. on p. 15).
- [90] J. Mafeni Mase, N. Leesakul, F. Yang, G. Figueredo, and M. Torres, *Towards privacy-preserving affect recognition: A two-level deep learning architecture*, Nov. 2021 (cit. on p. 15).
- [91] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, “Joint face detection and alignment using multitask cascaded convolutional networks”, *IEEE signal processing letters*, vol. 23, no. 10, pp. 1499–1503, 2016 (cit. on p. 15).
- [92] J. Deng, J. Guo, Y. Zhou, J. Yu, I. Kotsia, and S. Zafeiriou, “Retinaface: Single-stage dense face localisation in the wild”, *arXiv preprint arXiv:1905.00641*, 2019 (cit. on p. 15).
- [93] S. Zhang, X. Zhu, Z. Lei, H. Shi, X. Wang, and S. Z. Li, “Faceboxes: A cpu real-time face detector with high accuracy”, in *IJCB*, 2017 (cit. on p. 16).
- [94] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “Mobilenetv2: Inverted residuals and linear bottlenecks”, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510–4520 (cit. on p. 16).
- [95] S. Chen, Y. Liu, X. Gao, and Z. Han, “Mobilefacenet: Efficient cnns for accurate real-time face verification on mobile devices”, in *Biometric Recognition: 13th Chinese Conference, CCBR 2018, Urumqi, China, August 11-12, 2018, Proceedings 13*, Springer, 2018, pp. 428–438 (cit. on p. 16).
- [96] J. Martin, M. Hébert, E. Ledoux, M. Gaudreault, and L. Laberge, “Relationship of chronotype to sleep, light exposure, and work-related fatigue in student workers”, *Chronobiology international*, vol. 29, pp. 295–304, Apr. 2012. DOI: 10.3109/07420528.2011.653656 (cit. on p. 18).

- [97] Q. Massoz, T. Langohr, C. Francois, and J. G. Verly, "The ULg multimodality drowsiness database (called DROZY) and examples of use", in *2016 IEEE Winter Conference on Applications of Computer Vision, WACV 2016*, Institute of Electrical and Electronics Engineers Inc., May 2016, ISBN: 9781509006410. DOI: 10.1109/WACV.2016.7477715 (cit. on pp. 18, 19).
- [98] R. Ghoddoosian, M. Galib, and V. Athitsos, "A Realistic Dataset and Baseline Temporal Model for Early Drowsiness Detection", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, vol. 2019-June, pp. 178–187, Apr. 2019. arXiv: 1904.07312 (cit. on p. 18).
- [99] C. Fors, C. Ahlström, P. Sörner, *et al.*, *Camera-based sleepiness detection: final report of the project SleepEYE*. Statens väg-och transportforskningsinstitut, 2011 (cit. on p. 18).
- [100] C. H. Weng, Y. H. Lai, and S. H. Lai, "Driver drowsiness detection via a hierarchical temporal deep belief network", in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10118 LNCS, Springer Verlag, Nov. 2017, pp. 117–133, ISBN: 9783319545257. DOI: 10.1007/978-3-319-54526-4\_9 (cit. on p. 18).
- [101] J.-P. Couderc, X. Xiaojuan, W. Zareba, and A. J. Moss, "Assessment of the stability of the individual-based correction of qt interval for heart rate", *Annals of Noninvasive Electrocardiology*, vol. 10, no. 1, pp. 25–34, 2005 (cit. on p. 18).
- [102] L. Wolf, T. Hassner, and I. Maoz, "Face recognition in unconstrained videos with matched background similarity", in *CVPR 2011*, Jun. 2011, pp. 529–534. DOI: 10.1109/CVPR.2011.5995566 (cit. on p. 18).
- [103] J. A. Healey, "Wearable and automotive systems for affect recognition from physiology", Ph.D. dissertation, Massachusetts Institute of Technology, Massachusetts, United States of America, 2000 (cit. on p. 18).
- [104] M. Basner and D. F. Dinges, "Maximizing sensitivity of the psychomotor vigilance test (pvt) to sleep loss", *Sleep*, vol. 34, no. 5, pp. 581–591, 2011 (cit. on p. 19).
- [105] X. Wang and C. Xu, "Driver drowsiness detection based on non-intrusive metrics considering individual specifics", *Accident Analysis and Prevention*, vol. 95, pp. 350–357, 2016, ISSN: 00014575. DOI: 10.1016/j.aap.2015.09.002 (cit. on p. 21).
- [106] F. Song, X. Tan, X. Liu, and S. Chen, "Eyes closeness detection from still images with multi-scale histograms of principal oriented gradients", *Pattern Recognition*, vol. 47, no. 9, pp. 2825–2838, 2014 (cit. on p. 21).

- [107] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, “Labeled faces in the wild: A database for studying face recognition in unconstrained environments”, in *Workshop on faces in ‘Real-Life’ Images: detection, alignment, and recognition*, 2008 (cit. on p. 21).
- [108] J. F. Golding, “Predicting individual differences in motion sickness susceptibility by questionnaire”, *Personality and Individual Differences*, vol. 41, no. 2, pp. 237–248, 2006. DOI: <https://doi.org/10.1016/j.paid.2006.01.012> (cit. on p. 22).
- [109] J. Horne and L. Reyner, “Vehicle accidents related to sleep: a review.”, *Occupational and Environmental Medicine*, vol. 56, no. 5, pp. 289–294, 1999, ISSN: 1351-0711. DOI: 10.1136/oem.56.5.289 (cit. on p. 23).
- [110] A. Adan, S. Archer, M. P. Hidalgo, L. Di Milia, V. Natale, and C. Randler, “Circadian typology: A comprehensive review”, *Chronobiology International*, vol. 29, Nov. 2012. DOI: 10.3109/07420528.2012.719971 (cit. on p. 23).
- [111] C. F. Silva, J. Silvério, P. Rodrigues, *et al.*, “The portuguese version of the horne and ostberg morningness-eveningness questionnaire: Its role in education and psychology”, 2002 (cit. on p. 23).
- [112] B. G. Witmer and M. J. Singer, “Measuring presence in virtual environments: A presence questionnaire”, *Presence*, vol. 7, no. 3, pp. 225–240, 1998 (cit. on p. 24).
- [113] S. Rózsa, R. Hargitai, A. Láng, *et al.*, “Measuring immersion, involvement, and attention focusing tendencies in the mediated environment: The applicability of the immersive tendencies questionnaire”, *Frontiers in Psychology*, vol. 13, p. 931 955, Jul. 2022. DOI: 10.3389/fpsyg.2022.931955 (cit. on p. 24).
- [114] J. Brooke, “Sus: A quick and dirty usability scale”, *Usability Eval. Ind.*, vol. 189, Nov. 1995 (cit. on p. 24).
- [115] C. Chen, *PyTorch Face Landmark: A fast and accurate facial landmark detector*, Open-source software available at [https://github.com/cunjian/pytorch\\_face\\_landmark](https://github.com/cunjian/pytorch_face_landmark), 2021. [Online]. Available: [https://github.com/cunjian/pytorch\\_face\\_landmark](https://github.com/cunjian/pytorch_face_landmark) (cit. on p. 27).
- [116] J. W. Fernandes, “The legacy of art in plastic surgery.39 (1)”, *Plastic and Reconstructive Surgery (1946)*, vol. 9, e3519, May 2020. DOI: 10.1097/GOX.00000000000003519 (cit. on p. 28).
- [117] A. Dasgupta, A. George, S. L. Happy, and A. Routray, “A vision-based system for monitoring the loss of attention in automotive drivers”, *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, pp. 1825–1838, 2013, ISSN: 15249050. DOI: 10.1109/TITS.2013.2271052 (cit. on pp. 28, 30, 82).
- [118] K. Li, S. Wang, C. Du, Y. Huang, X. Feng, F. Zhou, *et al.*, “Accurate fatigue detection based on multiple facial morphological features”, *Journal of Sensors*, vol. 2019, 2019 (cit. on pp. 28, 30, 83).

- [119] P. A. Costa, A. Lourenço, C. Carreiras, D. Velez, P. Jorge, and A. Abrantes, “Multi-modal biometrics in vehicles: a simulator study”, 8th International Workshop on Biometrics and Forensics (IWBF2020), Apr. 2020 (cit. on p. 62).
- [120] A. Lourenço, A. P. Alves, C. Carreiras, R. P. Duarte, and A. Fred, “CardioWheel: ECG biometrics on the steering wheel”, in *Joint European conference on machine learning and knowledge discovery in databases*, Springer, 2015, pp. 267–270 (cit. on p. 62).
- [121] A. P. Alves and C. Carreiras, “CardioWheel : ECG Biometrics”, vol. 2, pp. 267–270, 2015. DOI: 10.1007/978-3-319-23461-8 (cit. on p. 62).
- [122] R. Delgado-Gonzalo, J. Parak, A. Tarniceriu, P. Renevey, M. Bertschi, and I. Korhonen, “Evaluation of accuracy and reliability of pulseon optical heart rate monitoring device”, in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2015, pp. 430–433 (cit. on p. 62).
- [123] J. Parak, A. Tarniceriu, P. Renevey, M. Bertschi, R. Delgado-Gonzalo, and I. Korhonen, “Evaluation of the beat-to-beat detection accuracy of pulseon wearable optical heart rate monitor”, in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2015, pp. 8099–8102 (cit. on p. 62).
- [124] P. A. Costa, A. Lourenço, P. M. Jorge, A. Abrantes, and P. Gamito, “Stream-based middleware platform for drowsiness detection”, *Proceedings of RECPAD 2019 (25th Portuguese Conference on Pattern Recognition)*, Oct. 2019, pp. 117–118 (cit. on pp. 62, 63).
- [125] S. Mohanty, S. V. Hegde, S. Prasad, and J. Manikandan, “Design of real-time drowsiness detection system using dlib”, in *2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, 2019, pp. 1–4. DOI: 10.1109/WIECON-ECE48653.2019.9019910 (cit. on pp. 66, 80).
- [126] L. Zhao, Z. Wang, X. Wang, and Q. Liu, “Driver drowsiness detection using facial dynamic fusion information and a dbn”, *IET Intelligent Transport Systems*, vol. 12, no. 2, pp. 127–133, 2018 (cit. on p. 66).
- [127] E. C. L. Santos, F. C. R. Figuinha, and F. Mastrocola, *Manual de eletrocardiografia - Cardiopapers*, 1st ed., E. Atheneu, Ed. 2017, ISBN: 8538807927 (cit. on p. 89).
- [128] G. Bakul and U. S. Tiwary, “Automated Risk Identification of Myocardial Infarction Using Relative Frequency Band Coefficient (RFBC) Features from ECG”, *The Open Biomedical Engineering Journal*, vol. 4, no. 1, pp. 217–222, Nov. 2010, ISSN: 1874-1207. DOI: 10.2174/1874120701004010217 (cit. on p. 88).
- [129] J. Pinto, “Continuous Biometric Identification on the Steering Wheel”, p. 197, 2017 (cit. on p. 89).

## BIBLIOGRAPHY

---

- [130] P. Healthcare, “Improving ECG Quality”, *Royal Phillips Electronics*, 2008 (cit. on p. 88).
- [131] *12 Lead Placement guide with diagram [VIDEO]*. [Online]. Available: <https://aimcardio.com/blog/12-lead-placement-guide-with-diagram/> (cit. on p. 90).
- [132] *File:EKG leads.png - Wikipedia*. [Online]. Available: [https://en.wikipedia.org/wiki/File:EKG%7B%5C\\_%7Dleads.png](https://en.wikipedia.org/wiki/File:EKG%7B%5C_%7Dleads.png) (cit. on p. 90).
- [133] A. Jorge and J. Cerca, “Fatigue and Drowsiness detection using inertial sensors and electrocardiogram”, 2018 (cit. on p. 90).



## THE *AUTOMOTIVE PROJECT*

### A.1 Introduction

The current work was integrated with the *AUTOMATIC multiMODAL drowsiness detectIOn for smart VEHICLES (AUTOMOTIVE) Project*. The *AUTOMOTIVE Project* was conducted in Portugal from October 2018 to November 2021 and led by the Institute for Systems and Computer Engineering, Technology and Science (INESC TEC), with the participation of the company CardioID Technologies Lda (CardioID), the Instituto Superior de Engenharia de Lisboa (ISEL), and the Lusophone University of Humanities and Technologies (ULHT). It was funded by the European Regional Development Fund (ERDF) through the Operational Programme for Competitiveness and Internationalization (COMPETE 2020 Programme), and by national funds through the Portuguese funding agency, Fundação para a Ciência e a Tecnologia (FCT).

The *AUTOMOTIVE Project* was an extensive study on applying signal processing and machine learning techniques for driver-specific drowsiness detection in smart vehicles. The main objectives behind the *AUTOMOTIVE Project* were the following:

- to apply digital signal and image processing technologies in adaptive automobiles, diagnosis, and maintenance;
- to monitor driver-status and detect distraction/stress;
- to develop computer vision methods for driver recognition and assisted driving;
- to apply multi-sensor fusion for driver identification and robust driver monitoring;
- to integrate human factors and cognitive science in the enhancement of vehicle and driver safety;

- to associate machine learning and data analytics with automotive systems;
- to address issues regarding security and privacy aspects for smart vehicle systems.

The current work, mainly focused on driver drowsiness monitoring, contributed to all the aforementioned objectives, apart from the last.

The *AUTOMOTIVE Project* encompassed the development of a simulated driving experience, which is presented in section A.2, in the interest of acquiring data for the project and advancing the state-of-the-art in driving research. It also integrated various contributions not presented in the current document. For a preview, Figure A.1 shows the different topics explored and how they interconnect. The *AUTOMOTIVE Project* case study research paper ([15]) offers a comprehensive view of all the research and development work conducted for the *AUTOMOTIVE Project*, highlighting its main achievements.

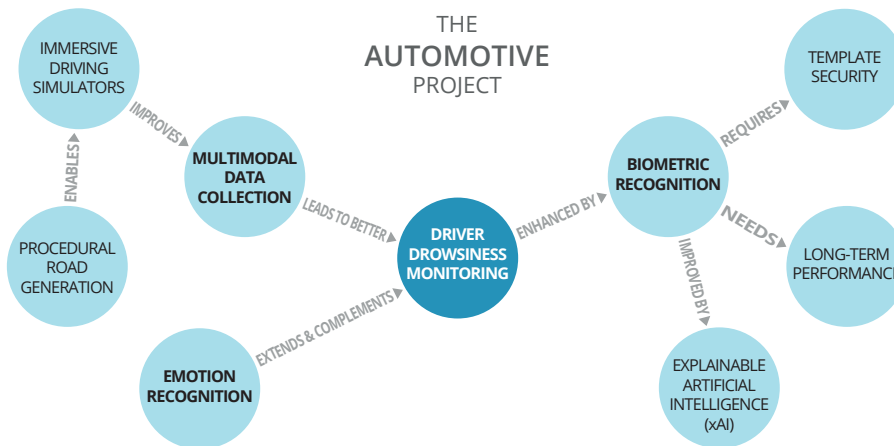



Figure A.1: A holistic view of the *AUTOMOTIVE Project*, showing the links between the explored topics and how they ultimately contribute to the central goal of improved driver drowsiness monitoring [15].

Of the four *AUTOMOTIVE Project* partners, the current work emerged mainly from the collaboration with CardioID and the Visual Computing and Machine Intelligence (VCMI) research group of INESC TEC.

### A.1.1 CardioID Technologies Lda

CardioID<sup>1</sup> is a spin-off of Instituto de Telecomunicações and Instituto Superior Técnico (University of Lisbon), born out of research on pattern recognition and machine learning methods applied to the study of biosignals. Of the various context-based applications of electrocardiogram (ECG) research, CardioID has selected the automotive industry as strategic, developing innovative products that take us closer to a non-intrusive and comprehensive driver monitoring reality and robust advanced driver-assistance systems

<sup>1</sup><https://www.cardio-id.com/> 

(ADAS). CardioID recurrently joins efforts in national and international partnerships, taking part in collaborative projects such as the i-DREAMS, Mobi, and, among others, the *AUTOMOTIVE Project*.

In the *AUTOMOTIVE Project*, CardioID was responsible for developing a driving simulator capable of mimicking the driving experience and collecting and storing data from different sensors, as well as conducting the simulated driving sessions resulting in the project's data. Furthermore, it shared with INESC TEC the leadership of the work on the automatic processing of the collected data.

### A.1.2 INESC TEC's VCMi Group

The VCMi research group <sup>2</sup> makes part of the Centre for Telecommunications and Multimedia (CTM) at INESC TEC. Its research mainly focuses on computer vision and pattern recognition and applies to endeavors such as computer-aided diagnosis, intelligent vehicle mobility, fatigue detection, breast cancer, cervical cancer, lung cancer, physiological infant monitoring, and biometric authentication.

In the *AUTOMOTIVE Project*, the VCMi research group was responsible for creating a structure with metrics to evaluate the data collected in the project's data acquisition. Furthermore, it shared with CardioID the leadership of the work on the automatic processing of the collected data.

## A.2 Data Acquisition

One of the *AUTOMOTIVE Project*'s objectives was to develop a platform to simulate a real driving experience and effectively collect multimodal data in real-time. The attributes of driving data acquired in simulated environments are approached in section 3.1. As a result, two simulator platforms were designed prior to the current work, where the main observable difference between the two relies on the immersiveness of the environment.

The first simulator, *AUTOMOTIVE Desk-Driver (DD) Simulator*, further presented in section A.2.2, runs on a simple desktop monitor with the driver seated on a chair. The second, *AUTOMOTIVE In-Half-Car (IHC) Simulator*, further presented in section A.2.3, includes a more realistic virtual driving environment, projected onto wide screens surrounding the driver. For added realism, the driver is behind the wheel of a real (parked) car cut in half. The experimental design was developed in the course of the current work - see section 3.5.

Both simulator platforms are located at CardioID's Lisbon office and integrate the same, multiple, sensory information, which is retrieved from different sources, as further explained in section A.2.1.

---

<sup>2</sup><https://vcmi.inesctec.pt/> 

### A.2.1 Sensors

The experimental setups of the *AUTOMOTIVE Project* are prepared to acquire data from the three measurement approaches presented in section 2.3: vehicle behavior, physiological signals, and driver behavior. This multimodal nature requires the setups to communicate with several sensors, each collecting data at different rates:

- **Face video** is acquired with the Intel® RealSense™<sup>3</sup> SR300 camera, which simultaneously captures an RGB channel (allowing the encoding of millions of colors), and infrared and depth video channels at 50 Hz. This camera’s software development kit (SDK) automatically extracts 78-point face landmarks [119];
- **ECG** is acquired at 1kHz with the CardioWheel [120], a latest generation device that has sensors built into a steering wheel, automatically extracts heart metrics, and recognizes the driver’s identity based on these [121]. A wristband (the PulseOn wristband) can also be added, acquiring the photoplethysmogram (PPG) at 25 Hz [122], [123], as well as a chest-band (the Movesense MD<sup>4</sup> medical sensor), acquiring heart rate with clinical precision at 512 Hz.
- **Collision-avoidance alerts** are obtained using the Mobileye® 6 Connect<sup>5</sup>, which is an ADAS that allows automatic detection of vehicles.
- **Performance telemetry** in simulated urban and highway environments (*e. g.*, dynamic signals such as acceleration and braking inputs, control of speed limits, lane following angle, obstacles, traffic violations, and collision events), is obtained by the simulation [119];

Managing these multiple data sources with specific sampling rates requires a performance-focused architecture, which the main contributor thoroughly explained in [124], and is also presented in the project’s research article( [15]).

In a nutshell, the middleware platform used for integrating inputs and outputs is Apache Kafka, which guarantees performance and redundancy in the real-time handling of the multiple data streams. This ability is paramount when dealing with biometric systems or obstacle detection alerts, both of which depend on millisecond-level response times to be effective. The inputs and outputs are integrated with a producer/consumer model, implemented using Message Queuing Telemetry Transport (MQTT) or advanced message queuing protocol (AMQP). The application programming interface (API) used for production/consumption is Unity 3D.

---

<sup>3</sup>Intel® RealSense™ Computer Vision: Depth and Tracking cameras. Available on <https://www.intelrealsense.com/> (visited on 09/11/2021).

<sup>4</sup>Movesense MD. Available on <https://www.movesense.com/wp-content/uploads/2020/12/Movesense-Medical-Spec-Sheet-2020-12-v1.0.pdf> (visited on 09/11/2021).

<sup>5</sup>Mobileye: Technical Specification Sheet. Available on <https://fleetsafe.com.au/wp-content/uploads/2019/06/Mobileye-6-Technical-Spec-v0.2-1.pdf> (visited on 09/11/2021).

A benchmarking test was conducted to test the platform’s capabilities, which it passed [124], proving that this system fulfills the requirements of a driving environment handling real-time analysis of multiple sources with variable sampling intervals.

### A.2.2 AUTOMOTIVE DD Simulator

This simulator’s graphics and physics are supported by Unity 3D, a rendering and physics engine that has gained widespread adoption in gaming and research and is featured in several state-of-the-art driving simulators. It is a general-purpose platform that enables fast development and prototyping of simulated environments while allowing for flexible scripting control of all virtual environment elements with C# or Javascript.

The designed virtual environment comprises two urban areas connected by a highway. The urban areas provide a set of complex road geometries and interaction opportunities with various agents (*e. g.*, pedestrians, traffic signs, crosswalks, intersections, buildings). The highway, in turn, provides flexibility, enabling experiments of customizable length and duration and varying road monotony. The procedural generation of the highway road is what creates this flexibility. It was implemented by adapting the MicroGSD RoadArchitect<sup>6</sup> module, which generates a static road geometry from a spline defined by user-placed nodes. The adaptation eliminated direct user inputs on road geometry and instead parametrized node location as a random distribution of 3D points. Such parameters are:

- Segment size, the longitudinal distance between consecutive nodes;
- Number of nodes, which together define the total highway length;
- Number of lanes, that can range from single to three-lane;
- Transverse position and height range: from these, a coordinate is randomly chosen, hence defining curve and elevation variations along the highway.

By defining these parameters, a highway is generated for every session, allowing us to control the experiment duration and the driving task’s complexity and ensure that infinite road geometries with equivalent difficulties are available. Furthermore, having the spline definition of road trajectory makes it possible to define a set of non-playable character (NPC) vehicles that travel the highway alongside the user. Generating cars that follow the defined splines and whose speeds are drawn from random distributions makes it possible to simulate and adjust traffic dynamics for a more realistic experience.

An example of the running simulation can be seen in Fig. A.2.

The resulting simulation is displayed on an ASUS Monitor (model VS197DE TN 18.5"16:9 60Hz FWXGA), which stands in front of the driver. The steering wheel is a Logitech G Dual-Motor Feedback Driving Force G29 Gaming Racing Wheel with Responsive Pedals. A photograph of the platform being used is shown in Fig. A.3.

---

<sup>6</sup>MicroGSD: Road Architect for Unity. Available on <https://github.com/MicroGSD/RoadArchitect>.



Figure A.2: Running simulation in the AUTOMOTIVE DD Simulator: view from inside the vehicle.



Figure A.3: Experimental setup of the AUTOMOTIVE DD Simulator during use. In the monitor, the simulation is being displayed in the view from outside the vehicle.

### A.2.3 AUTOMOTIVE IHC Simulator

Given the platform described in A.2.1 and A.2.2, the data acquisition and storage functionalities and the biometric recording, user identification, and authentication are available to be integrated into multiple simulation environments. Consecutively, a simulator developed externally to the *AUTOMOTIVE Project* by the ULHT partner was integrated to create a more immersive simulator, the AUTOMOTIVE IHC Simulator.

The ULHT simulator offers several driving scenarios for greater diversity during the

driving sessions. These include highways, suburban roads, and urban scenarios. Moreover, it includes the procedural generation of traffic events to analyze the driver's response to them. Accessing the database, the event manager can provide event timestamps, allowing the extraction of the physiological data that describes the driver's reaction to the corresponding events and, ultimately, their driving performance.

The lineup of the event occurrences must ensure some novelty across different sessions for the same user. As the route mandates a logical sequence in the highway and suburban road environments, these driving sessions are expected to include all planned events. On the other hand, to ensure a consistent experience across all users, the vehicle has a GPS available in the driving interface, indicating the route to follow in the urban environment. This environment also stores several invisible checkpoints in particular spots, such as road intersections or avenues, which trigger events according to the drivers' proximity.

Some implemented suburban event scenarios are:

- Moving emergency vehicles;
- Stopped vehicles (including after accidents);
- Vehicles breaking road laws;
- Weather, such as rain and fog.

Some urban event scenarios available in the simulator are:

- Pedestrians crossing the road (on crosswalks or not);
- Cyclists on the road;
- Traffic light intersections.

Summing up, with the integration of the ULHT simulator, an automated event generator enables the simulation of more realistic driving sessions (*e. g.*, commuting between urban and suburban environments with regular traffic events). An example of the running simulation can be seen in Fig. A.4. The resulting simulation was designed to be projected onto an angled display surrounding the driver, who is sitting on a real (parked) car with an integrated gaming steering wheel and pedals. An illustration of the planned setup can be seen in Fig. A.5.

### **A.3 Supervised Internships**

Following the collaboration with INESC TEC's CTM, during this dissertation an involvement in 2 internships also took place, specifically regarding demarcating their goals and providing guidance for methods.

### **A.3.1 Development of a Web-Application for Real-Time Drowsiness Monitoring**

This internship consisted primarily of the development of a web application, by the student Brenda Nogueira, in which a user could give access to their webcam or upload a video, and watch a drowsiness alarm system performing over the selected video input, in real-time.

This drowsiness alarm system constituted the first development approach conducted in this dissertation, and is explained as follows:

A drowsiness alarm system based on thresholds was implemented following an approach very commonly found in the Literature (for example, in the works of [46], [125], and [47]), using Dlib and OpenCV. The Dlib library was used to detect the face and predict the placement of facial landmarks, using Dlib's pre-trained facial landmark detector that places 68 landmarks within a face image.

Then, having the landmarks, the eye aspect ratio (EAR) was extracted as illustrated in figure A.6. The same approach was conducted for a mouth-opening ratio.

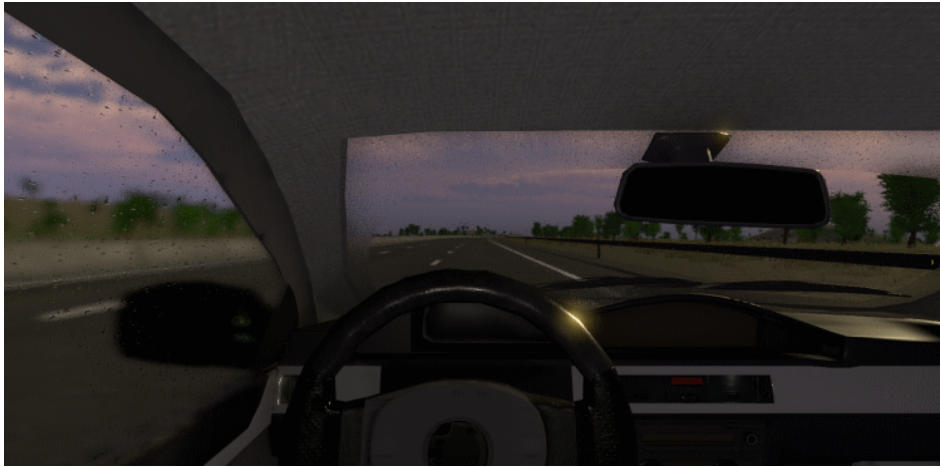
Given the real-time computation of these ratios, an alert system was activated where a siren sound played when the EAR was below a threshold for too long (over a certain time threshold), or the mouth-opening ratio was above a threshold, probably indicating a yawn.

This method was extremely quick to execute, making it usable for real-time performance, but it had these significant disadvantages:

- weak face detection (sometimes detecting more faces than there were in the frame);
- weak landmark placement, as shown in figure A.7.
- was based on thresholds which, understandably, will vary across individuals with different eye/mouth opening amplitudes.

### **A.3.2 A DBN Approach for Driver Drowsiness Detection**

This internship, performed by student Ana Maria Sousa, followed the approach from [126], using facial dynamic fusion information and a deep belief network (DBN). Ana found that the algorithms using dynamic fusion information outperformed those using single information.



(a) Running simulation on the Highway scenario in the AUTOMOTIVE IHC Simulator. Adverse weather conditions are active (rain).



(b) Running simulation on the Suburban Road scenario in the AUTOMOTIVE IHC Simulator.



(c) Running simulation on the Urban scenario in the AUTOMOTIVE IHC Simulator. Traffic lights, a pawn on the crosswalk, and a GPS indicating the route to follow can be seen.

Figure A.4: Running simulation on different scenarios in the AUTOMOTIVE IHC Simulator. Views from inside the vehicle.

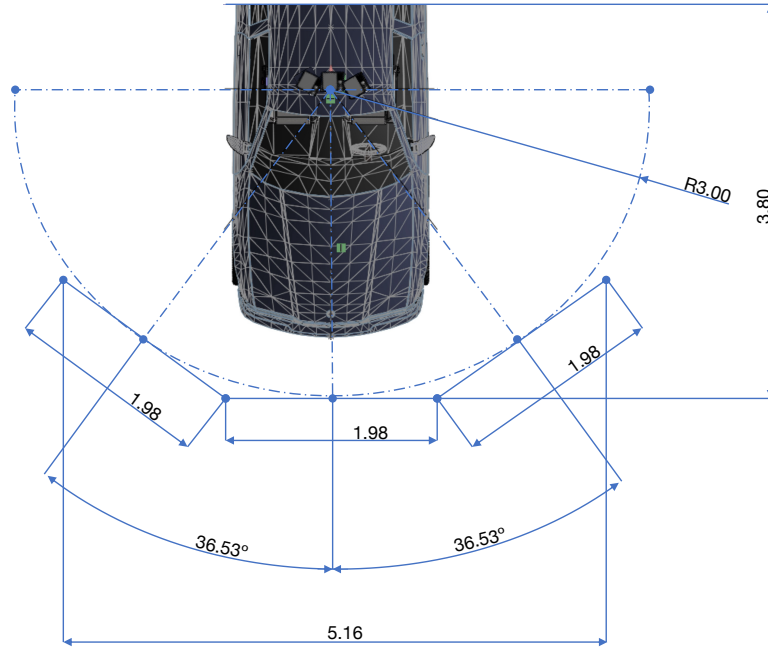


Figure A.5: Design of the experimental setup of AUTOMOTIVE IHC Simulator. 3 projectors are displayed on top of the car. The throw distances of these projectors are represented.

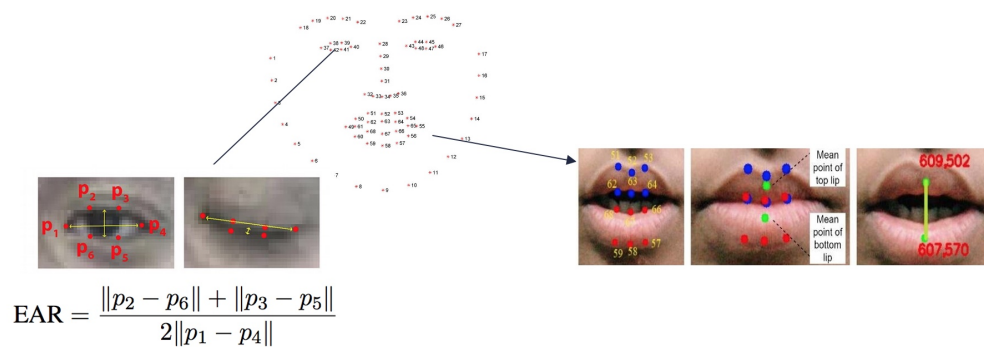


Figure A.6: Extraction of EAR and mouth landmarks performed.

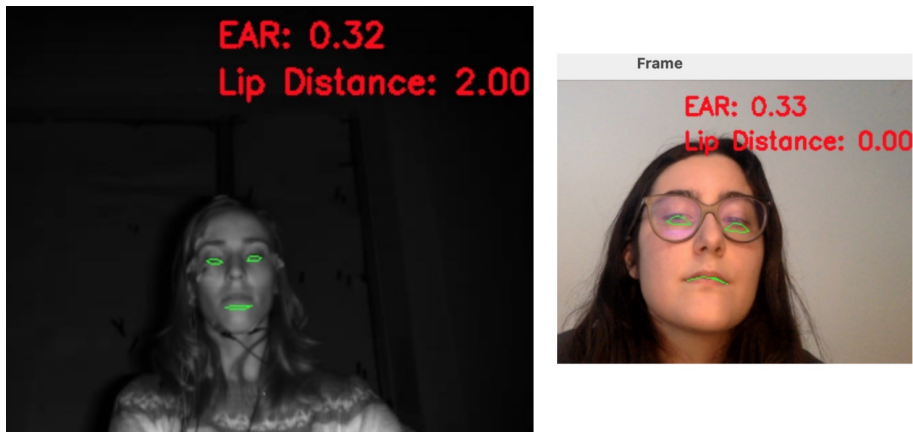
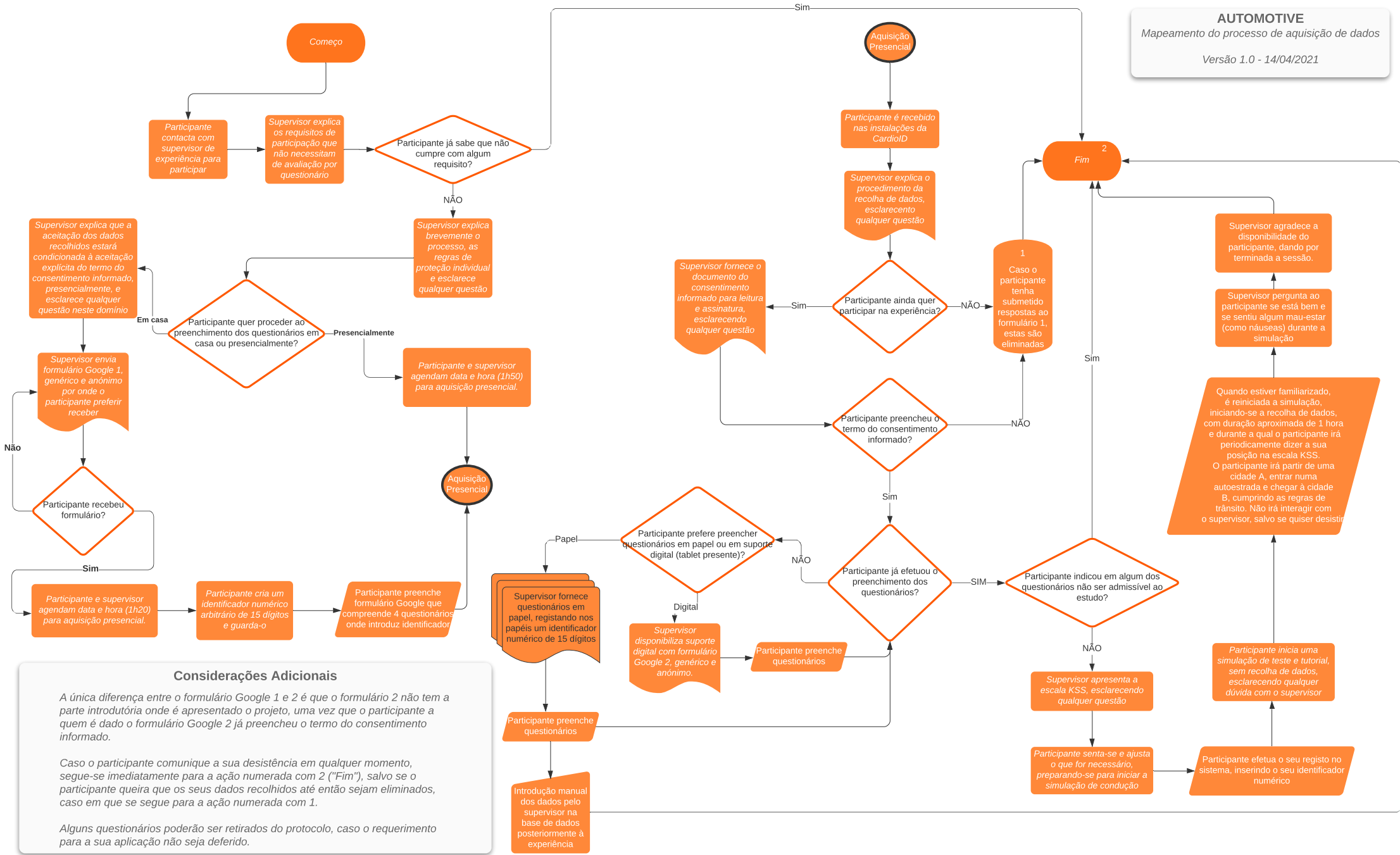


Figure A.7: An example of incorrect placement of landmarks in real-time, both in a frame retrieved from the DROZY database and in a frame retrieved from the webcam.

A P P E N D I X



## AUTOMOTIVE ACQUISITION MAPPING



A P P E N D I X



## AUTOMOTIVE ACQUISITION BRIEFING

ABRIL 2021

# SIMULADOR AUTOMOTIVE

PROTOCOLO  
DE AQUISIÇÃO  
DE DADOS

# 01

## OBJETIVO DA EXPERIÊNCIA

Adquirir dados que permitam o desenvolvimento de modelos de detecção de sonolência ao volante através de um conjunto alargado de sensores integrados no ambiente de condução, nomeadamente, uma câmara Intel® RealSense, uma banda de peito Movessense, uma pulseira pulseOn, um volante CardioWheel e um sistema Mobileye™.

## NATUREZA DOS DADOS

São recolhidos dados associados à dinâmica de condução do condutor (velocidade, aceleração e travagem, ângulo do volante), à sua fisiologia (eletrocardiograma), ao seu comportamento (abertura dos olhos, frequência do pestanejo, direção do olhar, postura), ao seu estado de sonolência (autoavaliado) e algumas informações adicionais.

O conjunto dos dados é guardado numa base de dados, sendo identificado por um número de sessão e um identificador único, que garantem a sua pseudoanonimização.

Todo o processo de aquisição de dados está condicionado pela aprovação explícita e informada do participante voluntário, à luz do Regulamento Geral sobre a Proteção de Dados.

# 02

## REQUISITOS

**Poderá participar na experiência o sujeito que, no dia da aquisição:**

- Não consumiu bebidas alcoólicas;
- Não consumiu café;
- Não consumiu chá cafeinado;
- Teve um período de descanso noturno entre 6 (seis) e 9 (nove) horas.

**Para além disso, o sujeito:**

- Tem carta de condução válida;
- Não é suscetível a enjoos, como atestado com o questionário MSSQ-Short, realizado previamente à aquisição no simulador;
- Não padece de epilepsia;
- Apresenta uma suscetibilidade a sonolência antes da aquisição, inferida através do questionário ESS em português, realizado previamente à aquisição no simulador, abaixo de 14 na Epworth Sleepiness Scale.

**03**

## PROCEDIMENTO (1/2)

### INICIÁVEL UNICAMENTE APÓS LEITURA, COMPREENSÃO E PREENCHIMENTO DO CONSENTIMENTO INFORMADO

1. O condutor (participante) deve inteirar-se do conteúdo deste protocolo, esclarecendo qualquer dúvida com o supervisor da experiência antes do início do exercício de condução simulada com aquisição de dados.
2. O condutor deve preencher o formulário de caracterização que lhe será entregue, relativo a informações adicionais, assim como realizar o questionário de tipologia circadiana de Horne and Ostberg, que também lhe será entregue.
3. O supervisor irá apresentar a escala utilizada para a autoavaliação do estado de sonolência (escala KSS) ao condutor, encontrável na página 5, devendo este inteirar-se da mesma, esclarecendo qualquer dúvida com o supervisor.
4. O condutor deve sentar-se em frente ao volante e ajustar a altura e a posição longitudinal da cadeira, assim como a posição dos pedais, de modo a estabelecer uma posição confortável para o exercício de condução.
5. Caso seja a primeira sessão do condutor, este deverá proceder ao seu registo no sistema sob orientação do supervisor.



# 04

## PROCEDIMENTO (2/2)

**O PARTICIPANTE PODE COMUNICAR A SUA DESISTÊNCIA DO ESTUDO A QUALQUER MOMENTO, SEM NECESSIDADE DE JUSTIFICAÇÃO**

6. Para familiarizar o condutor com a simulação de condução e esclarecer quaisquer dúvidas sobre a sua interação, esta será primeiramente iniciada sem aquisição de dados associada.

*A partir do início do exercício de condução com aquisição de dados e até ao seu término, nenhuma interação não prevista deverá existir entre o condutor e o supervisor ou elementos exteriores.*

7. Iniciada a aquisição de dados, o condutor irá começar a sua viagem na Cidade A, convergindo posteriormente para uma autoestrada.

8. O condutor irá permanecer na autoestrada aproximadamente 60 minutos, chegando posteriormente à Cidade B.

9. Na Cidade B o condutor deverá circular por cerca de 5 minutos até decidir terminar a aquisição.

Durante a aquisição irá ser periodicamente solicitado ao condutor que indique a sua posição atual na escala KSS. Neste momento, o condutor deverá proferir unicamente o número correspondente à sua autoavaliação.

*Durante toda a simulação o condutor deve respeitar as regras de trânsito e, especialmente, as marcações horizontais.*



# 05

## ESCALA KSS

### KAROLINSKA SLEEPINESS SCALE

A Karolinska Sleepiness Scale, ou KSS, é uma escala numérica que compreende valores de 1 a 9 (inclusive), correspondentes ao estado de sonolência do indivíduo, conforme apresentado em seguida:

1. *Em alerta total*
- 2.
3. *Em alerta*
- 4.
5. *Nem alerta nem sonolento/a*
- 6.
7. *Sonolento/a, mas sem dificuldades em estar acordado/a*
- 8.
9. *Claramente sonolento/a, a combater o sono*

Quando solicitado, o condutor deverá proferir qualquer um dos números apresentados supra, incluindo os valores intermédios, consoante a descrição que melhor representar o seu estado naquele momento.

APPENDIX



## PRIOR ART APPROACHES

Study	Data	Features	Main approach	Results	Observations
Mohanty's [125]	YawDD Dataset, for yawning videos; MRL Eye Dataset for approximately 90,000 images.	EAR; mouth aspect ratio.	Application of Dlib's shape detector to retrieve 68 facial landmarks on the input video. Computation of eye and mouth aspect ratios. Drowsiness detection based on thresholds for EAR (0.15) and mouth aspect ratio (0.83).	real-time accuracies: 82.02% for eyes and 85.44% for yawn; pre-recorded videos: 93.25% for eyes and 96.71% for yawn.	Fast enough to run in real time over a video stream.
Islam's [46]	Private video dataset.	EAR, blink frequency.	Face detection performed by the Viola-Jones method. Fatigue detection based on EAR using thresholds, followed by fatigue detection based on Eye Blinks, also using thresholds (less than 6 blinks per minute and it's considered drowsy).	98.39% accuracy in blink detection.	
Pasaribu's [47]	Private video dataset of 5 videos acquired at 30 FPS - during 105 minutes for each person, so they feel tired.	Modified-EAR (M-EAR), proposed in this study.	Threshold-based: sleepy if the number of blinks is below 10 per minute.		

Study	Data	Features	Main approach	Results	Observations
Naurois's [24]	Video recordings with drowsiness ground truths rated by trained observers.	Blink duration (average and standard deviation (std)); Blink frequency (avg and std); percentage of eye closure (PERCLOS) (avg and std); Head translation in x, y, and z (avg and std); Head rotation in x, y, and z (avg and std); Saccade frequency (avg and std); Driving time; Additional subject-specific information.	Feature Normalization [-1;1]; Ten-fold cross-validation; Grid-search; Simple feed-forward artificial neural network (ANN) to predict drowsiness-level every minute.	Drowsiness-level detection (5-class classification) with a mean square error of 0.22.	Best performance steemed from usage of driver behavioural data and additional subject-specific information.

Study	Data	Features	Main approach	Results	Observations
Danishman's [61]	JZU eye-blink database.	Eye blink duration	From the camera frame, face detection is performed, then eye detection, then eye blink duration is determined based on horizontal symmetry of the eyes. Drowsiness is detected based on the eye blink duration, applying thresholds (blink duration less than 400 ms is awake, and between 400 and 800 ms is drowsy).	Eye blink detection with 94.00% accuracy and 1.00% false positive rate.	
Dasgupta's [117]	Private dataset with 460 images used for training the Support Vector Machine (SVM).	local binary patterns (LBPs)'s histograms to classify the eye state in open or closed. Then PERCLOS retrieved over a 3 minute window interval.	Face detection through Haar-like features; eye detection through LBPs; and Principal Component Analysis (PCA); eye state classification through SVM; then retrieval of PERCLOS.	For the eye-state classification, linear SVM showed a hit rate of 98.6%, while quadratic SVM showed a hit rate of 98.6%. No machine learning drowsiness classification was performed.	Real-time face detection. Method performed over NIR lighting (similar to the setting of DROZY's videos).

Study	Data	Features	Main approach	Results	Observations
Li's [118]	Dataset built and released. Claim that there is no public database of fatigue visual images and videos.	LBP features, HOG features and LeNet features.	Examined 3 image feature extraction algorithms: LBP, HOG, and LeNet. SVM with radial basis function (RBF) kernel. Proposed REcognizing the Drowsy Expression (REDE) algorithm.	LBP achieved the best classification performances. accuracy 96.07%	

## ***DROZY DATABASE DATA STUDY***

This Appendix shows the DROZY table with the summary information of each sample, including video of origin, Karolinska Sleepiness Scale (KSS) score, and calculated driving time.

Sample	Video	Sample Frequency (FPS)	Frame Interval		KSS	Driving Time (s)	Driving set split
			start	end			
1	1-1	30	1	1976	3	0	train
2	1-1	30	15889	17865	6	596	train
3	1-2	15	1	988	6	596	test
4	1-2	15	8509	9497	7	1 229	train
5	1-3	15	1	988	7	1 229	test
6	2-1	30	1	1976	3	0	train
7	2-1	30	15923	17899	7	597	train
8	2-2	15	1	988	7	597	test
9	2-2	15	8413	9401	6	1 254	train
10	2-3	15	1	988	6	1 254	train
11	3-1	30	1	1976	2	0	train
12	3-1	30	15906	17882	3	596	train
13	3-2	15	1	988	3	596	Not Considered
14	3-2	15	7860	8848	4	1 186	train
15	3-3	15	1	988	4	1 186	train
16	4-1	30	1	1976	4	0	train
17	4-1	30	15813	17789	8	593	test
18	4-2	15	1	988	8	593	train
19	4-2	15	7475	8463	9	1 157	train
20	4-3	15	1	988	9	1 157	train

---

21	5-1	30	1	1976	3	0	train
22	5-1	30	15922	17898	7	597	train
23	5-2	15	1	988	7	597	Not considered
24	5-2	15	7187	8175	8	1 142	Not considered
25	5-3	15	1	988	8	1 142	Not considered
26	6-1	30	1	1976	2	0	train
27	6-1	30	15937	17913	3	597	test
28	6-2	15	1	988	3	597	Not considered
29	6-2	15	7864	8852	7	1 187	train
30	6-3	15	1	988	7	1 187	train
31	7-2	15	1	988	4	590	train
32	7-2	15	6924	7912	9	1 117	train
33	7-3	15	1	988	9	1 117	test
34	8-1	30	1	1976	2	0	test
35	8-1	30	15887	17863	6	595	train
36	8-2	15	1	988	6	595	train
37	8-2	15	7886	8874	8	1 187	train
38	8-3	15	1	988	8	1 187	test
39	9-2	30	1	1976	6	590	train
40	9-2	30	15932	17908	8	1 187	test
41	9-3	30	1	1976	8	1 187	train
42	10-1	30	1	1976	3	0	train
43	10-3	30	1	1976	7	1 181	train
44	11-1	30	1	1976	4	0	train
45	11-1	30	15938	17914	7	597	train
46	11-2	30	1	1976	7	597	train
47	11-2	30	15910	17886	7	1 193	test
48	11-3	30	1	1976	7	1 193	train
49	12-1	30	1	1976	2	0	train
50	13-1	30	1	1976	6	0	train
51	13-1	30	15926	17902	3	597	test
52	13-2	30	1	1976	3	597	train
53	14-1	30	1	1976	5	0	train
54	14-1	30	15907	17883	7	596	train
55	14-2	30	1	1976	7	596	train
56	14-2	30	15885	17861	8	1 191	train
57	14-3	30	1	1976	8	1 191	train

---

Table E.1: Summary information of each sample, including video of origin, KSS score, and calculated driving time.



## LANDMARK ERROR TABLE

Algorithm	min error	mean error	max error
FaceBoxes MobileNet	2,91	3,95	5,02
RetinaFace MobileNet	2,91	3,93	5,04
FaceBoxes MobileFaceNet	2,26	3,44	4,46
FaceBoxes MobileNet	2,37	3,35	4,18

Table F.1: Minimum, mean, and maximum distribution of error (by Euclidian distance of pixels), by algorithms applied, over all samples.

Algorithm	Facial subsection	min error	mean error	max error
MobileFaceNet_FaceBoxes	Face Outline	2.67	6.58	11.80
MobileFaceNet_FaceBoxes	EyeBrow Right	2.06	3.47	4.72
MobileFaceNet_FaceBoxes	EyeBrow Left	1.44	2.76	5.02
MobileFaceNet_FaceBoxes	Nose	2.07	3.25	5.24
MobileFaceNet_FaceBoxes	Eye Right	2.55	3.38	4.62
MobileFaceNet_FaceBoxes	Eye Left	1.99	2.83	4.21
MobileFaceNet_FaceBoxes	Mouth Ext	2.02	2.79	4.16
MobileFaceNet_FaceBoxes	Mouth Int	2.33	3.26	5.93
MobileFaceNet_Retinaface	Face Outline	2.65	6.51	11.57
MobileFaceNet_Retinaface	EyeBrow Right	2.09	3.46	4.71
MobileFaceNet_Retinaface	EyeBrow Left	1.46	2.68	4.79
MobileFaceNet_Retinaface	Nose	2.11	3.27	5.26
MobileFaceNet_Retinaface	Eye Right	2.53	3.38	4.65

---

MobileFaceNet_Retinaface	Eye Left	1.98	2.84	4.19
MobileFaceNet_Retinaface	Mouth Ext	2.04	2.79	4.21
MobileFaceNet_Retinaface	Mouth Int	2.31	3.25	5.88
MobileNet_FaceBoxes	Face Outline	2.60	5.10	8.23
MobileNet_FaceBoxes	EyeBrow Right	1.39	3.13	5.32
MobileNet_FaceBoxes	EyeBrow Left	1.20	2.93	5.80
MobileNet_FaceBoxes	Nose	2.20	3.01	4.82
MobileNet_FaceBoxes	Eye Right	2.20	2.97	4.30
MobileNet_FaceBoxes	Eye Left	1.88	2.59	3.77
MobileNet_FaceBoxes	Mouth Ext	1.83	2.70	4.38
MobileNet_FaceBoxes	Mouth Int	1.84	3.01	5.07
MobileNet_Retinaface	Face Outline	2.50	4.82	7.05
MobileNet_Retinaface	EyeBrow Right	1.43	3.14	5.31
MobileNet_Retinaface	EyeBrow Left	1.13	2.85	5.69
MobileNet_Retinaface	Nose	2.21	2.97	4.79
MobileNet_Retinaface	Eye Right	2.30	3.05	4.35
MobileNet_Retinaface	Eye Left	1.86	2.56	3.72
MobileNet_Retinaface	Mouth Ext	1.82	2.65	4.34
MobileNet_Retinaface	Mouth Int	1.81	2.98	5.12

---

Table F.2: Minimum, mean, and maximum distribution of error (by Euclidian distance of pixels), by facial subsection and algorithms applied, over all samples.



## MORE ON THE ECG

Electrocardiography is defined as the process of sensor recording the electrical activity of the heart, usually displaying it in a graph of micro voltage over time. This is done by measuring small electric changes created by the heart muscle tissue when its chambers are contracting and relaxing, during each cardiac cycle. Hence, a physiological signal is recorded. The heart has 4 chambers, namely, two atriums and two ventricles, anatomically organized as represented in figure G.1.

The ECG, the graph produced by electrocardiography, is widely used in medicine to assess cardiovascular health and to alert for a variety of possible heart diseases.

The signal is created due to the successive contractions and relaxations of the myocardium, which in turn create different depolarization and repolarization directions that are registered. As such, the ECG is defined by a periodic and very specific wave format, described in figure G.2, which is usually characterized by five different types of waves: P, Q, R, S, T, and in some cases a sixth wave, denominated as U. The health assessment is done by analyzing the ECG and, more specifically, by detecting variations on the properties of the electrocardiographic signal [128].

The acquisition of an ECG can be done either in an intrusive or a non-intrusive methodology. The intrusive acquisition method is done in a clinical setting, where several electrodes are attached to the human body. Electrodes are sensors that require the use of a conductive material to enable physical contact between them and the skin. This is usually the procedure most commonly used in a medical setting, as it accurately captures the electrocardiographic signal, given that the electrode type, application, and skin preparation are correct [130]. Nonetheless, this process has the limitation of requiring a controlled physical environment as an ambulance or a treatment room. In this type of acquisition configuration, 10 electrodes are attached, creating 12 ECG-Leads, as represented in figure G.3. This means that, under some circumstances, the placement of up

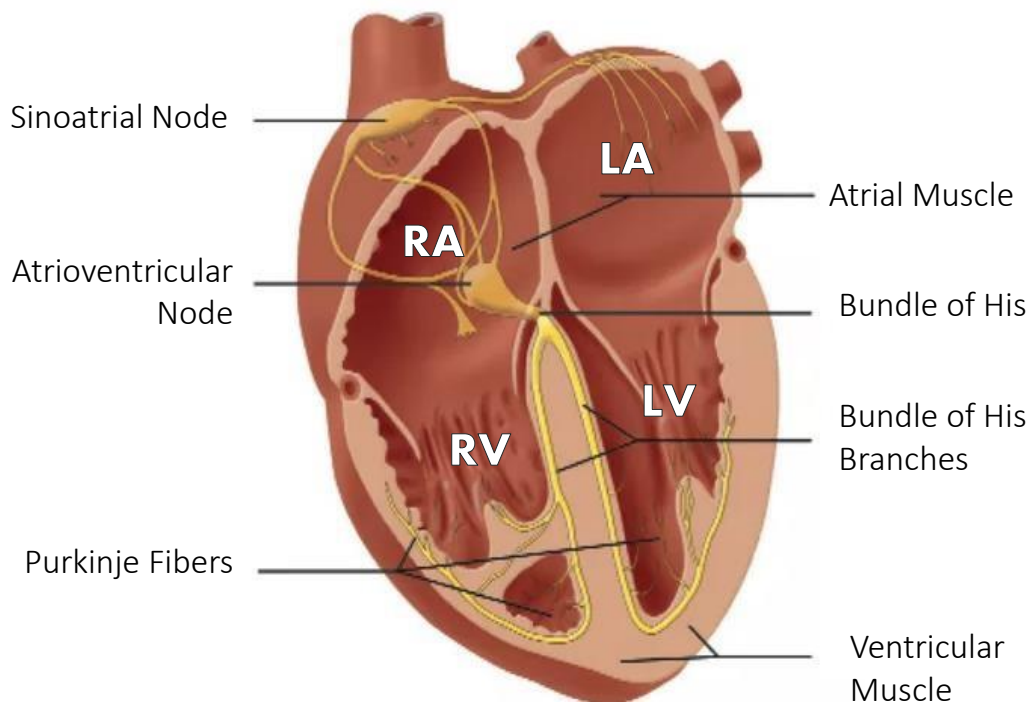


Figure G.1: Abbreviation's: RA - Right Atrium; LA - Left Atrium; LV - Left Ventricle; RV - Right Ventricle. Adapted of [127]

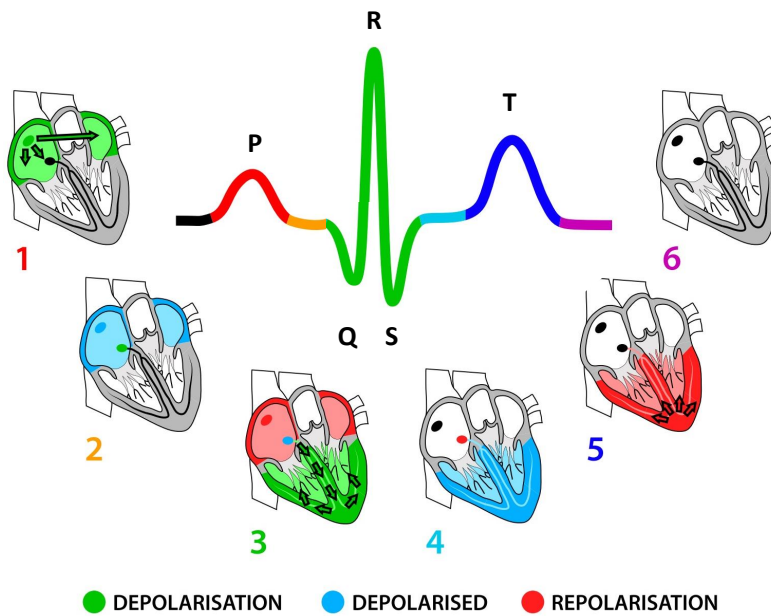


Figure G.2: Electrical conduction events in the heart. Diagrams 1-6 are color coded to match the respective phase on the ECG signal. Another color code is used to indicate the polarisation state in each heart cavity. [129]

to twelve electrodes is required. All these limitations are serious disadvantages for the acquisition of the ECG under a surveillance context, like the one this work will focus on, as it may affect the action performance of the individuals.

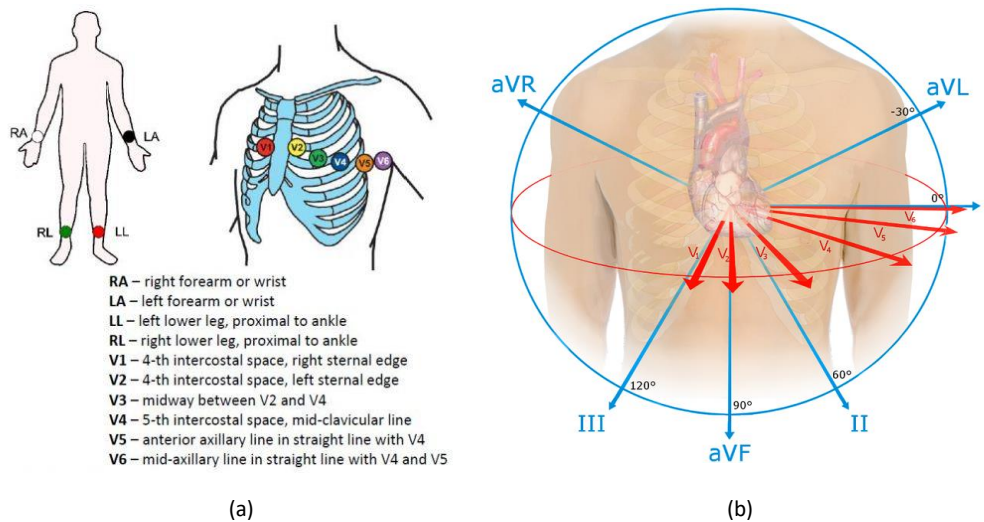


Figure G.3: (a) Schematic representation of the placement in the human body of the 10 electrodes as well as their names, under an invasive methodology (b) Schematic representation of the directions of each of the 12 ECG leads along the human body, with aVR, aVL, aVF, V1, V2, V3, V4 and V5 being unipolar and I, II and III being bipolar. Moreover the arrows in color red correspond to precordial leads, and the arrows in blue indicate the limb leads [131], [132]

On the other hand, non-intrusive methods acquire data without the sensors needing to be placed in the body. Instead, these are placed on objects which are used recurrently and are in contact with the human subject. These sensors take advantage of human perspiration to acquire signals through the use of dry electrodes, electrodes which do not require a conductor gel material to acquire. Non-intrusive ECG acquisition hence has the very significant advantage of not interfering with the human daily actions, despite being more susceptible to noise when compared with intrusive methods, and consequently requiring a more complex pre-processing step. [133]

## QUESTIONNAIRES FOR THE AUTOMOTIVE ACQUISITION PROTOCOL

Please find the following questionnaires enclosed in the following pages:

- The Horne and Ostberg's Morningness-Eveningness Questionnaire (MEQ);
- The Epworth Sleepiness Scale (ESS);
- The Immersive Tendencies Questionnaire (ITQ);
- The Subject Characterization Questionnaire;
- The Motion Sickness Susceptibility Questionnaire (MSSQ)-short;
- The System Usability Scale (SUS).

Notably, they are written in Portuguese, as this is the mother tongue of the designed experience's target group.

# Questionário de Horne e Ostberg

Este questionário destina-se à aplicação do "Morningness-Eveningness Type Questionnaire" de Horne e Ostberg, adaptado para Português de Portugal por Silvério, Silva e Macedo, em 1998.

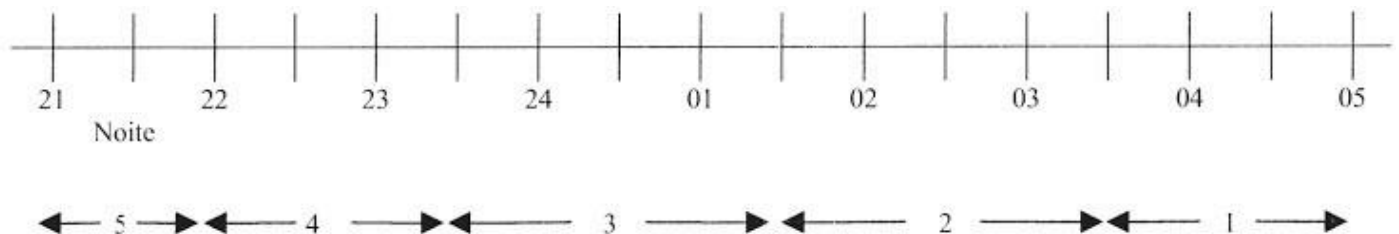
Este questionário é uma ferramenta utilizada para aferir a tipologia circadiana de um indivíduo.

Por favor responda o melhor que conseguir a todas as 16 (dezasseis) alíneas.

## \*Obrigatório

1. Considerando apenas o seu bem-estar pessoal e tendo liberdade total para planear a sua noite, a que horas se deitaria? \*

Por favor selecione o número, de 1 a 5, correspondente à sua resposta, tendo em conta a imagem abaixo em que a escala horizontal representa determinadas horas num dia.



Marcar apenas uma oval.

- 1
- 2
- 3
- 4
- 5

2. Até que ponto precisa do despertador para acordar a uma determinada hora de manhã? \*

Marcar apenas uma oval.

- Não preciso
- Preciso poucas vezes
- Preciso muitas vezes
- Preciso sempre

3. Quando está em boas condições mentais, físicas e ambientais (ex.: temperatura do quarto agradável), com que facilidade acha que se levanta de manhã? \*

*Marcar apenas uma oval.*

- Nada fácil
- Pouco fácil
- Fácil
- Muito fácil

4. Na primeira meia hora depois de ter acordado de manhã, em que medida se sente desperto/a? \*

*Marcar apenas uma oval.*

- Nada desperto
- Pouco desperto
- Desperto
- Muito desperto

5. Depois de acordar, como é o seu apetite durante a primeira meia hora? \*

*Marcar apenas uma oval.*

- Muito mau apetite
- Mau apetite
- Bom apetite
- Muito bom apetite

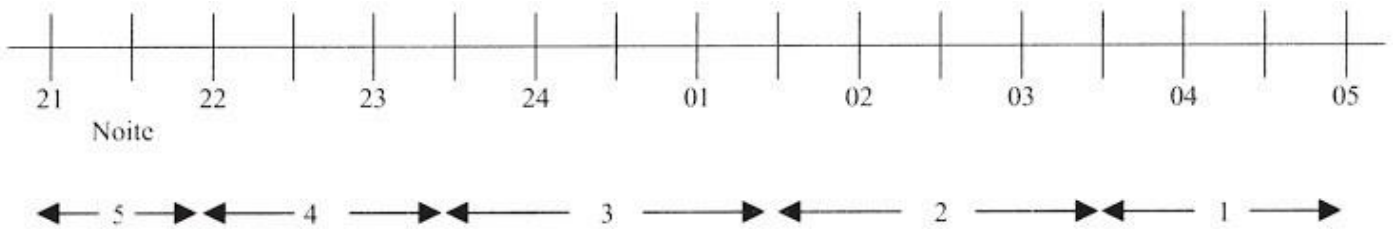
6. Depois de ter acordado de manhã, em que medida se sente cansado/a na primeira meia hora? \*

*Marcar apenas uma oval.*

- Muito cansado
- Cansado
- Fresco
- Muito fresco

7. Habitualmente, a que horas da noite se sente cansado/a e com vontade de dormir? \*

Por favor selecione o número, de 1 a 5, correspondente à sua resposta, tendo em conta a imagem abaixo em que a escala horizontal representa determinadas horas num dia.



Marcar apenas uma oval.

- 1
- 2
- 3
- 4
- 5

8. Quer estar no MÁXIMO DA SUA FORMA para fazer um teste que dura duas horas e sabe que é mentalmente cansativo. Considerando o seu bem-estar pessoal e com liberdade total para planear o seu dia, qual destes horários escolheria para fazer esse teste? \*

Marcar apenas uma oval.

- Das 08h00 às 12h00
- Das 12h00 às 16h00
- Das 16h00 às 20h00
- Das 20h00 às 24h00

9. Depois de um dia normal, qual seria o seu nível de cansaço se tivesse de se deitar às 23h00? \*

Marcar apenas uma oval.

- Nada cansado
- Cansado
- Muito cansado
- Extremamente cansado

10. Por alguma razão, foi dormir várias horas mais tarde do que é habitual. Se no dia seguinte não tiver hora certa para acordar, o que acontece? \*

*Marcar apenas uma oval.*

- Acordo à hora habitual, sem sono
- Acordo à hora habitual, com sono
- Acordo à hora habitual, mas adormeço novamente
- Acordo mais tarde do que é habitual

11. Se tiver de ficar acordado das 04h00 às 06h00 horas da manhã para realizar uma tarefa e não tiver nenhum compromisso no dia seguinte, o que faz? \*

*Marcar apenas uma oval.*

- Durmo só depois de realizar a tarefa
- Tiro uma soneca antes da tarefa e durmo depois
- Durmo bastante, antes de realizar a tarefa, e tiro uma soneca depois
- Só durmo antes de realizar a tarefa

12. Imagine que tem duas horas de exercício físico pesado para realizar. Considerando apenas o seu bem-estar pessoal e tendo total liberdade para planear o seu dia, qual destes horários escolheria? \*

*Marcar apenas uma oval.*

- Das 07h00 às 11h00
- Das 11h00 às 15h00
- Das 15h00 às 19h00
- Das 19h00 às 23h00

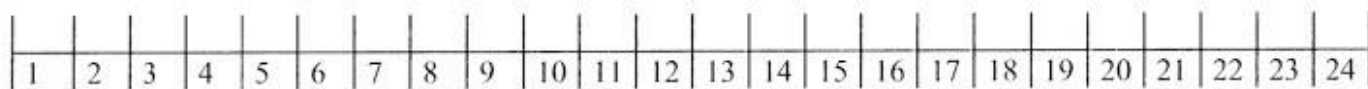
13. Decidiu fazer exercício físico "duro". Um amigo sugeriu o horário das 22h00 às 23h00, duas vezes por semana. Considerando apenas o seu bem-estar pessoal, como acha que seria o seu desempenho se fizesse exercício entre as 22h00 e as 23h00? \*

*Marcar apenas uma oval.*

- Seria excelente
- Seria bom
- Seria mau
- Seria muito mau

14. Suponha que poderia escolher o seu próprio horário de trabalho e que deveria trabalhar 5 (cinco) horas seguidas por dia, sendo possível efetuar pequenos intervalos nesse período. Imagina que seja um serviço interessante e que você ganharia pelos resultados da produção. A que horas começaria a trabalhar? \*

Por favor selecione o número, de 1 a 5, correspondente à sua resposta, tendo em conta a imagem abaixo em que a escala horizontal representa determinadas horas num dia.

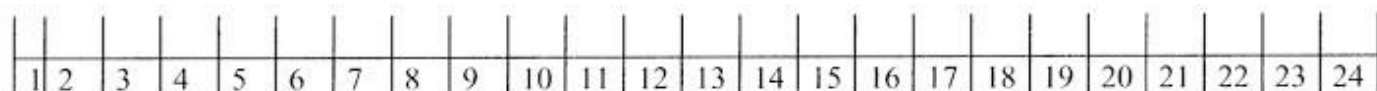


Marcar apenas uma oval.

- 1
- 2
- 3
- 4
- 5

15. Em termos de bem-estar geral a que hora do dia se sente no seu melhor? \*

Por favor selecione o número, de 1 a 5, correspondente à sua resposta, tendo em conta a imagem abaixo em que a escala horizontal representa determinadas horas num dia.



Marcar apenas uma oval.

- 1
- 2
- 3
- 4
- 5

16. Ouve-se dizer que há pessoas que funcionam melhor de manhã (tipo matutino) e pessoas que funcionam melhor à tarde/noite (tipo vespertino). Qual destes tipos acha que é? \*

*Marcar apenas uma oval.*

- Sem dúvida do tipo matutino
- Mais matutino do que vespertino
- Mais vespertino do que matutino
- Sem dúvida do tipo vespertino

---

Este conteúdo não foi criado nem aprovado pela Google.

Google Formulários

# Questionário da Escala de Sonolência de Epworth

Este questionário destina-se à extração da sua posição na Escala de Sonolência de Epworth.

O questionário da Escala de Sonolência de Epworth é uma ferramenta utilizada para aferir níveis excessivos de sonolência diurna.

Por favor responda o melhor que conseguir a todas as 8 (oito) alíneas.

**\*Obrigatório**

Este questionário refere-se ao seu modo de vida habitual nos últimos tempos. Mesmo que não tenha feito algumas destas coisas ultimamente, tente imaginar como é que elas o/a afetariam.

1. Qual a probabilidade de dormirar ("passar pelas brasas") ou de adormecer - e não apenas sentir-se cansado/a - nas seguintes situações? \*

*Marcar apenas uma oval por linha*

	Nenhuma probabilidade de dormirar	Ligeira probabilidade de dormirar	Moderada probabilidade de dormirar	Forte probabilidade de dormirar
Sentado/a a ler	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A ver televisão	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sentado/a inativo/a num lugar público (por exemplo, sala de espera, cinema ou reunião)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Como passageiro num carro durante uma hora, sem paragem	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Deitado/a a descansar à tarde quando as circunstâncias o permitem	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sentado/a a conversar com alguém	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sentado/a calmamente depois de um almoço sem ter bebido álcool	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ao volante, parado/a no trânsito durante uns minutos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Este conteúdo não foi criado nem aprovado pela Google.

# Immersive Tendencies Questionnaire

Este questionário destina-se à aplicação do "Immersive Tendencies Questionnaire" de Witmer e Singer, em Português.

Indique a sua resposta preferencial, assinalando a opção adequada da escala de sete pontos. Tenha em conta toda a escala ao dar as suas respostas, uma vez que os níveis intermédios se podem aplicar. Por exemplo, se a sua resposta for uma ou duas vezes, deve assinalar '2'. Se a sua resposta for muitas vezes, mas não com extrema frequência, então deve assinalar '6'.

Por favor responda o melhor que conseguir a todas as 18 (dezoito) alíneas.

**\*Obrigatório**

1. Fica facilmente envolvido/a em filmes ou séries de televisão? \*

*Marcar apenas uma oval.*

	1	2	3	4	5	6	7	
Nunca	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Frequentemente

2. Alguma vez se envolveu tanto num programa ou num livro que as pessoas têm dificuldade em captar a sua atenção? \*

*Marcar apenas uma oval.*

	1	2	3	4	5	6	7	
Nunca	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Frequentemente

3. Quão mentalmente alerta se sente neste momento? \*

*Marcar apenas uma oval.*

	1	2	3	4	5	6	7	
Nada alerta	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Completamente alerta





14. Com que frequência joga jogos de arcada ou de vídeo? (Entenda-se por frequentemente todos os dias ou de dois em dois dias, em média). \*

*Marcar apenas uma oval.*

	1	2	3	4	5	6	7	
Nunca	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Frequentemente

15. Alguma vez se entusiasmou durante uma cena de perseguição ou de luta na televisão ou no cinema? \*

*Marcar apenas uma oval.*

	1	2	3	4	5	6	7	
Nunca	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Frequentemente

16. Alguma vez se assustou com algo num programa de televisão ou num filme? \*

*Marcar apenas uma oval.*

	1	2	3	4	5	6	7	
Nunca	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Frequentemente

17. Alguma vez ficou apreensivo ou com medo muito tempo depois de ver um filme de terror? \*

*Marcar apenas uma oval.*

	1	2	3	4	5	6	7	
Nunca	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Frequentemente

18. Alguma vez se envolveu tanto a fazer algo que perdeu a noção do tempo? \*

*Marcar apenas uma oval.*

	1	2	3	4	5	6	7	
Nunca	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Frequentemente

---

Este conteúdo não foi criado nem aprovado pela Google.

# Motion Sickness Susceptibility Questionnaire Short-Form

Este questionário destina-se à aplicação do "Motion Sickness Susceptibility Questionnaire Short-Form" em Português (MSSQ-short), para prevenir episódios de cinetose (conjunto de sintomas associados ao enjoo e provocados por ambientes em movimento) na sua participação na componente presencial do estudo.

O MSSQ-short é um questionário que afere os movimentos responsáveis por iniciar sintomas de náusea e tonturas, considerando a sua relação com nove meios de transporte e entretenimento, atualmente e na sua infância.

Por favor responda o melhor que conseguir a todas as 18 (dezoito) alíneas.

**\*Obrigatório**

1. Na sua infância (antes dos 12 anos), quão frequentemente se sentiu adoentado/a ou com náuseas nas seguintes situações? \*

*Marcar apenas uma oval por linha.*

	Não aplicável (nunca viajei ou participei na ação descrita)	Nunca ficava enjoado/a	Raramente ficava enjoado/a	Às vezes ficava enjoado/a	Frequentemente ficava enjoado/a
Numa viagem de carro	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Numa viagem de autocarro	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Numa viagem de comboio	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Numa viagem de avião	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Numa viagem num barco pequeno	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Numa viagem num navio ou cacilheiro	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nos baloiços do parque infantil	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Na roda/gira-gira/chávena do parque infantil	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Em montanhas-russas ou grandes carrosséis	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2. Ao longo dos últimos 10 anos, quão frequentemente se sentiu adoentado/a ou com náuseas nas seguintes situações? \*

Marcar apenas uma oval por linha.

	Não aplicável (nunca viajei ou participei na ação descrita)	Nunca ficava enjoado/a	Raramente ficava enjoado/a	Às vezes ficava enjoado/a	Frequentemente ficava enjoado/a
Numa viagem de carro	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Numa viagem de autocarro	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Numa viagem de comboio	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Numa viagem de avião	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Numa viagem num barco pequeno	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Numa viagem num navio ou cacilheiro	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nos baloiços do parque infantil	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Na roda/gira-gira/chávena do parque infantil	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Em montanhas-russas ou grandes carrosséis	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Este conteúdo não foi criado nem aprovado pela Google.

Google Formulários

# Caracterização

Este questionário destina-se a recolher informação adicional sobre si com vista a caracterizá-lo, não sendo usável para o/a identificar.

Por favor responda da forma mais verosímil que conseguir.

**\*Obrigatório**

1. Idade (anos) \*

---

2. Sexo \*

*Marcar apenas uma oval.*

Feminino

Masculino

Outra: \_\_\_\_\_

3. Como avalia a qualidade do seu sono numa escala de 1 a 10? \*

*Marcar apenas uma oval.*

1      2      3      4      5      6      7      8      9      10

---

Péssima           Excelente

---

4. Com que frequência consome cafeína atualmente? \*

*Marcar apenas uma oval.*

Nunca

Raramente

1 ou 2 cafés por dia

Mais de 2 cafés por dia

5. Com que frequência conduz, num ano típico? \*

*Marcar apenas uma oval.*

- Ocasionalmente
- Várias vezes por mês
- Várias vezes por semana
- Várias vezes por dia

6. Como avalia a sua relação com o exercício de condução atualmente? \*

*Marcar apenas uma oval.*

- Extremamente desconfortável a conduzir
- Algum desconforto a conduzir
- Razoável desconforto a conduzir
- Total confiança a conduzir

7. Se usar pacemaker ou tomar alguma medicação que influencie o seu ritmo cardíaco, por favor indique-os.

Esta questão não é obrigatória. Caso não se aplique a si, por favor deixe em branco.

---

8. Se tiver algum distúrbio do sono diagnosticado, por favor indique-o.

Esta questão não é obrigatória. Caso não se aplique a si, por favor deixe em branco.

---

9. Horário de trabalho (ou equiparável) típico.

Se trabalhar por turnos por favor indique-o, referindo também se faz turnos durante a noite. Com "equiparável" pretende-se dar a entender que esta questão está focalizada na sua principal atividade diária, admitindo como resposta não só um horário laboral, como também horário de aulas ou de voluntariado, por exemplo, se for o que mais se aplica a si. Esta questão não é obrigatória. Caso não se aplique a si, por favor deixe em branco.

---

---

---

---

---

---

Este conteúdo não foi criado nem aprovado pela Google.

Google Formulários

# System Usability Scale (SUS)

Este questionário destina-se à aplicação da System Usability Scale (SUS), adaptada para Português de Portugal por Ana Isabel Martins et al. (2015), seguindo as linhas orientadoras estabelecidas internacionalmente, de forma a assegurar a qualidade da tradução resultante e a equivalência semântica, garantido a consistência do significado dos conceitos entre esta versão e a original.

**\*Obrigatório**

1. Por favor responda o melhor que conseguir a todas as 10 (dez) alíneas. \*

*Marcar apenas uma oval por linha.*

	1. Discordo Fortemente	2	3	4	5. Concordo Fortemente
1. Acho que gostaria de utilizar este produto com frequência.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Considerei o produto mais complexo do que necessário.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Achei o produto fácil de utilizar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Acho que necessitaria de ajuda de um técnico para conseguir utilizar este produto.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Considerei que as várias funcionalidades deste produto estavam bem integradas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Achei que este produto tinha muitas inconsistências.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Suponho que a maioria das pessoas aprenderia a utilizar rapidamente este produto.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. Considerei o produto muito complicado de utilizar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. Senti-me muito confiante a utilizar este produto.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. Tive que aprender muito antes de conseguir lidar com este produto.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Este conteúdo não foi criado nem aprovado pela Google.

Google Formulários

