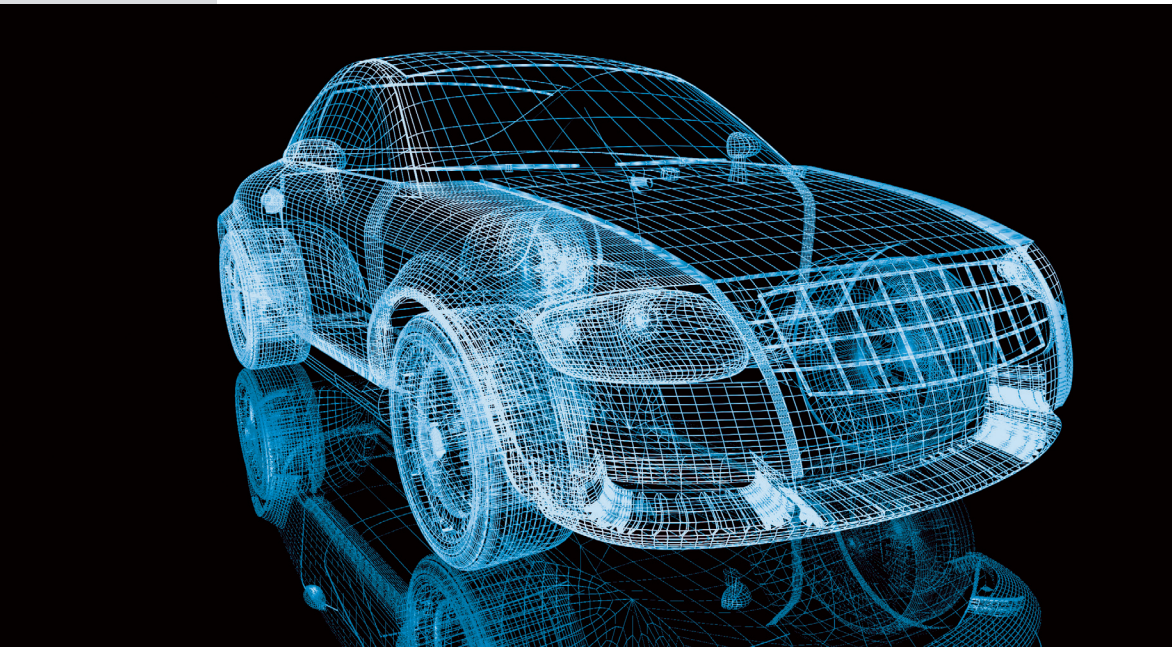
VERLAG DER TECHNISCHEN UNIVERSITÄT GRAZ

Sadegh Arefnezhad

Evaluation of Driver Performance in Semi-automated Driving by Physiologic, Driver Behaviour and Video-based Sensors

FTG 10

MONOGRAPHIC SERIES TU GRAZ
REIHE FAHRZEUGTECHNIK / AUTOMOTIVE SERIES



Sadegh Arefnezhad

Evaluation of Driver Performance in Semi-automated Driving by Physiologic, Driver Behaviour and Video-based Sensors

Monographic Series TU Graz

Reihe Fahrzeugtechnik / Automotive Series

Series Editor:

Institute of Automotive Engineering

Monographic Series TU Graz

Reihe Fahrzeugtechnik / Automotive Series **Volume 10**

Sadegh Arefnezhad

Evaluation of Driver Performance in Semi-automated Driving by Physiologic, Driver Behaviour and Video-based Sensors

---

This work is based on the dissertation *"Evaluation of Driver Performance in Semi-automated Driving by Physiologic, Driver Behaviour and Video-based Sensors"*, presented at Graz University of Technology, Institute of Automotive Engineering in 2021.

Supervision / Assessment:
Arno Eichberger (Graz University of Technology)
Wen-Hua Chen (Loughborough University)

Cover photo	'concept car' by ArchMen (modified) Fotolia / Adobe Stock
Cover layout	Verlag der Technischen Universität Graz
Print	DATAFORM Media Ges.m.b.H.

www.tugraz-verlag.at

Print
ISBN 978-3-85125-862-2

E-Book
978-3-85125-863-9
DOI 10.3217/978-3-85125-862-2

Acknowledgement

Throughout the writing of this dissertation, I have received a great deal of support and assistance.
I would first like to thank my supervisor, Assoc. Prof. Arno Eichberger, whose expertise was invaluable in formulating the research questions and methodology.
I would also like to thank my second examiner, Prof. Wen-Hua Chen from the Department of Aeronautical and Automotive Engineering, at Loughborough University, UK. His valuable comments helped me to improve my thesis and remove its shortcomings.
I would like to acknowledge my colleagues Dr. Sajjad Samiee, Dipl.-Ing. Matthias Frühwirth, Priv.-Doz. Dr.rer.nat. Ioana Victoria Koglbauer, and B.Sc. Clemens Kaufmann for their great help and support to prepare the data set and scientific publications.
This dissertation is written in the structure of the WACHSens project that was carried out by a consortium including Human Research Institut für Gesundheit-stechnologie und Präventionsforschung GmbH, Graz University of Technology, AVL Powertrain UK Limited, and Factum apptec ventures GmbH. This project was co-funded by the Austrian Research Promotion Agency (FFG) via the Future Mobility Program (grant no. 860875). I am indebted to all drivers who participated in the experiments, the experimenters, and the many people who helped set up the tests.
Finally, I would like to thank my family for their wise counsel and sympathetic ear. You are always there for me.

Abstract

Drowsy driving is an important cause of road accidents that can lead to many fatalities and monetary losses. Moreover, in the upcoming SAE level 3 (conditionally automated driving), the state of the drivers must be monitored since the driver must be attentive to drive manually when the automated driving system cannot control the car any more. To detect driver drowsiness, three data sources have been generally used in the literature: vehicle-based data, facial-based data, and biosignals. Recent studies mostly focused on designing driver drowsiness detection systems using binary classifiers that report the driver's vigilance into two classes, alert and drowsy. However, adding a middle level of drowsiness can help better estimate the transition between alertness and drowsiness to warn the driver early enough to prevent impaired driving. In addition, previous works mainly concentrated on driver drowsiness detection in manual driving mode, whereas there is no input from the driver in SAE level 3 automated driving. Therefore, the drowsiness detection system cannot utilize vehicle-based data to estimate the drowsiness in automated driving. To address the issues of the previous works, this thesis proposes three new approaches to classify driver drowsiness in simulated driving tests:

(1) The first approach is data fusion of vehicle-based data, electrocardiogram (ECG) signals, and facial-based data (eyelid opening and pupil diameter) using traditional machine learning methods. The input data are preprocessed, their statistical features are extracted, a feature selection algorithm is applied to remove the redundant and irrelevant features. Finally, random forest and K-nearest neighbors are employed to classify the drowsiness into three classes of alert, moderately drowsy, and extremely drowsy. Results show that data fusion of the different data sources applying the feature selection method outperforms the systems using only one individual data source and the system without a feature selection algorithm.

(2) The second approach is based on training deep convolutional neural networks using ECG signals and eyelid opening data. One neural network is constructed for every input data where the soft voting ensemble method is also applied to ensemble the results of the two different neural networks. The Bayesian optimization method is employed to optimize the hyperparameters of the neural networks. Outcomes present that deep convolutional neural networks significantly outperform the traditional machine learning methods applied to the same input data. Moreover, the results show that by using only ECG signals, a balanced accuracy of about 80% is achieved for drowsiness classification.

(3) The third approach is based on electroencephalogram (EEG) data processing for drowsiness detection. The percentage of eyelid closure (PERCLOS) is employed as ground truth for driver drowsiness, and a neural encoder-decoder modeling framework is proposed to select the essential EEG features that encode the PERCLOS progression in drivers. Results show that this framework obtains satisfying performance with an average RMSE of 0.11 to estimate the real PERCLOS. This framework also provides a set of EEG features that are consistently significant across different drivers for drowsiness estimation.

Kurzfassung

Schläfrigkeit von Fahrer*innen ist eine bedeutende Ursache für schwere Verkehrsunfälle, die zu Todesfällen und anderen Verlusten führen können. Darüber hinaus muss in der kommenden SAE-Stufe 3 (bedingt automatisiertes Fahren) der Zustand des Fahrers überwacht werden, da der Fahrer übernahmebereit sein muss, wenn der Fahrroboter die Fahrzeugführung nicht mehr selbst beherrscht. Zur Erkennung der Müdigkeit des Fahrers wurden in der Literatur im Allgemeinen drei Datenquellen verwendet: Fahrzeugdaten, gesichtsbasierte Merkmale und Biosignale. Wissenschaftliche Studien konzentrierten sich hauptsächlich auf die Entwicklung von Systemen die binäre Klassifikatoren verwendeten und die Vigilanz in zwei Klassen einteilten: wach und schläfrig. Die Hinzunahme eines mittleren Schläfrigkeitsgrades kann jedoch dazu beitragen, den Übergang zwischen Wachsamkeit und Schläfrigkeit besser einzuschätzen, um den Fahrer rechtzeitig über bevorstehende Fahruntüchtigkeit zu warnen. Zusätzlich konzentrierten sich bisherige Arbeiten auf die Erkennung von Müdigkeit beim manuellen Fahren, während beim automatisierten Fahren nach SAE Level 3 keine Eingaben des Fahrers vorkommen. Daher können Systeme zur Müdigkeitserkennung keine Fahrzeugdaten nutzen, um die Müdigkeit beim automatisierten Fahren einzuschätzen. Um dies zu lösen, werden hier drei neue Ansätze zur Klassifizierung der Müdigkeit des Fahrers im Fahrsimulator vorgeschlagen:

(1) Der erste Ansatz ist die Fusion von Fahrzeugdaten, Elektrokardiogramm (EKG)-Signalen und gesichtsbasierten Merkmalen (Augenlidöffnung und Pupillendurchmesser) unter Verwendung traditioneller maschineller Lernmethoden. Die Eingabedaten werden vorverarbeitet, ihre Merkmale extrahiert und ein Feature Selection Algorithmus angewendet, um redundante und irrelevante Merkmale zu entfernen. Schließlich werden Random Forest und K-nearest Neighbors Klassifikatoren eingesetzt um die Schläfrigkeit in drei Klassen zu unterscheiden: wach, mäßig schläfrig und extrem schläfrig. Die Ergebnisse zeigen, dass die Datenfusion der verschiedenen Datenquellen unter Anwendung der Feature Selection Methode die Systeme, die nur eine einzelne Datenquelle verwenden übertrifft, ebenso das System ohne dem Feature Selection Algorithmus.

(2) Der zweite Ansatz basiert auf dem Training von Deep Convolutional Neural Networks unter Verwendung von EKG-Signalen und Augenliddaten. Ein neuronales Netz wird für jede dieser Eingabedaten aufgebaut, wobei die Soft-Voting-Ensemble-Methode angewendet wird, um die Ergebnisse der beiden verschiedenen neuronalen Netze zusammenzuführen. Die Bayes'sche Optimierungsmethode wird zur Optimierung der Hyperparameter der neuronalen Netze eingesetzt. Die Ergebnisse zeigen, dass Deep Convolutional Neural Networks die traditionellen Methoden des maschinellen Lernens, angewandt auf dieselben Eingabedaten, deutlich übertreffen. Darüber hinaus zeigen die Ergebnisse, dass durch die ausschließliche Verwendung von EKG-Signalen eine mittlere Genauigkeit von etwa 80% bei der Klassifizierung von Schläfrigkeit erreicht wird.

(3) Der dritte Ansatz basiert auf der Verarbeitung von Elektroenzephalogramm (EEG)-Daten zur Erkennung von Müdigkeit. Percentage of Eyelid Closure (PERCLOS) wird als Ground Truth für die Schläfrigkeit des Fahrers verwendet und ein neuronales Encoder-Decoder-Framework vorgeschlagen um die relevanten EEG-Merkmale auszuwählen, die den PERCLOS-Verlauf bei Fahrern beschreiben. Die Ergebnisse zeigen, dass dieses Framework mit einem RMSE von 0.11 eine zufriedenstellende Genauigkeit bei der Klassifikation von Schläfrigkeit erzielt. Dieses Framework liefert auch EEG Merkmale, die bei verschiedenen Fahrern durchgängig kennzeichnend sind um Schläfrigkeit abzuschätzen.

Contents

Abbreviations

AAA	American Automobile Association
ADAM	ADAptive Moment estimation
Adaptive Boosting	Ada-Boost
ADAS	Advanced Driver Assistance System
ADSG	Automated Driving Simulator of Graz
AL	Alert
AMPD	Automatic Multiscale-based Peak Detection
AmpD2Theta	Area under steering angle graph
ANN	Artificial Neural Network
ANS	Autonomic Nervous System
ApEn	Approximate Entropy
AR	Autoregressive (AR) modeling
Avg	Average heart rate
BN	Batch Normalization
BP	Back Propagation
CNN	Convolution Neural Network
Conv	Convolution layer
CWT	Continuous Wavelet Transform
DBN	Beep Belief Network
DCNN	Deep Convolution Neural Network
DVR	German Road Council
ECG	Electrocardiography
ED	Extremely Drowsy
EDA	Electrodermal Activity
EEG	Electroencephalography
EMD	Empirical Mode Decomposition
EMG	Electromyography
EOG	Electrooculography
FastICA	Fast Independent Component Analysis
FC	Fully Connected layer
GMM	Gaussian Mixture Modes
GSR	Galvanic Skin Response
HF	Power of high-frequency band (0.15 Hz-0.40 Hz) in the PSD of the RRI signal

HPD	High Probability Density
HR	Heart Rate
HRV	Heart Rate Variability
ICA	Independent Component Analysis
KNN	K-Nearest Neighbors
KSS	Karolinska Sleepiness Scale
LF	Power of low-frequency band (0.04 Hz-0.15 Hz) in the PSD of the RRI signal
LF/HF	Ratio of LF to HF
Max-Pool	Max-pooling Layer
MD	Moderately Drowsy
MeanRR	Mean of R-to-R Interval (RRI)
MICROSTEERINGS	Small steering adjustment rate
MOL	Multilevel Ordered Logit
MSPC	Multivariate Statistical Process Control
NCA	Neighborhood Component Analysis
NHTSA	National Highway Transportation Safety Administration
NMRHOLD	Number of times steering wheel is hold longer than a predefined duration
NMWRONG	Number of times steering wheel is corrected
ORD	Observer Rating of Drowsiness
PCA	Principal Component Analysis
PERCLOS	PERcentage of eyelid CLOSure
PPG	Photoplethysmogram
PRR50	The percentage of RR50 of all RR intervals
PSD	Power Spectral Density
RBF	Radial Basis Function
RE-A	REsted-Automated
ReLU	Rectified Linear Unit
RE-M	REsted-Manual
RMSE	Root Mean Square Error
RMSSD	Root means square of the difference of adjacent RRI
RR50	The number of pairs of adjacent RRI whose difference is more than 50ms
RRV	Respiratory Rate Variability

SAE	Society of Automotive Engineers
SamEn	Sample Entropy
SDRR	Standard deviation of RRI
sEMG	Surface EMG
SFFS	Sequential Floating Forward Selection
ShEn	Shannon Entropy
Sparse-DBN	Sparse-deep belief networks
SpEn	Spectral Entropy
STFT	Short-Time-Fourier-Transform
STV25	First quartile of steering velocity
STV50	Second quartile of steering velocity
STV75	Third quartile of steering velocity
STWVELZCR	Steering velocity zero crossing rate
STWZCR	Steering zero crossing rate
SVM	Support Vector Machine
TI-A	TIred-Automated
TI-M	TIred-Manual
TP	Total Power (Variance) of RRI
VHAL	Ratio of high to low steering wheel corrections
WFZ	Weight Flat Zero
WPT	Wavelet Packet Transform

Chapter 1

Introduction and Motivation

Drowsiness is an intermediate condition that fluctuates between alertness and sleep. It reduces the consciousness level and hinders a person from responding quickly on important road safety issues [1]. The American Automobile Association (AAA) has reported that about 24% of 2,714 drivers that participated in a survey revealed being extremely drowsy while driving, at least once in the last month [2]. In 2017, the National Highway Transportation Safety Administration (NHTSA) also reported 795 fatalities in traffic accidents which involve drowsy driving [3]. Drowsy driving has caused about 2.5% of fatal accidents from 2011 through 2015 in the USA, and it is estimated to produce an economic loss of USD 230 billion annually [4]. Klauer et al. have found in their study that drowsy drivers contributed to 22-24% of the crashes or near-crash risks [5]. The German Road Safety Council (DVR) has reported that drowsy drivers have caused one out of four fatal highway crashes [6]. In a study carried out in 2015, it has been reported that the average prevalence of falling asleep while driving in the previous two years was about 17% in the 19 European countries [7]. The results of these studies emphasize the importance of detecting drowsiness early enough to initiate preventive measures. Drowsiness detection systems are intended to warn the drivers before an upcoming level of drowsiness gets critical to prevent drowsiness-related accidents.

Intelligent systems that automate motor vehicle driving on the roads are being introduced to the market step-wise. The Society of Automotive Engineers (SAE) issued a standard defining six levels ranging from no driving automation (level 0) to full driving automation (level 5) [8]. While the SAE levels 0-2 require that an attentive driver carries out or at least monitors the dynamic driving task, in the SAE level 3 of automated driving, drivers will be allowed to do a secondary task allowing the system to control the vehicle under limited conditions, e.g., on a motorway. Still, the automation system has to hand back the vehicle guidance to the driver whenever it cannot control the state of the vehicle anymore. However, the handover of vehicle control to a drowsy driver is not safe. Therefore, the system should be informed about the state of the driver.

To date, different Advanced Driver Assistance Systems (ADAS) have been made by car manufactures and researchers to improve driving safety and manage the traffic flow. ADAS systems have been benefited from advanced machine perception methods, improved computing hardware systems, and intelligent vehicle control algorithms. By recently increasing the availability of vast amounts of sensor data to ADAS, data-driven approaches are extensively exploited to enhance their performance. The driver drowsiness detection systems have gained much attention from researchers. Before its use in the development of driving automation, drowsiness warning systems have been produced for the direct benefit of avoiding accidents.

The rest of this chapter first explains the state-of-the-art of driver drowsiness detection methods. Then, it will describe the contribution of this thesis to enhance the previous works.

Table 1.1: Karolinska sleepiness scale (KSS) [9]

Rating	Verbal Description
1	Extremely alert
2	Very alert
3	Alert
4	Fairly alert
5	Neither alert nor sleepy
6	Some signs of sleepiness
7	Sleepy, but no effort to keep alert
8	Sleepy, some effort to keep alert
9	Very sleepy, great effort to keep alert, fighting sleep

1.1 State-of-the-art of Driver Drowsiness Detection Methods

1.1.1 Subjective Measures for Driver Drowsiness

Subjective measures determine the drowsiness using questionnaires that drivers fill to report their levels of vigilance. One of the most commonly used subjective scales for drowsiness is the Karolinska Sleepiness Scale (KSS) [9] that is a nine-point scale that each scale describes the driver's vigilance. Table 1.1 presents the scales of KSS.

The KSS was validated using objective scores derived by brain activities in [10]. Results show that sleepiness can be measured using the relationship between brain signals and the KSS. Results of [11] also show that median reaction time and alpha and theta power densities of brain signals are highly correlated with the KSS.

Datasets that have been exploited to design a driver drowsiness detection system can be categorized into three main groups: 1) Vehicle-based data, 2) Facial-based data, and 3) Bio-signals. The following sections will explain each of these data sources and their corresponding recent studies to develop a driver drowsiness detection system.

1.1.2 Driver Drowsiness Detection using Vehicle-based Data

Researchers have utilized vehicle-based data to design non-invasive driver drowsiness detection systems. The most frequently exploited vehicle-based data for this goal include lateral and longitudinal accelerations, steering wheel angle, steering wheel angular velocity, yaw angle, lateral deviation from road center-line, and speed. In this section, recent studies that have used vehicle-based data to design driver drowsiness detection systems are described.

Steering wheel data has been commonly employed in many studies for this application. For example, Li et al. [12] designed an online driver drowsiness detection system using the extracted approximate entropy [13] of the steering wheel angle signal in real-world driving tests. They employed the rates collected using video observation of the driving tests as ground truth for driver drowsiness. In this method, approximated entropy features were first extracted using fixed-length sliding windows on real-time steering wheel angle data. Then, an adaptive piece-wise linear fitting was employed to linearize the extracted entropy, followed by calculation of the wrapping distance [14] between the linear feature series. Finally, the wrapping distance was used as input to a binary decision classifier to determine the vigilance state (alert or drowsy). Results of that study showed that the accuracy of this method for drowsiness detection is about 78%. Figure 1.1 shows the flowchart of the proposed method in [12].

Table 1.2: Proposed steering wheel angle features to detect the drowsiness in [16].

Feature Name	Description
ELLIPSE	Steering angle and absolute steering velocity
NMWRONG	Number of times steering wheel is corrected
NMRHOLD	Number of times steering wheel is hold longer than a predefined duration
AmpD2Theta	Area under steering angle graph
VHAL	Ratio of high to low steering wheel corrections
MICROSTEERINGS	Small steering adjustment rate
STWZCR	Steering zero crossing rate
STWVELZCR	Steering velocity zero crossing rate
STV25	First quartile of steering velocity
STV50	Second quartile of steering velocity
STV75	Third quartile of steering velocity

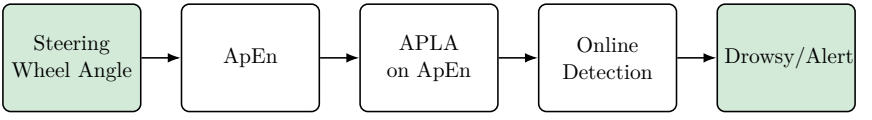


Figure 1.1: Proposed method in [12] for driver drowsiness detection. ApEn: approximate entropy; APLA: adaptive piece wise linear approximation.

Meng et al. [15] extracted 11 features from the steering wheel angle signal and applied multivariate analysis of variance to find parameters that have a significant correlation with the level of drowsiness. By applying this method, the number of parameters was reduced to the four most important features. The KSS was used to define the ground truth for three levels of drowsiness: (1) 1≤KSS≤6, (2) KSS = 7, and (3) KSS = 8-9. Three different classifiers, including Back Propagation (BP) neural network, Support Vector Machine (SVM), and Multilevel Ordered Logit (MOL), were built based on these selected parameters. Results showed that the MOL achieved a higher recognition accuracy than two other models to classify three levels of driver drowsiness. The balanced accuracies of the MOL, SVM, and BP were 72.92%, 63.86%, and 62.10%, respectively.

Friedrichs and Yang [16] proposed time-domain features of steering wheel angle for drowsiness detection obtained from real-world driving tests. These features are listed in Table 1.2. In their study, the effects of the external environment, such as road bumps and cross-wind, were also considered. They used KSS reports as ground truth for driver drowsiness and defined three classes based on this scale: (1) alert (KSS<6), (2) questionable (KSS = 6-7), and (3) drowsy (KSS = 8-9). Correlation of features with the subjective Karolinska Sleepiness Scale (KSS) was analyzed, and their performance was investigated using K-Nearest Neighbors (KNN; k=5) [17], Gaussian Mixture Modes (GMM; 3 Modes) [18], linear discriminant [19], Bayes classifier [20] combined with Sequential Floating Forward Selection (SFFS) [21] and Artificial Neural Network (ANN). The results showed that the ANN outperformed other classifiers. The balanced accuracy (average accuracy of three classes) was 73.5%; however, the accuracy of the drowsy class was only 54.6%. Therefore, extreme drowsiness levels can be misclassified in the questionable or alert classes.

Wang and Xu implemented MOL modeling on vehicle-based, and eye movement measures

[22]. In their study, 19 driving behavior variables (such as the average of lateral position, lane departure frequency, steering wheel reversals, etc.) and four eye feature variables, including average blink frequency per second, average blink duration per second, PERcentage of eyelid CLOSure (PERCLOS) and average pupil diameter were measured. To record the driver drowsiness levels, drivers were asked to report their KSS, and they are used to define three classes of drowsiness: (1) KSS range from 1 to 6, no drowsiness or low-level drowsiness, (2) KSS is 7, moderate-level drowsiness, (3) KSS is 8 or 9, high-level drowsiness. Their results showed that considering the individual differences between drivers in the structure of MOL improves the models' ability for drowsiness detection.

McDonald et al. [23] proposed a contextual and temporal algorithm for drowsiness detection that exploited steering wheel angle, gas pedal input, brake pedal input, vehicle speed, and acceleration as input data. Speed and acceleration were used to build a real-time measure of driving context. In their study, dynamic Bayesian networks [24] were employed for considering the time dependencies in the transition between drowsiness and awake states. The drowsiness-related lane departures, which were identified using video observation of all lane departures, were employed to define the ground truth of drowsiness. The modeling approach considered in that study consists of two types of features: driver behavioral measures (steering wheel angle, gas pedal input, and brake pedal input) and road context measures (speed and acceleration) while random forest algorithm [25] and Symbolic Aggregate Approximation (SAX) [26] were employed to generate features form each type of this data, respectively. Results showed that the number of false-positive of drowsiness detection was decreased in highway and rural environments. According to the results, the false positive rate and true positive rate of their method were 0.89 and 1.00, respectively.

An ensemble learning for detection of drowsiness-related lane departures was presented in [27]. In that study, they identified the drowsiness-related lane departures and used the steering data from 60 s to 6 s before a lane departure as input data. Observer Rating Drowsiness (ORD) scale [28] was used as the ground truth for driver drowsiness that is a continuous scale between 0 and 100, and it was separated into five levels: not drowsy (ORD$<$12), slightly drowsy (12$\leq$ORD$<$37), moderately drowsy (37$\leq$ORD$<$62), very drowsy (62$\leq$ORD$<$90), and extremely drowsy (ORD$\geq$90). The ORD scale was matched with the lane departures, and the segments which observed as moderately, very, or extremely drowsy and contained lane departures were labeled as drowsy. Other lane departures were considered as members of the alert class. Thus, they labeled the drowsiness in two classes: drowsy and alert. Finally, the random forest algorithm was utilized as a classifier by inputting the segmented drowsiness-related raw steering data. Results showed that the overall accuracy of this approach is about 79% for binary classification of drowsiness.

Three feature sets in time, frequency, and state-space domains were extracted from steering angle data in [29] to capture impaired driving caused by drowsiness. Each feature set was separately used as input to 5 different classifiers, and the output of every single classifier was combined in an ensemble classifier. Finally, a meta-ensemble classifier was employed to combine the outputs of three ensemble classifiers by averaging their outputs. The results showed that this approach obtained an accuracy of 86.1%.

Standard Deviation of Lane Position (SDLP) is another vehicle-based feature used in previous research for driver drowsiness classification. For example, Ingre et al. [30] studied the relationship between SDLP and KSS ratings. They designed two 2-hours driving tests for ten shift workers, five male, and five female. Drivers performed one of the tests after a normal night sleep and another one after working a night shift. The subjective drowsiness was monitored using KSS in 5-min intervals during the driving tests. A mixed ANalysis Of VAriance (ANOVA) was employed to study the relationship between the average of SDLP and KSS levels. Results show that the average SDLP (average of 20 driving tests) increased in higher KSS levels. For example, the KSS levels of 1, 5, and 9 are associated to mean SDLP of 0.19, 0.26, 0.47, re-

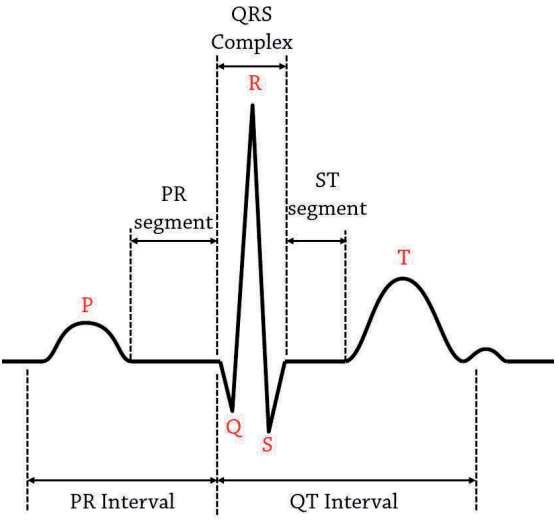


Figure 1.2: Schematic representation of ECG wave, adopted from [31].

spectively. However, SDLP is not higher than 0.25 for some drivers, even at the KSS rating of 9.

1.1.3 Driver Drowsiness Detection using Bio-signals

Biosignals that have been commonly employed for driver drowsiness detection include Electroencephalography (EEG), Electrocardiography (ECG), Photoplethysmogram (PPG), Electromyography (EMG), Electrooculography (EOG), Electrodermal (EDA), and respiration. In this subsection, the structure of each one of these signals and some of the corresponding studies for driver drowsiness detection are explained.

- **ECG**

Electrocardiography (ECG) is a method to measure the electric activity of the heart over a period of time which is recorded by electrodes connected to body [31]. Figure 1.2 shows the scheme of a normal ECG recording.

Detection of R Peaks is the first step to extract the Heart Rate (HR) and Heart Rate Variability (HRV), which are employed in medical applications to detect heart diseases [32, 33]. The HRV can represent the Autonomic Nervous System (ANS) activity related to drowsiness and stress levels. Activities of this system are classified into two general parts: sympathetic and parasympathetic activities. Alertness states are characterized by an increase in sympathetic activity and/or a decrease in parasympathetic activity, while extreme relaxation states are characterized by an increase in parasympathetic activity and/or a decrease in sympathetic activity [34, 35]. Therefore, HRV signals have been used to detect drowsiness in drivers in different studies [36, 37, 38]. The most widely extracted features from HRV signals are as follows [36, 37]:

- Avg: Average heart rate.
- MeanRR: Mean of R-to-R Interval (RRI).

- RMSSD: Root mean squares of the difference of adjacent RRI.
- RR50: The number of pairs of adjacent RRI whose difference is more than 50ms.
- PRR50: The percentage of RR50 of all RR intervals.
- SDRR: Standard deviation of RRI.
- Total Power (TP): Variance of RRI.
- LF: Power of low-frequency band (0.04 Hz-0.15 Hz) in the Power Spectral Density (PSD) calculated in each sliding window of the RRI signal. This feature reflects both the parasympathetic and sympathetic nervous system activities.
- HF: Power of high-frequency band (0.15 Hz-0.40 Hz) in the PSD of the RRI signal. This feature reflects parasympathetic nervous system activity only.
- LF/HF: Ratio of LF to HF. This feature presents the balance between parasympathetic and sympathetic nervous systems' activities.

To discriminate between the HRV dynamics in two states of fatigued (caused by sleep deprivation) and drowsy (caused by monotonous driving), two different monitoring systems were proposed in [39] based on extracted features from HRV and respiration signals. One of these systems is a binary classifier (alert/drowsy) for assessing the level of driver vigilance every minute. Another one detects the driver's sleep deprivation in the first three minutes of his driving. Thirty drivers participated in their driving tests, and every driver was classified as sleep-deprived if either s/he slept less than 4 hours in the night before the test (partially sleep-deprived) or s/he was awake for at least 20 hours before conducting the test (fully sleep-deprived). Thirteen drivers were sleep-deprived (either partially or fully) out of thirty drivers. A combination of EEG signal analysis, video observation-based rating, and PERcentage of eyelid CLOsure (PERCLOS) was employed to define the ground truth for driver drowsiness [40]. Results of this study showed that the balanced accuracy of the drowsiness detection system that only used the HRV-based features is about 65.5%. However, by adding the extracted features from respiration signals, this system achieves a balanced accuracy of 78.5% (about 13% improvement). The balanced accuracy of the sleep deprivation system was also about 75% (it detected eight out of thirteen sleep-deprived drivers correctly).

The behavior of the extracted features from HRV dynamics in different levels of driver drowsiness was studied in [41] where ECG signals were collected using embedded ECG sensors on the steering wheel. Two male drivers (aged 27 and 31) conducted a driving test with a duration of 2 hours. The driver was vigilant at the beginning of the test while they showed clear signs of drowsiness, such as yawning at the end of the test. Results showed that heart rate, SDNN, and RMSSD decreased by increasing the drowsiness level, whereas the LF/HF ratio increased by the transition from alertness to the extremely drowsy level.

Li and Chung [42] employed the combination of HRV wavelet analysis and SVM classifier to detect the driver drowsiness. The PERCLOS was used to define the two driving states. The PERCLOS between 0% to 30% and upper that 30% indicated the alert and drowsy states of the drivers, respectively. Receiver Operating Curve (ORC) [43] analysis was also used to select the extracted features from HRV data, and its results represented that the wavelet-based method outperformed the FFT-based method (LF/HF ratio) regardless of the duration of the HRV data (1-min, 2-min, and 3-min). Classification results showed that the wavelet-based feature system obtained an overall accuracy of 95%. Figure 1.3 presents the flowchart of the proposed method in [42].

Three features, including LF, HF, and LF/HF ratio, were used in [1] to detect the drowsiness states during a monotonous 80-min length driving experiment performed by 22 subjects. The

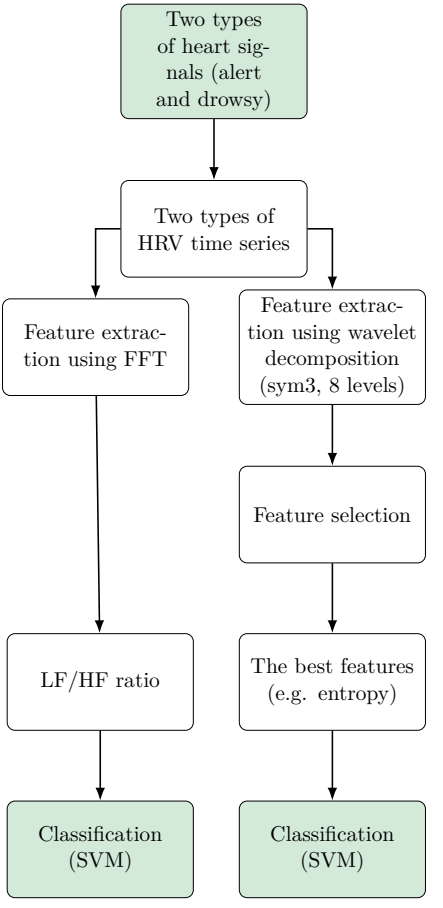


Figure 1.3: Flowchart of the proposed method in [42].

SVM classifier was employed to classify the driver states, and paired t-test was used to select the statistically significant features ($p < 0.05$). Video recording during the driving test was used as a measure of ground truth, and drowsiness-related events were detected using observation of facial features such as eye blink duration, facial tone, head nodding, and yawning. Video observations showed that only 11 out of 22 drivers felt drowsy during the test. When the video observer detected a drowsy event, the 5-min period before the event and 5-min period after the event was annotated as alert and drowsy states, respectively. Results showed that the SVM obtained an overall accuracy of 70% for detecting the drowsy conditions using only HRV-based features. However, results demonstrated that using both of the ECG and EEG features increased the classification accuracy to about 81%.

Patel et al. [35] suggested an ANN-based system to detect the early onset of fatigue in drivers using HRV. This study investigated the relationship between LF and HF components of HRV and driver drowsiness. To classify the HRV signals, a single-layer neural network using a combination of bipolar logistic function as an activation function and delta learning rule has been designed. The spectral image, plotted from the PSD, was the input given to the neural network that yielded an accuracy of 90%. Figure 1.4 shows the LF/HF ratio of alert and fatigue states. It shows that this ratio is higher for 11 out of 12 drivers in the alert state than in the drowsy state.

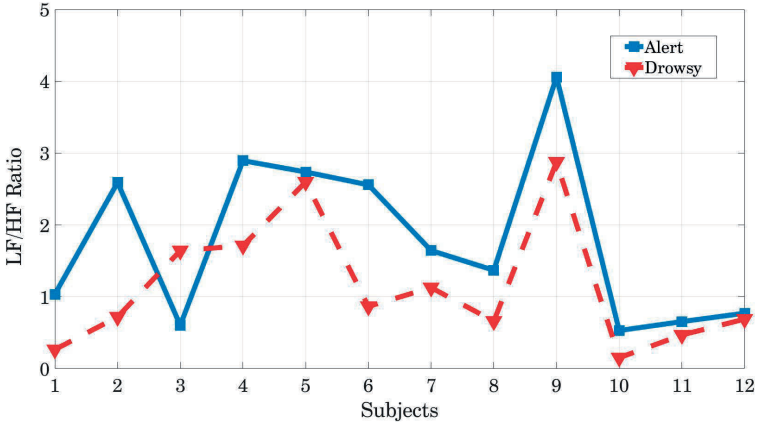


Figure 1.4: LF/HF ratio of 12 participants in two states: Alert and Drowsy. It shows that this ratio is higher for 11 out of 12 drivers in alert state, adopted from [35].

Multivariate statistical process control (MSPC) [44] was employed in [38] to detect driver drowsiness using eight different HRV-based features. EEG-based sleep scoring was conducted by a sleep specialist and utilized as ground truth for drowsiness detection. In that study, drowsy driving was considered as an anomaly in driving and MSPC was applied to detect this anomaly. The performance of the method was evaluated using EEG signals and the results showed that the proposed method detected 12 out of 13 drowsiness onset and false-positive rate of the anomaly detection system was about 1.7 times per hour.

Buendia et al. [45] had two objectives in their study. The first goal was the evaluation of various signal processing methods for removing noisy heartbeats and examine the effects of spectral transformation of HRV signals on drowsiness detection. The second goal was extracting the relationship between HRV data and KSS data that were reported by drivers every 5 minutes during the real-world driving tests. Results showed that the average heart rate decreased with

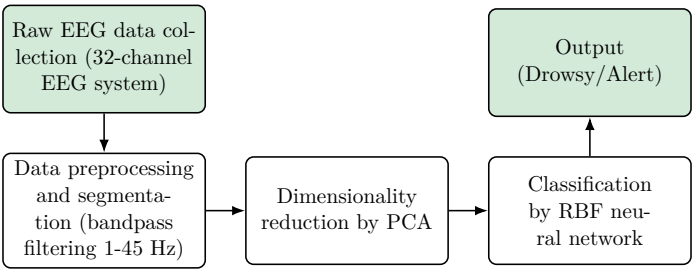


Figure 1.5: Proposed method in [54] to classify the driver drowsiness using EEG data and RBF neural network.

increasing the KSS, whereas variance of the heart rate increased.

- **EEG**

Brain signals are generally promising biosignals for drowsiness classification. Besides the effort to measure it, EEG signals are contaminated by various noise sources such as muscle activities and eye movements. EEG signals are usually decomposed into their sub-bands for being used in drowsiness classification. Important sub-bands of EEG include Delta (0-4Hz), Theta (4-8 Hz), Alpha (8-12Hz), Beta (12-30 Hz) and Gamma (30-60 Hz) [46]. Recent studies showed that the dynamic behaviors of these sub-bands in time and frequency domains have strong relationships with different levels of drowsiness in drivers [47, 48, 49]. For example, Taran and Bajaj [50] employed adaptive Hermite decomposition [51, 52] for feature extraction from EEG signals. They utilized the MIT/BIH Polysomnographic data set [53] which is developed to study the sleep apnea disorder, and no driving test was conducted while collecting this data set. In their proposed method, evolutionary optimization algorithms including Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Artificial Bee Colony (ABC) were first used to optimize the Hermite decomposition parameters (order of Hermite polynomial and dilation factor) where ABC provided the minimum reconstruction error of the EEG signals. Second, the Hermite coefficients were calculated, and their statistical measures were extracted as output features. Finally, the extracted features were tested with different classifiers such as extreme learning machine, decision tree, k-nearest neighbor, least-square support vector machine, artificial neural network, and naïve Bayes algorithms to detect alertness drowsiness using EEG signals. The extreme learning machine classifier achieved better performance than other classifiers that obtained 95.45% and 87.92% detection accuracies for alertness and drowsiness states, respectively.

Two-Level Learning Hierarchy Radial Basis Function (RBF-TLLH) was exploited in [54] to classify the driver vigilance state into two classes (drowsy and alert) using EEG signals. They used EEG data were collected from six participants that conducted the driving tests in two vigilance states: alert and drowsy. In the alert tests, the drivers were asked to have a usual sleep time (about 8 hours) during the night before the tests. On the other hand, to conduct the drowsy tests, drivers were required to sleep for only 4 hours during the night before the test. By considering this approach, no ground truth based on the driving performance was needed. In this study, first, EEG signals were preprocessed and band-pass filtered (1-45 Hz). Then, the PCA was employed to reduce the dimensions of the preprocessed EEG signals and extract features from them. Finally, RBF-TLLH was utilized to classify the driver's vigilance state. The results demonstrated that the proposed method achieved the mean accuracy of 92.71% for drowsiness classification. The flowchart of this method is shown in Figure 1.5.

EEG power spectrum analysis was used in [55] to study the brain activities during a

monotonous driving test. Nine university students participated in the experiments. Observation of recorded videos during the driving experiment was exploited as the ground truth. The PSD of different EEG sub-bands were computed, and results showed that Alpha and Theta band powers increased significantly during the transition from alert to the drowsy state. This result was more significant in the occipital and parietal regions [56] than other regions of the brain.

Time, spectral, and wavelet analysis of EEG signals were used in [4] to detect the drowsiness state. In that study, the MIT-BIH Polysomnographic Database [53] was used that all subjects were male and aged between 32 and 56 years old. In the proposed method, first, signals were band-pass filtered with cut-off frequencies of 0.5-60 Hz, and power-line interference (50 Hz) was removed by using the cascade of adaptive filters [57]. Second, signals were segmented into 5 seconds windows. Third, 19 different features were extracted from EEG signals, such as standard deviation in time analysis, central frequency in spectral analysis, and zero-crossing rate of wavelet decomposed levels. Fourth, Linear Decomposition Analysis (LDA) was exploited to select the most informative features to the drowsiness level. Finally, an ANN with one hidden layer was employed as a binary classifier which was trained using the Levenberg-Marquardt backpropagation algorithm. The number of neurons in the hidden layer was varied between 10 to 40 to select the optimal architecture of the ANN classifier. The flowchart of the proposed method in [4] is shown in Figure 1.6. Results demonstrated that this method achieved the correct detection rates of 87.4% and 83.6% for alertness and drowsiness states, respectively.

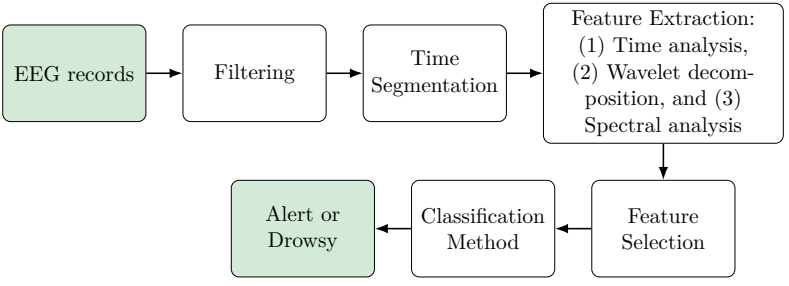


Figure 1.6: Flowchart of the proposed method in [4] for drowsiness detection using EEG signals.

Kurt et al. [58] used EEG, EMG, and EOG signals to estimate three different vigilance levels: alert, drowsy, and asleep. EEG signals were decomposed into their sub-bands using wavelet transformation in their study, while EMG and EOG signals were employed to eliminate movement artifacts from EEG sub-bands. After preprocessing the EEG signals, left and right EOG signals, EEG sub-bands (Delta, Theta, Alpha, and Beta), and EMG signal were used as inputs (seven different input signals) to an ANN classifier with two hidden layers that each contained ten neurons. Ten subjects participated in their driving tests, and two experts rated the collected data and classified the driver's vigilance level into three classes: awake, drowsy, and sleep. The overall accuracy of the method was 97-98% for driver drowsiness classification.

A real-time wireless brain-computer interface system was designed in [59] for drowsiness detection. In that study, an embedded drowsiness detection was also implemented based on Alpha and Theta sub-bands spectra for binary classification of drowsiness. They assumed that the driver is alert in the first few minutes of his driving tests and used the first 3-min of EEG data to build the alertness model. This model was built under the assumption that EEG spectra in Theta and Alpha bands follow a multivariate normal distribution. The deviation from the alert model was assessed continuously by calculating the Mahalanobis distance [60] between the current distribution of Alpha and Theta bands and alertness model. Finally, the driver state was classified as drowsy if the calculated distance was higher than a predefined threshold.

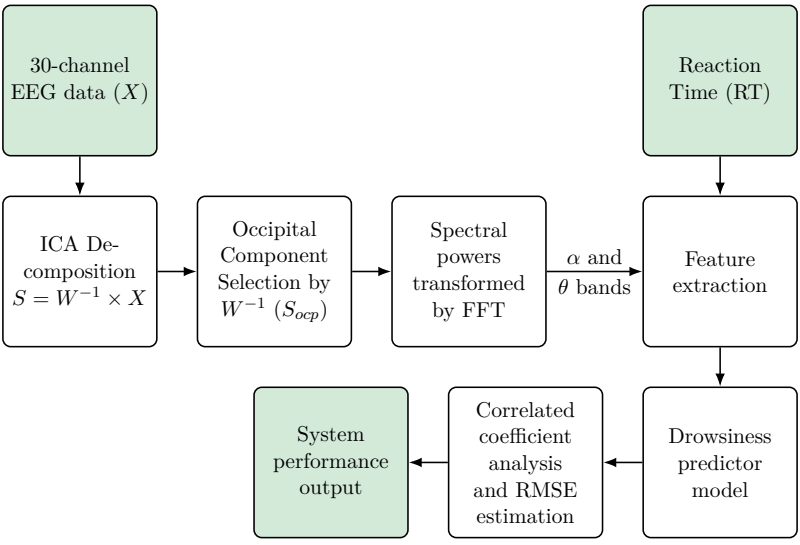


Figure 1.7: Flowchart of the proposed method in [61] for drowsiness detection using EEG data. Alpha and Theta powers of the EEG signals are obtained by applying Fast Fourier Transform (FFT) and extracted features from these sub-bands are exploited to predict the driver drowsiness.

Driver's reaction-time to the simulated lane deviation onset was employed to define the ground truth of the driver's drowsiness. If the reaction time was within 0.2-1 sec, the driver would be classified as an alert. Results showed that the F1-score of this method was 76.7%.

A generalized EEG-based drowsiness monitoring and prediction system using a self-organizing neural fuzzy network was proposed in [61]. In that study, simulated lane departures were implemented in the 1-h driving experiments, and drivers were required to correct the lateral position of the car by inputting the appropriate steering angle. Six drivers participated in the study, and their reaction times to these events were recorded to be used as ground truth for driver drowsiness, and a drowsiness predictor model was developed to estimate this reaction time (regression problem). Figure 1.7 illustrates the flowchart of the proposed method in [61]. As this Figure shows, the method composed of these main parts: (1) Independent Component Analysis (ICA) [62], (2) power spectral analysis, (3) feature extraction, (4) drowsiness predictor model, and (5) estimation of drowsiness detection model performance. First, the ICA method was used to select the occipital components of the EEG signals (S_{ocp}) as the region of interest for spectral analysis and feature extraction. Second, Fast-Fourier Transform (FFT) was applied to derive the power spectral density of Theta and Alpha bands over frequency. The driver's reaction time to the simulated events was also employed as the actual output to evaluate the model performance. Third, four different models including (1) Support Vector Regression (SVR) [63], (2) Multi-Layer Perceptron Neural Network (MLPNN) [64], (3) Radial Basis Function Neural Network (RBFNN) [65], (4) Self-organizing Neural Fuzzy Inference Network (SONFIN) [66] were utilized to estimate the reaction time of the driver. Finally, the performance of these models was evaluated using correlation analysis and root mean square error (RMSE) between the estimated reaction time and the actual one. Results showed that the SONFIN models obtained the lowest average RMSE (0.076 ± 0.022) and highest average correlation coefficients ($97.2\% \pm 1.6$) among other proposed models.

A driving drowsiness detection system was designed based on EEG signals using the com-

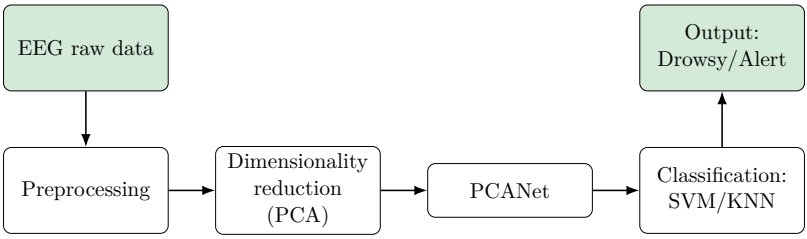


Figure 1.8: Structure of the proposed method in [67] for driver drowsiness detection by applying a combination of principal component analysis (PCA) and deep networks to the EEG data.

bination of principal component analysis (PCA) and a deep learning model called PCANet in [67]; see Figure 1.8. The dimension of EEG data was reduced using PCA to improve the feasibility of the final structure of the model. Two classification algorithms, including KNN and SVM, were exploited to generate the final binary output. Six drivers participated in the simulated driving tests. Two driving tests were organized for every driver: alert and drowsy. An observer seated two meters beside the drivers during the experiments and recorded their vigilance class by observing the drowsiness signs such as head nodding and long eye closure. Results demonstrated that the developed method obtained robust performance that provided a classification accuracy of about 95%. Moreover, results showed that parietal and occipital lobes were strongly associated with driver drowsiness.

Wang et al. [68] designed a real-time fatigue detection based on collected 24 EEG signals using dry electrodes. Ten subjects (seven males and three females) participated in the simulated driving tests with the length of 90 min, carried out in the afternoon from 3 to 5 pm. The drives' reaction time to the brake signals of the leading car was recorded during the experiments. Two methods were investigated for drowsiness detection: PSD analysis and sample entropy (SE) analysis. Figure 1.9 presents the flowchart of these methods. As this Figure shows, signal preprocessing and discrete wavelet transform were first applied to the EEG channels. EEG signals were bandpass filtered to the frequency band of 0.5 to 100 Hz in the preprocessing step. The Daubechies-4 (db4) wavelet was used to decompose the EEG signals into six different levels. In the PSD analysis method, the computed wavelet coefficients were used to provide the Theta, Alpha, and Beta sub-bands frequency components of EEG signals and subsequently their spectral powers: p_θ, p_α and p_β. Two drowsiness indexes were calculated using these powers: (1) $(p_\theta + p_\alpha)/p_\beta$ and (2) p_θ/p_β. The correlation between calculated drowsiness indexes and reaction times was computed, and if this correlation was higher than a predefined threshold, that specific channel was selected to be used for drowsiness detection. This process was performed for each drowsiness index separately. This step provided the sensitive EEG channels for the detection of the onset of driver drowsiness. Finally, an integrated drowsiness metrics (IM) was calculated as the average of both drowsiness indexes presented by Equation (1.1).

$$IM = \frac{1}{2}\left(\sum_{n=1}^{N}\frac{p_\theta(n)+p_\alpha(n)}{p_\beta(n)} + \sum_{m=1}^{M}\frac{p_\theta(m)}{p_\beta(m)}\right) \tag{1.1}$$

where N and M are the number of sensitive EEG channels for the $(p_\theta + p_\alpha)/p_\beta$ and p_θ/p_β, respectively. In the SE method, the sample entropy [13] was calculated by averaging the sample entropies of the two EEG channels from the occipital region of the brain: O1h and O2h channels. The calculated sample entropy was also employed to predict the onset of drowsiness. Results showed that the IM was increasing during the driving test by increasing the reaction time. On the other hand, the SE was decreasing by increasing the drivers' reaction time to the brake signals.

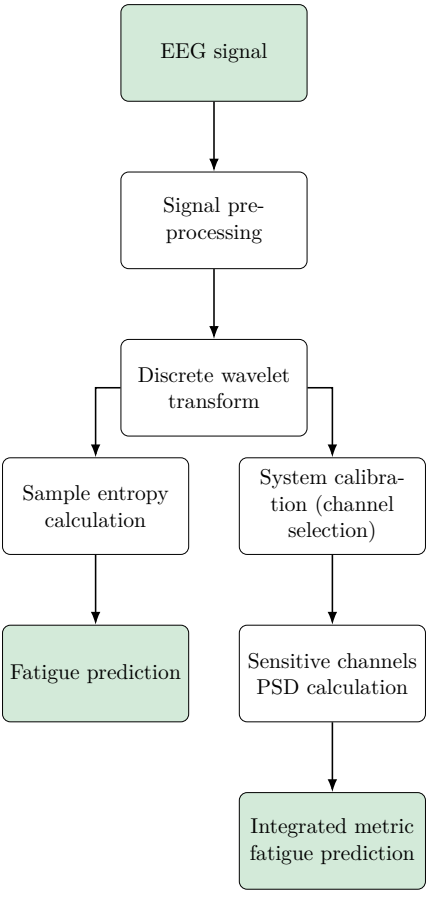


Figure 1.9: Flowchart of the proposed method in [68].

- **EMG**

EMG is a method to detect the electrical activity of the muscle using connecting the electrodes to the skin surface. Surface EMG (sEMG) was analyzed in [69] to determine the dynamical behavior of muscle activity during simulated driving. Electrodes were placed on the shoulder and neck muscles of the driver.

The vigilance level was estimated in [70] using both EEG and EMG signals to increase the accuracy of the estimation. Thirty subjects (14 females and 16 males) with a mean age of 33.5 years participated in the tests. No driving test was performed in that study, and participants laid down on a testbed to collect the data. Every test took about 7 hours for every participant, and collected data were segmented into 20 min parts. Two experts evaluated the recorded EEG data, and they reported the three vigilance levels: awake, drowsy, and sleep. EMG signals were collected using electrodes that were connected to the chin. These signals were used to verify and remove motion artifacts from EEG signals. An ANN was utilized to investigate the changes of EEG and EMG from alert to drowsy states. Training and testing data sets of ANN consist of four EEG sub-bands (Delta, Theta, Alpha, and Beta) and the power spectral density of EMG signals. The output of the ANN was three levels of vigilance: awake, drowsy, and asleep. Results show that the accuracy of the method is 98-99%. Figure 1.10 shows the structure of the used ANN.

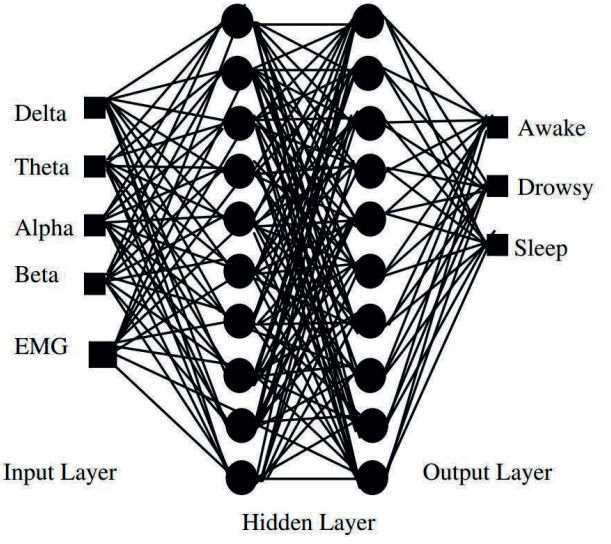


Figure 1.10: The method used in [70] to classify driver drowsiness using power spectral densities of Delta, Theta, Alpha and Beta sub-bands of EEG signals, and power spectral density of EMG data as inputs to a neural network.

Fu and Wang [71] proposed a method for driver fatigue detection based on sEMG and ECG data. Eight subjects (five males and three females; age: 24.75±2.76) participated in the driving tests. The length of every test was between 2 to 2.5 hours and conducted between 12pm to 3pm. Subjects raised their hand when they feel drowsy during the tests, which is used to define the ground truth for driver drowsiness. Collected data were first preprocessed by applying the Fast independent component analysis (FastICA) [72]. This algorithm was used to separate the sEMG and ECG signals by removing the noisy components and reconstructing the clean signals

by remained components. However, the baseline low-frequency noise was still present in the outputs of the FastICA. Therefore, band-pass filtering was applied to the signals (sEMG: 10 to 300 Hz and ECG: 0.5 to 50 Hz) to remove the baseline noise. Some features were extracted from these signals, and results showed that some of them were not informative to the driver's drowsiness. Thus, Kolmogorov-Smirnov z-test [73] was employed as the feature selector and peak factor of sEMG, and the maximum of cross-correlation curve of sEMG and ECG showed the most significant difference between two alert and drowsy classes, so these two features were selected for drowsiness detection. Finally, the Mahalanobis distance (MD) was employed as the discriminator between drowsy and alert states. The MD provided the decision boundary between two classes when two selected features were utilized as its inputs. Results showed that this method achieved an accuracy of about 87% to discriminate between drowsy and alert states.

In [74] a similar method based on a combination of EMG and ECG features was proposed for drowsiness detection. In that study, twelve male participants (20-30 years old) performed the simulated driving tests with the length of 120 min. Drivers filled the subjective questionnaire of SOFI-25 (Swedish Occupational Fatigue Inventory) [75] during the tests at 10 min intervals that were used to define the ground truth for drowsiness. Firstly, EMG and ECG were separated using FastICA and Empirical Mode Decomposition (EMD) [76] was exploited to denoise the signals. Then, three features, including the complexity of EMG, the complexity of ECG, and the sample entropy of the ECG signal, were extracted. Principal component analysis was used to eliminate the redundant information between extracted features. Finally, a mathematical model using multiple linear regression theory was established while the principal components of three features were used as input data, and output was the drowsiness level of the driver. Results showed that this approach achieves an accuracy of 91% for binary classification of driver drowsiness.

- **Respiration**

A drowsiness detection method using analysis of the respiratory signal was proposed in [77]. The respiratory signal was collected using an inductive plethysmography belt [78] from twenty adult participants (ten males and ten females, aged 20 to 60 years). The subjects performed the two driving tests on two different days. In one of the tests, drivers had a normal sleep situation (at least 6 hours of sleep in the night before the test). In another test, drivers had no sleep in the 24 hours before starting the test. To generate the ground truth, trained external observers were asked to rate the driver's vigilance levels as drowsy or non-drowsy per minute by observing the recorded videos. The drowsiness detection method was based on Respiratory Rate Variability (RRV) analysis to detect the fights against drowsiness and named as Thoracic Effort Derived Drowsiness index (TEDD). This algorithm is presented in Figure 1.11. A low-pass filter filtered the collected respiration signals with a cut-off frequency of 0.5 Hz. This low-pass filter was followed by a high-pass filter with a cut-off frequency of 0.05 Hz. The low-pass and high-pass filters removed the noise spikes and baseline of the respiration signals, respectively. The quality of the filtered signal was assessed, and it was excluded from further processing if its quality was not approved. Characterization block found the reference dynamics of every individual driver to deal with differences between different subjects. The block searched to find a stable region used to extract a reference pattern for every driver. In the Breath-to-Breath respiratory rate estimation block, a quasi-peak detector was applied to estimate the changes in the RRV signal. Finally, a drowsiness detector was applied to the RRV signal, and its performance was assessed using different classification metrics. Results showed that this method obtained the specificity of 96.6% and the sensitivity of 90.3%.

Kiashari et al. [79] presented a non-intrusive method using facial thermal imaging to analyze drivers' respiration signals. Thirty subjects participated in the driving tests, and they were assumed to be awake at the beginning of the test and show drowsiness signs during the experiment. The mean and standard deviation of the respiration rate and the inspiration-to-

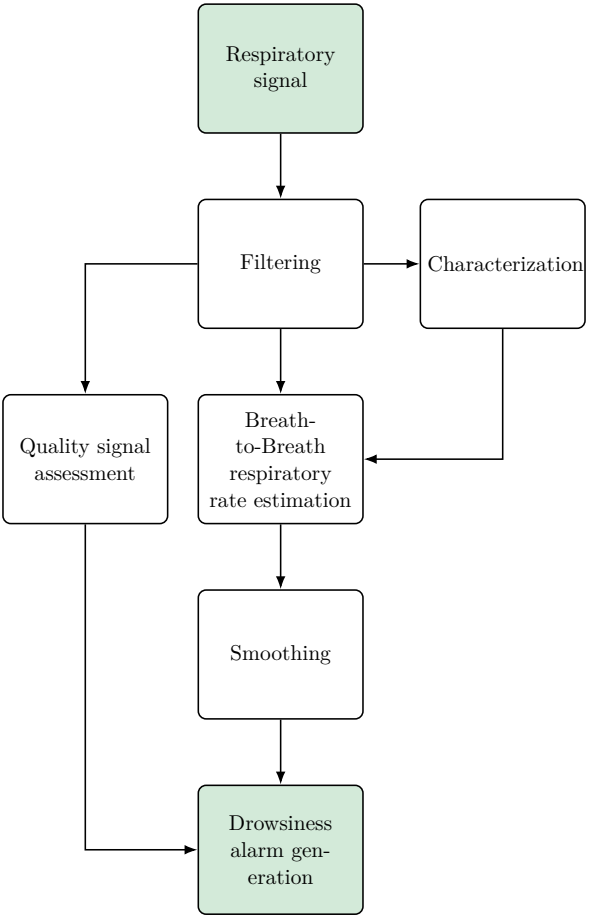


Figure 1.11: Flowchart of the proposed method in [77] for driver drowsiness detection by estimation of breath-to-breath respiratory rate.

expiration time ratio were extracted from signals. The ORD was employed as the ground truth for drowsiness detection obtained by evaluations performed by three expert observers. The observers reported drowsiness into five levels: not-drowsy, slightly drowsy, moderately drowsy, very drowsy, and extremely drowsy. The final drowsiness score was provided by averaging the three levels reported by three observers. The ORD score of three and higher was labeled as a drowsy class, and the ORD lower than three was considered as the alert class. The SVM was also employed for drowsiness classification. Results showed that this method obtained the overall accuracy and precision of 90% and 91%, respectively.

Solaz et al. [80] proposed a driver drowsiness detection system using Kinect cameras that were installed in the car to be used as a breathing rate sensor. Output images of this sensor were processed to investigate the movement of the driver's chest. Five male subjects participated in the driving tests aged between 22 and 38 years old. The tests were conducted in two different states: (1) one test with normal sleep, and (2) another one with sleep-deprived drivers (at least 24 hours of being awake). To validate the cameras, the plethysmography band was used as a gold standard that obtained precise chest and abdomen movements. Results showed that the breath rate measured with the cameras was correlated with the plethysmographic measure greater than 90%.

Dosario et al. [81] used different biosignals such as ECG and respiration to discriminate between three levels of driver's vigilance states (1) attentive, (2) fatigued, and (3) drowsy. Twenty subjects participated in driving tests with the length of 1 h 45 min. The combination of EEG dynamics and PERCLOS data were exploited to define a ground truth for driver's vigilance states during the experiment. The respiration signal was measured using a plethysmography band. Results showed that the respiration amplitude was 5% higher during the fatigue and drowsy states than during the attentive state.

A respiration monitoring system based on Radio-frequency IDentification (RFID) technology was proposed in [82]. This system was composed of multiple RFID tags attached to the seat belt, which were used to estimate the respiration frequency of the driver. In order to mitigate the influence of random sampling and external noises such as vehicle vibration, tensor completion [83] and Canonical Polyadic Decomposition (CPD) [84] were applied to the RFID phases. The results showed that the proposed systems outperformed the Kinect cameras and Ultra-Wide-Band (UWB) radar. The mean estimation error of Kinect camera, UWB radar, and RFID tags were 0.79, 0.31, and 0.11 breath per minute, respectively.

1.1.4 Driver Drowsiness Detection using Facial-based Data

Different facial-based drowsiness signs were exploited in recent studies to detect driver drowsiness, such as rapid and constant blinking, head nodding, and yawning. For example, A CCD camera was used in [85] to recognize and track the driver's mouth movements for yawning detection during the driving test. In that study, Gravity-Center template [86], and greyscale projection were used to detect the driver's face and his mouth corners, respectively. The Gabor wavelets [87] were then exploited to extract features from detected left and right corners of the mouth. Finally, a Linear Discriminate Analysis (LDA) was employed as a classifier to detect yawning using extracted features. This method is used to detect the yawning in 400 images, and results showed that this method achieved an accuracy of about 95% for yawning recognition during driving tests.

A facial-based feature called Eye Aspect Ratio (EAR) was used in [88] for developing a real-time drowsiness classification system. The first step of this method was eye detection in the driver's image. The landmark estimation method described in [89] was applied for this purpose. After eye detection and for every video frame, the EAR was estimated as the average of eye's height divided by eye's width. Figure 1.12 shows the used eye landmarks to calculate the EAR. Equation (1.2) also presents how to calculate the EAR. Thus, the EAR is equal to a small value when eyes are closed and increasing back when the eyes are open. This measure

can also be used to calculate the blinking rate. In order to classify the driver drowsiness, fifteen consecutive EARs were used as inputs to three different classifiers, including Multi-Layer Perceptron (MLP), random forest, and SVM. Every classifier classified the eye blinking dynamics into three categories: (1) open eye, (2) short blink, and (3) long blink. To validate the driver drowsiness model, a public database called "DROZY" [90] was exploited. Results showed that SVM outperformed other classifiers and achieved the classification accuracy of 94.9% for drowsiness detection.

$$EAR = \frac{||p_2 - p_6|| + ||p_3 - p_5||}{2||p_1 - p_4||} \tag{1.2}$$

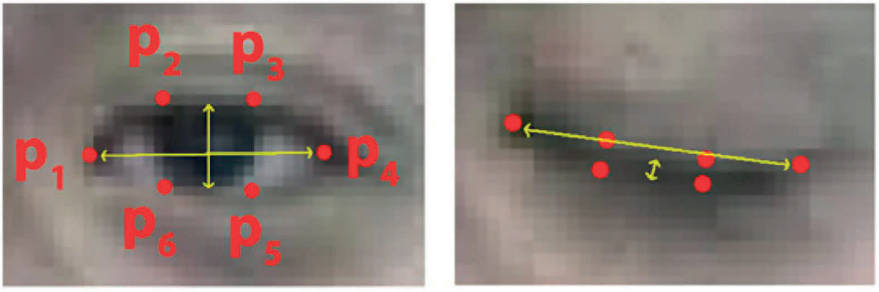


Figure 1.12: Used eye landmarks to calculate the EAR measure in [88].

Massoz et al. in [91] developed a multi-time scale drowsiness detection system to deal with the trade-off between accuracy and responsiveness. They explained that the better accuracy is obtained by longer timescales (long sliding windows) while, the better responsiveness is achieved by shorter timescales (short sliding windows). Twenty-nine subjects (18 females and 11 males) participated in the standard Psychomotor Vigilance Tasks (PVT). Their reaction times while performing the tests were employed to introduce a ground truth of driver drowsiness. Face images of the subjects were also collected using a camera. The participants were sleep-deprived for more than 30 hours before performing the PVTs. Participants completed three 10-min PVTs at different times of the day. Four SVM binary classifiers were trained jointly using four different time scales. The flowchart of their method is presented in Figure 1.13. This method is composed of three main modules: (1) the 'Eye image' module that was used to detect the left and right eyes in the face image. (2) 'Eyelid distance' module that was a spatial CNN with greyscale eye images and eyelid distance estimation of the left and right eyes (d_l and d_r) as input and outputs, respectively. (3) 'Drowsiness' module that was temporal CNN with a 1-min sequence of eyelid distances of both eyes as inputs and probabilities of drowsiness in different time scales as outputs. Results showed that the accuracy of the classifiers using different time scales with the length of 5, 15, 30, and 60 seconds were 70.68%, 85.45%, 89.82%, and 94.22%, respectively. According to these results, longer sliding windows achieved better drowsiness detection accuracies; however, longer sliding windows decrease the capability of fast response and warning in case of a critical level of driver drowsiness.

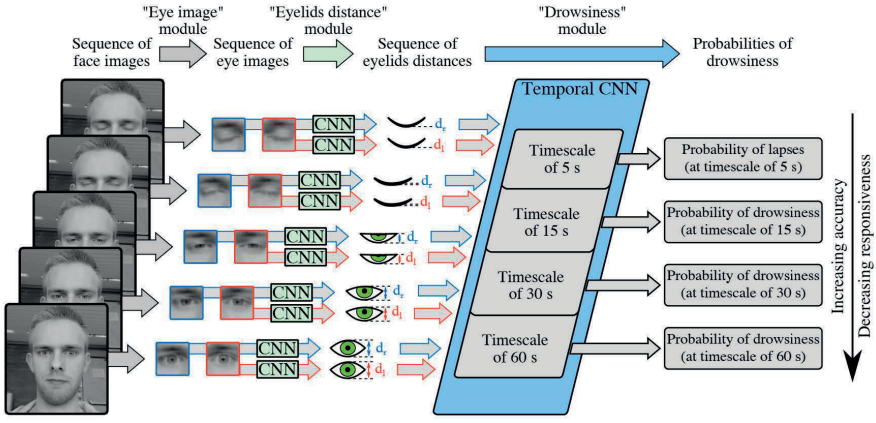


Figure 1.13: Overview of the multi-timescale drowsiness detection system developed based on 1-min sequence of face images in [91].

Deng and Wu [92] proposed a system called DriCare which detected the driver drowsiness states such as yawning and duration of blinking using recorded videos during driving tests. Ten subjects participated in the 1 hour long tests to simulate the drowsy and alert driving states. The flowchart of the DriCare is presented in Figure 1.14. This system is composed of three main parts: face tracking, feature extraction, and drowsiness evaluation. In the face tracking part, a combination of convolutional neural networks (CNN) and Kernelized Correlation Filters (KFC) [93] was used to build the Multiple Convolutional Neural Networks(CNN)-KCF (MC-KCF) algorithm. Moreover, a preprocessing method based on illumination enhancement and histogram equalization [94] was proposed to overcome the issue with changes in illumination intensity during rain or night situations. After preprocessing, the situations of the eyes and mouth were detected using extracted features by CNN. Finally, three different metrics were employed to detect the driver's drowsiness: (1) eye closure duration, (2) blinking rate, and (3) yawning. Results showed that DriCare achieved an accuracy of about 92%.

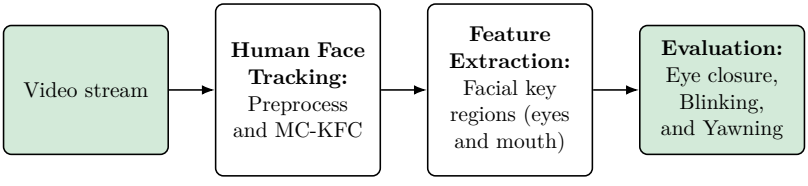


Figure 1.14: DriCare workflow, proposed in [92].

Unscented Kalman Filtering (UKF) [95] was employed in [96] to design a driver drowsiness detection using eye blinking dynamics. Twenty drivers (aged between 25 and 50 years; 15 males, 5 females, 9 with glasses, and 11 without glasses) participated in four different driving tests: (1) eight hours daytime motorway driving, (2) eight hours daytime rural driving, (3) six hours night-time motorway driving, and (4) six hours daytime rural driving. Several video cameras were utilized to capture the subject's face images. Flowchart of the proposed method in [96] is presented by Figure 1.15. As this Figure shows, first, the driver's face was located by applying the Haar algorithm [97]. Second, eye geometric features and image projection techniques were used to detect the eye in the driver's face image. Third, UKF was applied to

track the movements of the detected eyes. Finally, PERCLOS was calculated using the UKF tracking outputs and used to generate the alert sound if the driver was drowsy. Results showed that the proposed method achieved a tracking accuracy of about 99% in all different driving conditions.

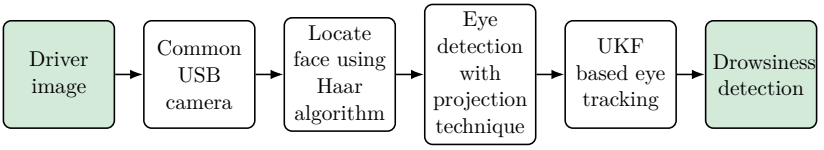


Figure 1.15: flowchart of the UKF-based eye tracking system proposed in [96] for driver drowsiness detection.

A deep belief network (DBN) [98] was employed in [99] for drowsiness detection using facial data. Videos of 30 subjects (aged from 20 to 55 years) were recorded during the driving tests, where drivers simulated different facial behaviors such as talking poses and head rotations. Drivers were subjected to 6-8 hours of sleep deprivation before starting the test. Figure 1.16 shows the structure of the proposed method. In this method, firstly, the landmarks and facial textures were extracted from captured videos using Viola-Jones face detection method [100]. A DBN was then built to classify the driver's drowsiness. Results showed that the average accuracy of this method was 96.7% for drowsiness detection.

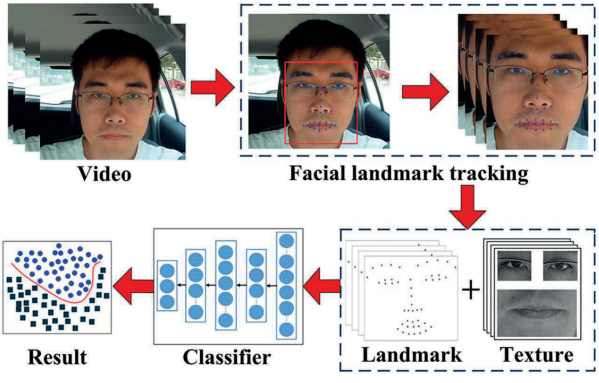


Figure 1.16: Proposed framework in [99] for driver drowsiness expression recognition.

1.2 Objectives of Thesis

Objectives of this thesis are planned in a way to enhance the previous works and address their issues. These objectives are as follows:

1. This thesis proposes a complete pipeline from data preprocessing to classification for driver drowsiness classification using machine learning methods. In this pipeline, multiple features are extracted from different signals, including vehicle-based, ECG, and facial-based data, and the significance of every feature for drowsiness classification is measured using feature selection methods. This pipeline finds important features from input data and reduces the computation burden to design a driver drowsiness detection system.

2. Previous works mainly focused on drowsiness detection in the manual driving mode. However, in the SAE level 3, the driver is allowed to perform a secondary task when the conditional automation system is turned on. Therefore, the driver's vigilance level should be monitored during automated driving to provide a safe control transition from the automated system to the driver when the automated system cannot control the vehicle anymore. In this thesis, the driver drowsiness in the automated mode is evaluated using data fusion of facial-based and biosignals. Vehicle-based data cannot be used for driver state monitoring in the automated mode as the driver inserts no control input to the vehicle.

3. Traditional machine learning models have been widely used in the previous works to classify the driver drowsiness. In those models, features need to be extracted either based on experts' knowledge or by applying dimensionality reduction techniques such as PCA to make the problem tractable for a traditional classifier. However, individual differences between driving styles, lack of experts' knowledge in the field of driver drowsiness detection, and also the risk of losing important information by using dimensionality reduction methods can reduce the classification accuracy of the final system. To remove these limitations, deep convolutional neural networks are employed in this thesis to classify the driver drowsiness using ECG and eyelid data and their ensemble. Deep networks automatically extract the significant features from input data and outperform the traditional machine learning for drowsiness classification.

4. This thesis proposes data fusion frameworks for both of traditional machine learning methods and deep network models. These frameworks investigate the importance of every individual data source and present how using two or more data sources together enhances the performance of the system.

5. This thesis proposes a neural encoder-decoder framework to find significant EEG features that are informative regarding different levels of drowsiness across different drivers. These features are used to design a real-time drowsiness detection using EEG features. By using this framework, the computation load of EEG feature extraction and selection is significantly reduced.

The rest of this thesis is structured as follows:

- **Chapter 2** presents the experimental setup and driving test procedure to collect the data set. The derivation of the ground truth for driver drowsiness is also described by the video observation of the driving tests.

  In the chapters 3 to 5, three different methodologies are proposed to address the shortcomings of the previous studies: (1) traditional machine learning methods, (2) deep neural networks, and (3) encoder-decoder framework applied to the EEG features.

- **Chapter 3** explains the traditional machine learning methods for driver drowsiness classification. First, preprocessing methods are utilized to denoise and prepare the raw collected data, especially biosignals. Then, various hand-crafted features are extracted from different data sources. A feature selection method is applied to improve the system's performance by removing redundancy and irrelevancy in the feature set. Finally, K-nearest neighbours and random forest are employed as two broadly used classifiers to classify the driver drowsiness into three classes: (1) alert, (2) moderately drowsy, and (3) extremely drowsy.

- **Chapter 4** describes the application of different deep neural network architectures for the classification of driver drowsiness. Scalogram 2-d images are extracted from ECG and

eyelid signals and are used separately as inputs to deep convolution neural networks. An ensemble method is exploited to enhance the performance of the individual networks and increase the balanced accuracy of classification.

- **Chapter 5** presents an encoder-decoder dynamical modeling framework applied to the EEG channels for driver drowsiness detection. The percentage of the eyelid closure (PERCLOS) is considered as the ground truth in this framework. Independent Component Analysis (ICA) is used to preprocess the EEG channels and remove the noisy artifacts and eyelid movements. Every EEG channel is decomposed into its sub-bands, including Delta, Theta, Alpha, and Beta waves, and various features are extracted from every sub-band. The extracted features are used in the encoder-decoder framework, and statistically significant features that are consistent across different driving tests are investigated.

- **Chapter 6** summarizes the thesis and suggests some future steps to enhance the proposed methods.

1.2.1 Previous published, submitted or under preparation publications of the thesis

During performing this project, some scientific papers were published or submitted. Some parts of these papers were used in the structure of this thesis that are presented as follows:

- Arefnezhad, S., Hamet, J., Eichberger, A., Frühwirth, M., Ischebeck, A., Koglbauer, I. V., Moser, M., Yousefi, A. (2021). *Driver Drowsiness Estimation Using EEG Signals with a Dynamical Encoder-Decoder Modeling Framework*, Submitted to the Scientific Reports in July 2021 (under review).

  The methods and results of this paper are used to write Chapter 5 that explains the driver drowsiness tracking using EEG data.

- Kaufmann, C., Frühwirth, M., Messerschmidt, D., Moser, M., Eichberger, A., Arefnezhad, S. (2020). *Driving and tiredness: Results of the behaviour observation of a simulator study with special focus on automated driving.* Transactions on Transport Sciences, 11(2), 51-63, DOI: 10.5507/tots.2020.011 [101].

  This paper explains the video-observation process for preparation of the ground truth for driver drowsiness. Some parts of this paper is used to prepare Chapter 2.

- Arefnezhad, S., Eichberger, A., Frühwirth, M., Kaufmann, C., Moser , M. (2020). *Driver Drowsiness Classification Using Data Fusion of Vehicle-based Measures and ECG Signals*, Proceedings of IEEE International Conference on Systems, Man, and Cybernetics 2020 , Virtual Conference, 451-456. DOI: 10.1109/SMC42975.2020.9282867 [102].

  This paper explains the feature extraction from ECG and vehicle-based data for drowsiness classification. The extracted features are used to write the Chapter 3.

- Arefnezhad, S., Samiee, S., Eichberger, A., Frühwirth, M., Kaufmann, C., Klotz, E. (2020). *Applying Deep Neural Networks for Multi-level Classification of Driver Drowsiness Using Vehicle-based Measures.* Expert Systems with Applications, 162 [103].

  This paper describes the structure of deep learning methods used for drowsiness classification. Some parts of this paper are presented in the introductory of Chapter 1, in subsection 2.2 (Driving Tests' Procedure) and also in the subsection of 4.1 (Introduction) and 4.2 (Convolutional Neural Network) in Chapter 4.

- Arefnezhad, S., Samiee, S., Eichberger, A., Nahvi, A. (2019). *Driver Drowsiness Detection Based on Steering Wheel Data Applying Adaptive Neuro-Fuzzy Feature Selection*, Sensors, 19(4), [943]. DOI: 10.3390/s19040943 [104].

  This paper describes the extracted features from steering wheel angle data used for drowsiness classification. Some of these features are presented in Chapter 2.

Chapter 2

Experimental Setup

This chapter describes the experimental setup to collect the used data set in this thesis. The used apparatus (driving simulator), the procedure of driving tests, collected data set, and used methodology to define a ground truth for driver drowsiness based on video observation are explained as follows:

2.1 Apparatus

This thesis is based on a funded project titled WACHsens that was started officially on 01.05.2017 and finished on 31.10.2019. The research was carried out by the consortium composed of Human Research Institut für Gesundheitstechnologie und Präventionsforschung GmbH, Institute of Automotive Engineering at Graz University of Technology (FTG), AVL Powertrain UK, and Factum apptec ventures GmbH and funded by the Austrian Research Promotion Agency (FFG) via the program of mobility of the future (Grant No. 860875).

The WACHsens aims to develop a big database for driver drowsiness classification using a large sample of 92 volunteer drivers and a variety of measured data channels. The collected database is used for the data fusion of vehicle-based, facial-based and physiological signals of the drivers by considering the different human factors such as age and gender.

The study took place using a fixed-base driving simulator named Automated Driving Simulator of Graz (ADSG) at the Graz University of Technology (TU Graz), which is based on an altered production vehicle; see Figure 2.1. The visual cues are simulated by eight LCD panels, covering 180 degrees field of view and the rear screen, which is observed by the inner mirror. The side mirrors are also implemented in the LCDs covering the side windows. The system was designed for autostereoscopic visualization, but this feature was disabled to reduce the simulator sickness during the driving tests. The acoustic cue is simulated by generating engine and wind noise applied at the car's sound system. Moreover, four bass shakers generate the vibration in the car chassis and the driver and passenger seats. Haptic feedback is provided by the Sensodrive simulator steering wheel [105], and an active brake pedal simulator, gas pedal, and gear-shift input are taken from the vehicle unmodified controls. The vehicle dynamics is calculated by a complete vehicle software AVL-VSM™[106], parametrized with a middle-class passenger car. The vehicle model calculates dynamics states as well as engine speed and torque for the acoustic simulation.

Table 2.1: Distribution of drivers in terms of gender and age groups. Std.: standard deviation

Gender	Age range	Mean of age	Std. of age	Number
Female	18-39	25.2	5.3	16
Female	40-59	50.4	6.5	16
Female	60+	65.4	4.3	12
Male	18-39	24.7	3.7	16
Male	40-59	51.9	4.2	16
Male	60+	69.0	7.3	16
—	—	47.0	18.4	Sum: 92

Figure 2.1: The Advanced Driving Simulator of Graz (ADSG) is a modified production car pictured here without the external housing that separates the whole simulator from the environment during the tests (left). The test track from the driver's view simulated a night drive on a highway (right).

2.2 Driving Tests' Procedure

The WACHSens project collected data from 92 test persons in six groups, balanced based on gender and age. Three age groups are considered: 18-39, 40-59, and 60+. Table 2.1 shows the distribution of drivers in terms of gender and age groups. As this Table shows, the lowest number of participants belongs to the Female-60+ group since more participants in this group could not be recruited despite considerable effort.

Participants were enrolled either using a contact database of participants in earlier studies or by recruitment requests in newspapers, local television programs, and hanging posters in the city and at the university campus. Drivers were required to avoid drinking alcohol or an unusual amount of caffeine-content drinks before the tests.

The study was conducted according to the ethical guidelines of the Declaration of Helsinki and the General Data Protection Regulation of the European Union [107]. The study protocol was approved by the Ethics Committee of the Medical University of Graz in vote 30-409 ex 17/18 dated June 1, 2018. Informed consent was obtained from participants before the experiments, and they were also compensated by EUR 50 after finishing the sessions.

Each driver performs two different driving modes, including manual driving and automated driving. In the automated mode, lane-keeping and cruise control systems adjust the vehicle's lateral position and longitudinal speed in the test track, respectively. Drivers complete a 30-min highway track in each of these modes in two states: rested and fatigued. Figure 2.2 shows the simulated test track used for performing the tests. The data of four driving tests per participant

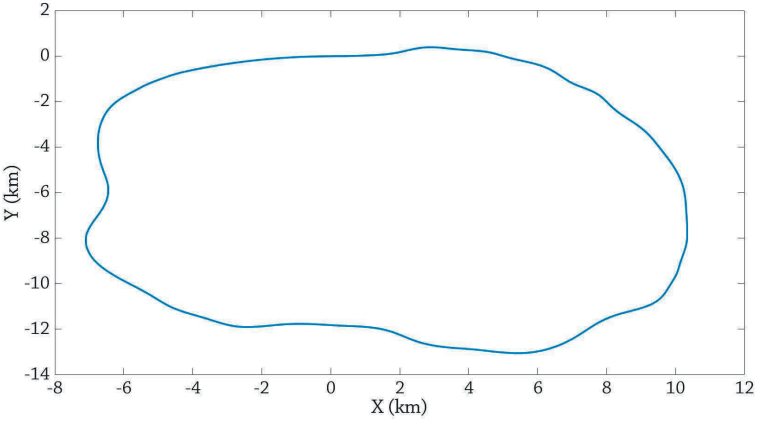


Figure 2.2: The simulated test track in the ADSG to perform the driving tests.

was recorded: rested-manual (RE-M), fatigued-manual (FA-M), rested-automated (RE-A), and fatigued-automated (FA-A). Consequently, for each driving mode, 92 driving tests have been conducted that resulted in a total of 368 tests with a total length of about 183 hours. For the rested state test, participants were asked to stick to a full night's sleep before the test and not deviate from their usual circadian cycle.

To induce drowsiness in drivers for the fatigued state test, researchers have proposed different procedures. For instance, Kartsch et al. [108] reported that in their study, drivers were under sleep deprivation (they slept only three hours the day before the test), and experiments were also conducted at late night hours. McDonald et al. [23] conducted the fatigued state test after a minimum of 18 h of wakefulness for their participants. In the study of Braua et al. [109], three fatigued state tests were run between 22.00 h and 06.15 h, and drivers were instructed to go to bed no later than 24.00 h and to get up no later than 9.00 h. In the DROZY dataset [90], subjects completed three (Psychomotor Vigilance Test) PVTs within about 36 hours. They had a sleep deprivation of 28 to 30 hours for the third PVT and were instructed to stay at the lab for about 30 hours for conducting the three PVTs. In the PVTs, drivers were instructed to press a button on a screen when they noticed a red box over the black background. There is also a counter that measures the reaction time of the drivers to these red boxes [110].

In the WACHSens project, two options were offered to the participants to be prepared for the fatigued state tests: one option was to stay awake for at least 16 h continuously before starting the test procedure (resulting in a drive after at least 17 to 18 hours of wakefulness) and to take the test at their usual bed-time. Another option was a sleep restriction of at least 50% (max. 4 hours of sleep) the night before. This latter option was generally preferred by younger people, as they were more likely used to late bed-times, whereas elderly participants more often felt uncomfortable with sleep restrictions and preferred delaying their usual bed-time. In both cases, the setting was used to increase the probability of a marked contrast of alertness and drowsiness during the driving tests but was not used for labeling the data. Effective drowsiness labels are independent of the setting. Levels of drowsiness and corresponding labels change throughout the drive and are not bound to specific times or the setting. A particular participant might be classified as falling asleep during the alert test, or he/she might not be classified as drowsy during the fatigued test at all. Such cases exist but are rare. Derivation of the drowsiness labels as ground truth for classification is explained in section 2.3.

Table 2.2: Measured vehicle-based data in each driving test.

Number	Measured Signal	Symbol	Unit
1	Absolute position X	x	m
2	Absolute position Y	y	m
3	Absolute position Z	z	m
4	Roll angle	ϕ	rad
5	Pitch angle	θ	rad
6	Yaw angle	ψ	rad
7	Longitudinal speed	v_x	m/s
8	Longitudinal acceleration	a_x	m/s^2
9	Lateral acceleration	a_y	m/s^2
10	Steering wheel angle	δ	deg
11	Lateral deviation from road center line	d	m
12	Brake pedal	p_b	–
13	Throttle pedal	p_t	–
14	Engine speed	n	rpm
15	Current gear	i_g	–

Before starting the tests, subjects fill out a questionnaire form about their demographic information and driving experiences. Furthermore, experimenters implement biosignal sensors and prepare the driving scenario in the simulator. These tasks also take about 1 to 1.5 hours. Before the start of each driving test, the driver's drowsiness is subjectively evaluated using the Karolinska Sleepiness Scale (KSS) [111] and objectively by PVT [112].

A monotonous driving test track that simulates a night drive on a highway without traffic events has been designed to induce drowsiness. The driven track simulates a three-lane motorway that has some smooth turns, and participants drive about 50 km during the test. More details of the driving test procedure are explained in [113].

2.3 Collected Dataset

The collected data in the driving tests are categorized into three groups: (1) vehicle-based data, (2) facial-based data, and (3) biosignals. The collected signals in each of these groups are described as follows:

Vehicle-based data: This type of data was collected and verified using the AVL-VSM™vehicle dynamics simulation software [106]. The sampling frequency for collecting this data is 100 Hz. Table 2.2 presents the vehicle-based data collected in the driving tests.

Facial-based data: This non-invasive type of data is collected using an eye-tracking system called SmartEye™[114]. This device captures the head movements, eyelid movements, pupil diameter, and gaze vectors. The quality of the gathered signals is also provided to check the reliability of every sample of these signals. Figure 2.3 shows a snapshot of the SmartEye™output. Two squares on the left-top and right-top of this figure show the recognized pupils and their diameters. Two curvy blue lines around the eyes are showing the detected eyes which are used for measuring the distance between eyelids. Two red lines originated in the eyes are also utilized to collect the gaze vectors. Table 2.3 presents the collected facial-based data using SmartEye™system. The sampling frequency of this data is also 100 Hz.

Table 2.3: Measured facial-based data from SmartEye device in each driving test.

Number	Measured Signal	Unit
1	Time stamp of SmartEye	-
2	Frame number of SmartEye	-
3	Head position X in coordinate system	m
4	Head position Y in coordinate system	m
5	Head position Z in coordinate system	m
6	Head position quality	-
7	Gaze direction X in coordinate system	m
8	Gaze direction Y in coordinate system	m
9	Gaze direction Z in coordinate system	m
10	Gaze direction quality	-
11	Eyelid opening of left eye	m
12	Quality of eyelid opening of left eye	-
13	Eyelid opening of right eye	m
14	Quality of eyelid opening of right eye	-
15	Mean of eyelid opening	m
16	Quality of mean of eyelid opening	-
17	Pupil diameter of left eye	m
18	Quality of pupil diameter of left eye	-
19	Pupil diameter of right eye	m
20	Quality of pupil diameter of right eye	-
21	Mean of pupil diameter	m
22	Quality of mean of pupil diameter	-
23	Head rotation X	rad
24	Head position Y	rad
25	Head position Z	rad
26	Gaze origin X in coordinate system	m
27	Gaze origin Y in coordinate system	m
28	Gaze origin Z in coordinate system	m

Figure 2.3: Performance of the SmartEye™eye tracking system.

Biosignals: The biosignals collected in every driving test include 8 channels of EEG, two channels of EOG (installed in the vertical direction on the upper and lower parts of the left eye), respiration signal, electrodermal activity, and two ECG electrodes. The EEG channels measure brain activity using an g.Nautilus Research ™EEG cap [115] worn by drivers.

Figure 2.4 shows the positions of the collected EEG channels that include Fz, T7, T8, C3, C4, Cz, PO7, and PO8. To gather ECG signals and heart rate, 5 electrodes are attached to the chest of the drivers in the locations shown in Figure 2.5. As shown in this Figure, two blue electrodes are used to measure the heart rate using ChronoCord™(designed by Human Research Institute) [116] and three red electrodes are utilized for the g.Nautilus device to gather ECG channels. Electrodermal activity is measured using a system depicted in Figure 2.6. As this figure shows, the used electrodes are fastened to the left food the drivers (A and B positions in Figure 2.6) and then are attached to a Galvanic Skin Response (GSR) sensor. The left foot was chosen since no movements are expected from this foot in the automated gear-shift vehicle. Figure 2.7 shows the exploited sensor to collect respiration data when the reference clip is connected to the right ear. Biosignals are gathered with a sampling frequency of 500Hz. The g.Nautilus data were transferred via Bluetooth and recorded synchronously with vehicle and facial data. Table 2.4 shows the gathered biosignals in each driving test.

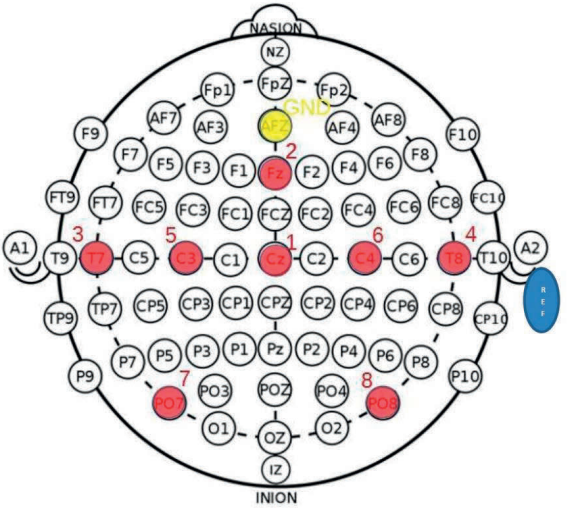


Figure 2.4: Locations of the collected EEG channels on EEG cap in the 10-20 system [117].

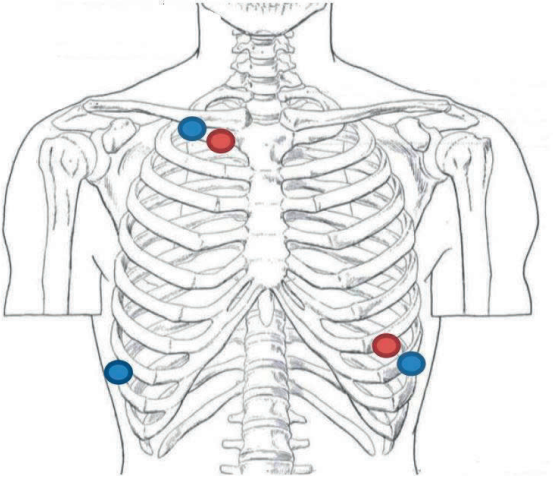


Figure 2.5: Locations of the attached electrodes to collect heart rate and ECG signals [118].

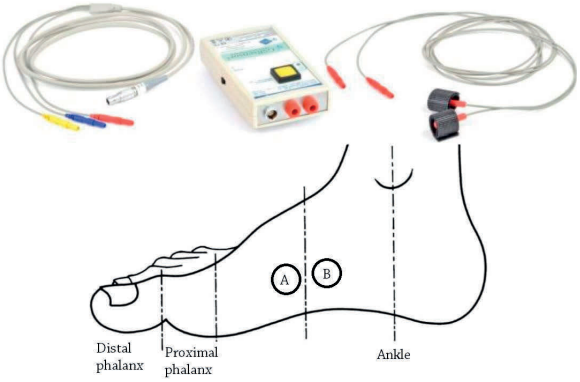


Figure 2.6: Used setup to measure electrodermal activity in the driving tests, adopted from [119]

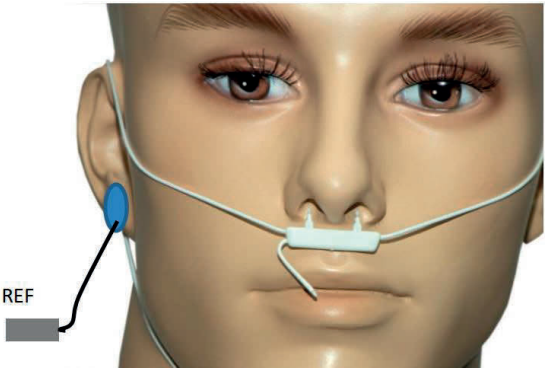


Figure 2.7: Used setup to measure respiration in the driving tests [120].

Table 2.4: Measured biosignals in each driving test.

Number	Measured Signal	Unit
1	EEG channel Cz	µV
2	EEG channel Fz	µV
3	EEG channel T7	µV
4	EEG channel T8	µV
5	EEG channel C3	µV
6	EEG channel C4	µV
7	EEG channel P07	µV
8	EEG channel P08	µV
9	EEG channel Ref.	µV
10	ECG lower lead	µV
11	ECG upper lead	µV
12	EOG lower lead	µV
13	EOG upper lead	µV
14	Respiration	µV
15	Electrodermal activity	µV

2.4 Ground Truth of Drowsiness based on Video Observation

To monitor the driving behavior, four cameras are placed inside the vehicle, and each of these cameras provides information about the actions of the driver from different views, see Figure 2.8. In literature, various strategies are exploited to obtain a ground truth for driver drowsiness, including subjective evaluation using questionnaires (e.g., Karolinska Sleepiness Scale), EEG signal processing, and driving video observation. In this study, video recordings are visually rated by an expert in driving observation. These drowsiness evaluation rules have been designed based on the driving video observation strategies proposed in [28] and [121].

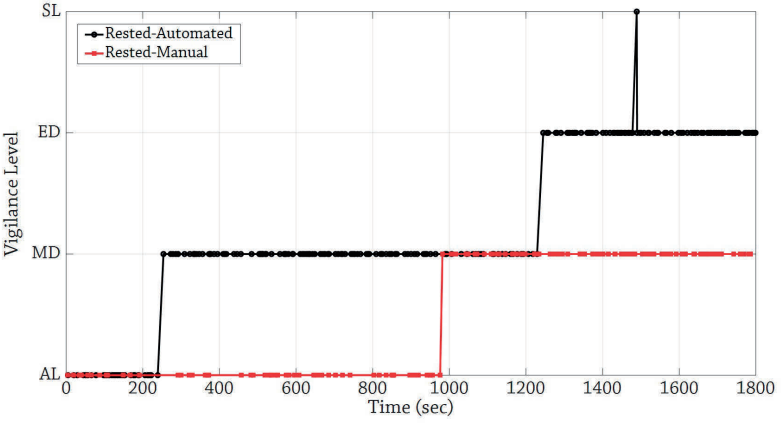
Figure 2.8: Four different views from a driver that are gathered using installed cameras inside the simulator.

For the drowsiness rating, the position and activities of the arms, hands, upper body, head, and facial movements are considered. The output of the driving video observation is one of the considered levels of drowsiness that include alert, moderately drowsy, and extremely drowsy. The indicators for each of these drowsiness levels are explained as follows.

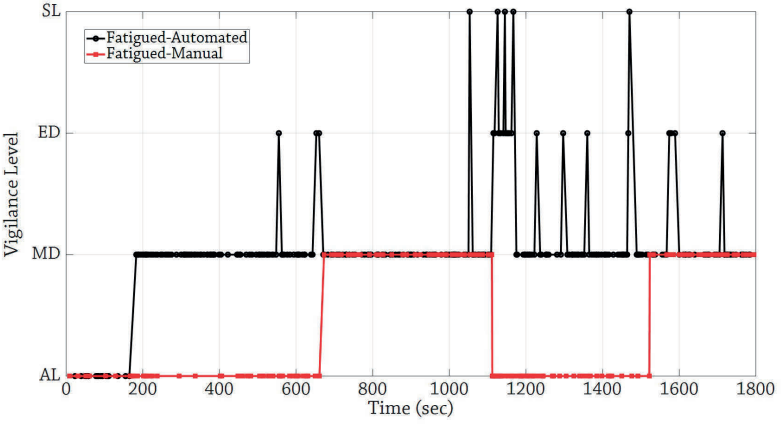
- Alert: normal facial tone, evident focus on driving, fast eye blinking.
- Moderately drowsy: face or eyes rubbing, restless posture, yawning, rigid look forward, hardly any movements within a short period, reduction of eye blink duration, and the degree of eyelid opening.
- Extremely drowsy: a rapid decline in the ability to respond, clear signs that the subject is struggling to keep awake, change in eyelid movements (faster and heavier), a significant increase in the duration of eye closure, and keeping the eyes closed for either a micro-sleep or longer.

Rated vigilance levels of the same driver in four different driving modes are shown in Figure 2.9. As this Figure presents, micro-sleep events (SL) are also reported by video observers. However, the SL events are merged with the class of extremely drowsy (ED) since the number of SL events are too low to be considered as a separate class. This Figure also shows that even in the rested tests, the driver has shown the signs of moderately drowsy (MD) or extremely drowsy (ED) states.

Overall, driver drowsiness has been rated in 196 driving tests (out of 368) and for 50 drivers (out of 92 drivers) by video observation. Table 2.5 shows the distribution of the video-observed tests per each driving mode. More details of the defined ground truth based on the video observation have been explained in [101].



(a)



(b)

Figure 2.9: Drowsiness levels rated using video observation in different driving modes performed by the same driver: in the rested (a) and in the fatigued (b) tests; AL: Alert, MD: Moderately Drowsy, ED: Extremely Drowsy, SL: Sleep

Table 2.5: Distribution of the video-observed tests per each driving mode.

Driving mode	No. of video-observed driving tests
RE-A	48
RE-M	49
FA-A	49
FA-M	50

Chapter 3

Driver Drowsiness Classification using Traditional Machine Learning Methods

3.1 Introduction

In this chapter, a traditional machine learning workflow is proposed to design a driver drowsiness classification system where labels of video observation are used as ground truth for drowsiness. The main processes of the traditional machine learning methods for classification problems are categorized into four steps: 1) Preprocessing, 2) Feature extraction, 3) Feature selection, and 4) Adjusting or training the classifier. Figure 3.1 shows the flowchart of the structure of traditional machine learning methods.

The goal of the preprocessing step is to reduce the influence of noise and artifacts on the raw input data. This step helps to improve the value of the signal-to-noise ratio. Standardization of input data is also categorized as a preprocessing method to normalize the effect of input data on the output results. Feature extraction is the process of extracting significant information from input data using data mining methods. These features can be extracted either based on expert knowledge in a specific application or by using data dimensionality reduction methods such as Independent Component Analysis (ICA) [62] and deep autoencoders [122]. Feature selection aims to select a subset of features that carry more important information than other features for mapping the input data to the output class labels.

Feature selection also helps to reduce the computation burden by removing redundant and irrelative features in classification and regression problems. In supervised learning, both the input data and class labels are assumed to be available, and classifiers are designed to reduce the defined error function.

In the rest of this chapter, a methodology based on traditional machine learning is proposed, and the classification results of this methodology are discussed.

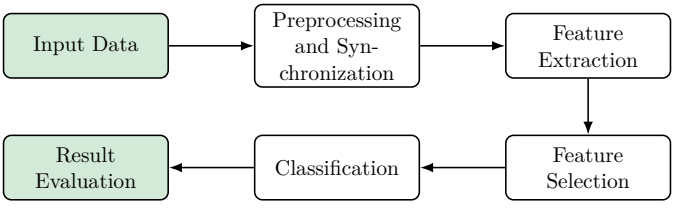


Figure 3.1: Flowchart of the traditional machine learning process.

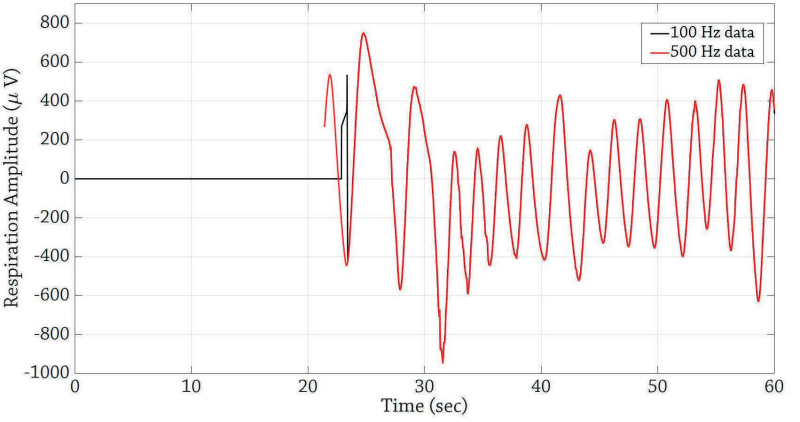


Figure 3.2: Data synchronization using respiration signals collected with two different frequencies of 100 Hz and 500 Hz. In this example, the 500 Hz data is shifted about 21.4 s forward to be synced with the 100 Hz data.

3.2 Preprocessing and Synchronization

Vehicle-based and ECG signals need to be preprocessed before feature extraction and classification of drowsiness. As SmartyEye™[114] provides already preprocessed facial-based information, its raw outputs are used as inputs to the feature extraction process. This section explains the synchronization of the data channels and preprocessing of vehicle-based and ECG signals.

3.2.1 Data Synchronization

Ground truth is defined based on the video observation and it is recorded using the frame number information of SmartEye data that are collected with a sampling frequency of 100 Hz. However, used ECG and EEG signals are gathered with a sampling frequency of 500 Hz. Therefore, these two sources of data should be synchronized to classify the driver drowsiness using ECG signals. To perform this synchronization, the respiration signal that is collected with two sampling rates of 100 Hz and 500 Hz is utilized. The normalized cross-correlation between the two respiration signals is calculated at all possible lags. The delay between these two signals is calculated as the lag that achieves the largest absolute value of normalized cross-correlation. Figure 3.2 shows an example of data synchronization where 500Hz respiration data is shifted about 21.4 s forward to be synchronized with the 100Hz respiration data. The exact time shift is also applied to the ECG signals collected with the sampling frequency of 500 Hz to sync them with the video observations.

3.2.2 Derivation of Heart Rate Variability Information

ECG signals are collected with a sampling frequency of 500Hz. A second-order Infinite Impulse Response (IIR) notch filter [123] is used to remove the powerline noise (50 Hz) from ECG signals. ECG signals are corrupted with baseline wander noise [124] that is a low-frequency noise. A high pass filter with a passband frequency of 0.5 Hz is employed to remove this low-frequency noise. Figure 3.3 shows the noisy and denoised ECG signal after preprocessing step. This Figure shows that this filter has a suitable performance to remove baseline wander noise from ECG signals.

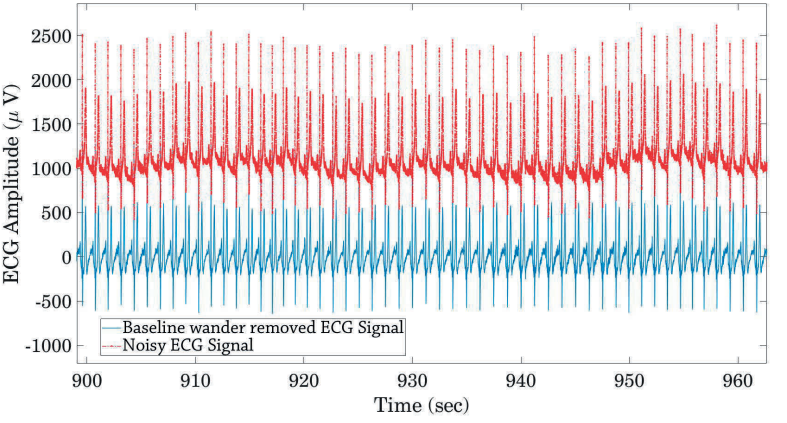


Figure 3.3: The performance of the baseline wander removal filter applied to ECG signals; Red: Noisy ECG signal and Blue: denoised ECG signals

Heart rate variability signal is extracted from ECG signals by using an R-peak detection algorithm. First, the Automatic Multiscale-based Peak Detection (AMPD) method that has been presented in [125] is applied as an ECG R-peak detector, then RR Intervals (RRIs) as the times passed between two adjacent R-peaks are calculated. Since the RRI series is not a signal sampled in equal time intervals, a spline interpolator is exploited to generate an equally sampled RRI signal. Figure 3.4 shows the detected R-peaks in the denoised ECG signal using the AMPD method. A part of the derived RRI signal is also shown in Figure 3.5.

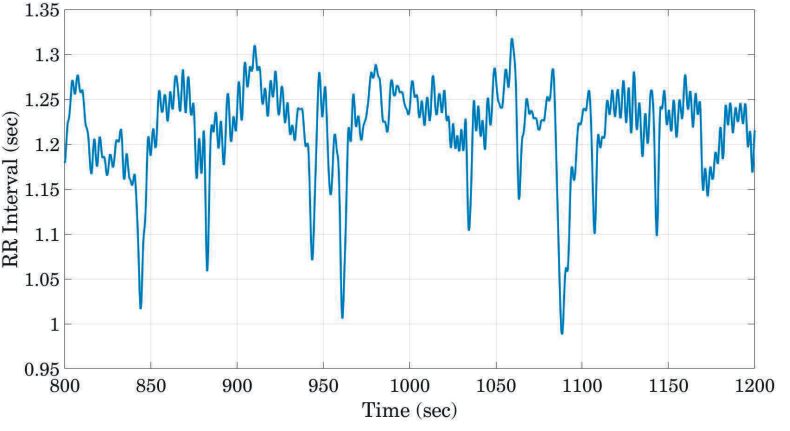


Figure 3.5: A part of the derived RRI signal after detecting the R-Peaks of the ECG signal in a rested-automated test. Since the time interval between adjacent R-peaks is different for every peak, spline interpolation is used to generate a uniformly-sampled signal.

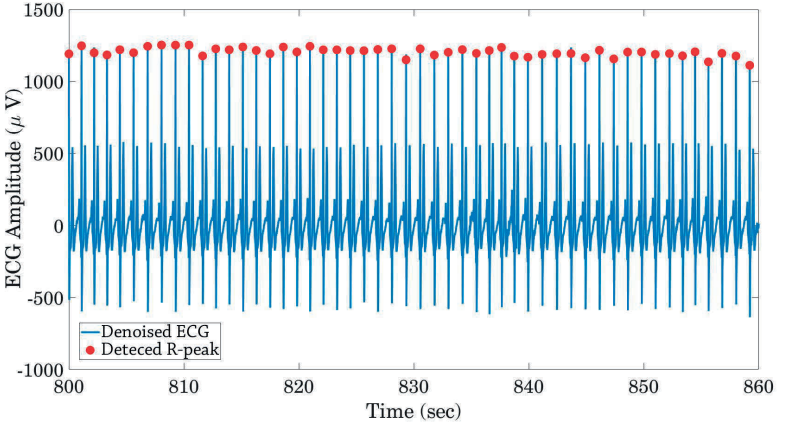


Figure 3.4: Detected R-peaks of the denoised ECG signal using the AMPD method in a rested-automated test. The detected R-peaks are used to obtain the heart rate variability signal based on the time interval between every two adjacent R-peaks.

3.3 Feature Extraction

3.3.1 Extracted Features from Vehicle-based Data

Vehicle-based data include steering wheel angle, steering wheel angular velocity, lateral deviation from road centre-line, lateral acceleration, and yaw rate. The extracted features from vehicle-based data are divided into two main groups: (1) Features that are extracted from all of the vehicle-based, and (2) Features that are extracted exclusively from steering wheel angle and steering angular velocity. These featrues are extracted from every sliding window of the data.

Features that are extracted from all of the vehicle-based:

Consider $X = \{x_1, x_2, ..., x_N\}$ to be a sliding window of any of the vehicle-based data where N is the number of samples in every sliding window. The following features are extracted from all of the vehicle-based data:

1. Average (avg): The average value of vehicle-based data.

$$Avg(x) = \frac{1}{N}\sum_{i=1}^{N} x_i, \tag{3.1}$$

2. Standard deviation (Std): Dispersion of the data around average value.

$$Std(x) = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (x_i - Avg(x))^2}, \tag{3.2}$$

3. Energy (Ene): Sum of the square of signal magnitude.

$$Ene(x) = \sum_{i=1}^{N} (x_i)^2, \tag{3.3}$$

4. Zero Crossing Rate (ZCR): This feature explains the rate of the signal sign changes [126].

$$ZCR(x) = \frac{1}{N}\sum_{i=1}^{N}|s(x_i) - s(x_{i-1})|, \tag{3.4}$$

where $s(x_i) = 1$ if the signal $x(i) > 0$ and $s(x_i) = 0$ if $x(i) \leq 0$.

5. First Quartile (Q_1): Middle number between the smallest number and the median of the signal.

6. Second Quartile (Q_2): Median of the signal in the sliding window.

7. Third Quartile (Q_3): Middle value between the median and the highest value of the signal.

8. Skewness ($Skew$): This is a measure of the asymmetry of the data probability distribution.

$$Skew(x) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - avg(x)}{std(x)}\right)^3, \tag{3.5}$$

9. Approximate Entropy ($ApEn$): This feature describes the complexity and unpredictability of the the data variations [127]. Larger values of the $ApEn$ indicate the difficulty of prediction of new samples and more irregularity in the signal. In order to calculate the $ApEn$, two inputs are required: (1) m that is a positive integer which presents the embedded dimension of the vector produced for calculating the $ApEn$ and (2) r that presents the filter factor. First, m dimensional vectors of $Y(1)$, $Y(2)$, ..., $Y(N-m+1)$ are generated from every sliding window that are defined by $Y(i) = [x(i), x(i+1), ..., x(i+m-1)]$; $i = 1, 2, ...N-m+1$. Second, the distance between $Y(i)$ and $Y(j)$ are calculated as:

$$d[Y(i), Y(j)] = \max_{k=0,..,m-1}[|x(i+k) - x(j+k)|], \tag{3.6}$$

Third, For each $i = 1, .., N-m+1$ the $C_r^m(i)$ are computed as

$$C_r^m(i) = \frac{P}{N-m+1}, \tag{3.7}$$

where P is the number of times that $d[Y(i), Y(j)]$ is less than or equal to r.

Forth, the quantity of $\phi^m(r)$ is calculated as:

$$\phi^m(r) = \frac{1}{N-m+1}\sum_{i=1}^{N-m+1}\ln C_r^m(i), \tag{3.8}$$

This process is repeated $m+1$ times. Finally, the $ApEn$ is calculated:

$$ApEn(m, r, N) = \phi^m(r) - \phi^{m+1}(r). \tag{3.9}$$

Here, m is equal to 2 and r is set to be $0.2 \times Std(x)$.

10. Shannon Entropy ($ShEn$): This feature explains the average level of information carried out by data [128].

$$ShEn(x) = -\sum_{i=1}^{N} p(x_i)\log_2 p(x_i), \tag{3.10}$$

where $p(x_i)$ is the probability of the occurring x_i in the X. To calculate this probability, number of repeats of every samples is saved and normalized in a way that the sum of probability over the samples is one.

11. Spectral Entropy ($SpEn$): This feature is the Shannon Entropy of the signal's normalized power distribution [129]. To calculate this feature, the discrete Fourier transform of the signal is first computed. Then, power spectrum of the signal is calculated as the square magnitude of the computed discrete Fourier transform. This power spectrum is normalized and treated as the probability input in the Shannon entropy equation. The spectral entropy is calculated as:

$$SpEn(x) = -\sum_{i=1}^{N} s_x(x_i) \log_2 s_x(x_i), \tag{3.11}$$

where s_x is the normalized power distribution of the signal.

12. Dominant Frequency (DF): The frequency that has the largest magnitude of the power spectral density [130].

Features that are extracted exclusively from steering wheel angle and steering angular velocity:

1. Amplitude duration squared delta (AmpD2Delta) [131]: To calculate this feature in every sliding window, firstly, the mean of steering angle (δ_m) should be subtracted from steering angle data (δ) to have $\delta - \delta_m$. Every two consecutive time points that their corresponding $\delta - \delta_m$ are zero make a block, and the area under each block of the curve of $\delta - \delta_m$ is calculated. Finally, AmpD2Delta is computed as presented in Equation (3.12).

$$AmpD2Delta = \frac{100}{N} \sum_{j=1}^{J} (A_j^{\delta}.t_j^{\delta}), \tag{3.12}$$

where A_j^{δ} is the are of the j-th block under the $\delta - \delta_m$ curve, t_j^{δ} is the length of the j-th block, J is the total number of blocks, and N is the total number of samples in the sliding window [131]. Figure 3.6 shows the curve of the $\delta - \delta_m$ and its corresponding variables to calculate the AmpD2Delta.

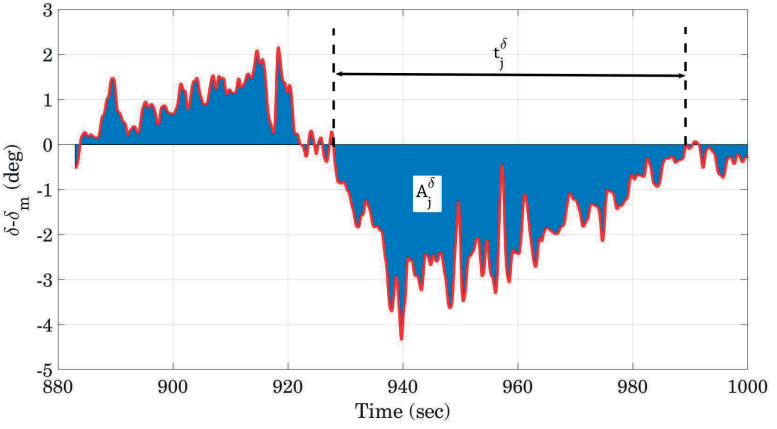


Figure 3.6: The used variables to calculate AmpD2Delta from steering wheel angle in a rested-manual test.

2. Ellipse-based features [131]: These types of features are based on the phase plot of the steering wheel angle and steering wheel angular velocity and a predefined control ellipse. Figure 3.7 shows the plot of steering wheel angular velocity ($\dot{\delta}$) versus steering wheel angle (δ). This figure contains the control ellipse that penalizes the steering behavior outside the ellipse. Steering behavior inside the ellipse is considered normal. Here, the centre of the control ellipse is located in the origin (0,0) and the major and minor semi-axes are set as 20 deg/s and 5 deg, respectively.

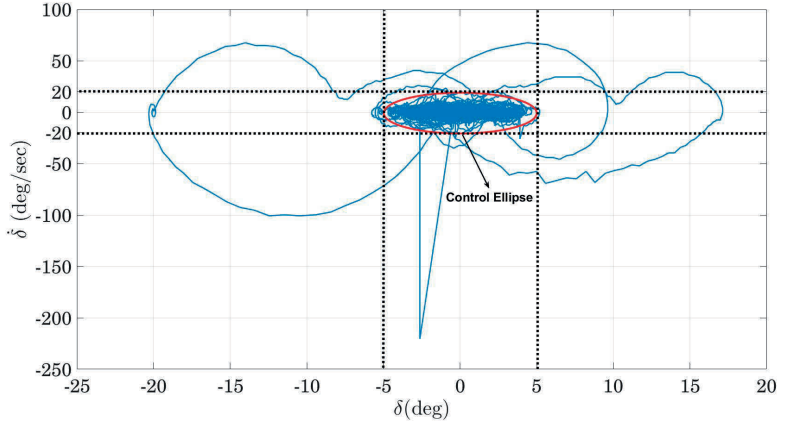


Figure 3.7: The phase plot of steering wheel angle and steering wheel angular velocity with control ellipse in a rested-manual test.

Two features are extracted using the control ellipse and the phase plot:

(a) Outside Percentage (P_{out}) [131]: Outside Percentage is defined as the percentage of

$(\delta, \dot{\delta})$ points per window outside the control ellipse as presented in Equation (3.13).

$$P_{out} = 100\frac{n}{N}, \tag{3.13}$$

where n is the number of points per window outside of the control ellipse, and N is the total number of points in each window.

(b) Weight Flat Zero (W_{FZ}) [131]: Only points $(\delta_i, \dot{\delta}_i)$ in the phase space satisfying the condition $|\delta| \leqslant \delta_c$ with $\dot{\delta}_c = 20$ deg/s are included in calculation. All points satisfying this condition are weighted by the square of the distance from the origin. Equation (3.14) shows how to calculate the W_{FZ} feature (only if $|\dot{\delta}_i| \leqslant \delta_c$).

$$W_{FZ} = \frac{100}{N}\sum_{i=1}^{N}(\frac{(\delta_i - \delta_m)^2}{a^2} + \frac{\dot{\delta}_i^2}{b^2}), \tag{3.14}$$

Where δ_i is the i-th value of steering angle, δ_m is the average of steering angle in the window, a is the half axis length of control ellipse in δ dimension, b is the half axis length of control ellipse in $\dot{\delta}$ dimension, $\dot{\delta}_i$ is the i-th value of steering angular velocity and N is the total number of points in the window.

3. NMRHold: NMRHold is the number of times that the amplitude of steering wheel angle is held less than 0.5 degrees ($|\delta| \leqslant 0.5$ deg) for longer than 0.04 s [16].

Table 3.1 presents the extracted features from vehicle-based data for driver drowsiness classification.

Table 3.1: Extracted features from vehicle-based data

Feature	Description
Average	Average value of the signal in every sliding window
Standard Deviation	Dispersion of the data around mean value
Energy	Sum of the square of signal magnitude
Zero-Crossing Rate	Number of steering or steering velocity direction changes per second
First Quartile	Middle number between the smallest number and the median of the signal in sliding window
Second Quartile	Median of the signal in the sliding window
Third Quartile	Middle value between the median and the highest value of the signal in sliding window
Skewness	A measure for signal similarity
Approximate Entropy	Complexity of signal in time domain based on distance in embedding dimension
Shannon Entropy	Complexity of signal in time domain based on probability function
Spectral Entropy	Complexity of signal in frequency domain
Dominant Frequency	The frequency that has the maximum value of the PSD
Amplitude duration squared delta	Described in 3.3.1, see Equation (3.12)
Outside Percentage	Described in 3.3.1, see Equation (3.13)
Weight Flat Zero	Described in 3.3.1, see Equation (3.14)
NMRHold	Described in 3.3.1

3.3.2 Extracted Features from ECG Signals

Literature has proposed some features to be extracted from RR intervals for driver drowsiness detection [132]. Some of these features are based on a visualization technique called the Poincaré plot. In this subsection, firstly, this plot is introduced, then commonly extracted features from RR intervals are explained.

- **Poincaré plot:** This plot is a type of recurrence plot to investigate the similarity in time series that can be used to analyze the nonlinear properties of HRV data [133]. Consider $X = [RR_1, RR_2, ..., RR_M]$ as a RR interval time series presenting M heartbeats. The Poincaré plot first plots (RR_1, RR_2), then plots (RR_2, RR_3), then plots (RR_3, RR_4) and so on. This plot provides information about the short-term and long-term dynamics of the RR interval. An ellipse is fitted to the plotted data points, and the minor and major semi-axes of the ellipse are associated with short-term and long-term HRV, respectively. Figure 3.8 shows the Poincaré plot for RR intervals collected in a rested-automated driving test. The least-square method has been employed to fit an ellipse on given RR intervals [134] and geometrical properties of this ellipse are extracted as features to describe the HRV dynamics.

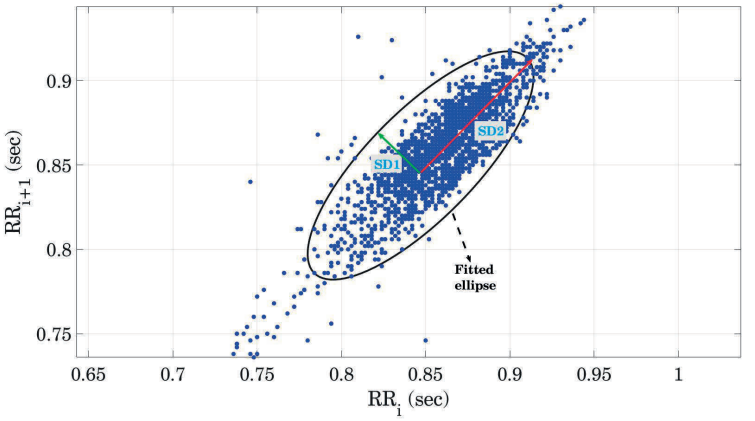


Figure 3.8: Poincaré plot and fitted ellipse for RR Intervals during a rested-automated driving test. Minor and major semi-axes of the fitted ellipse, SD1 and SD2, are calculated as two features to capture the dynamics of HRV data.

Three features are extracted from this plot:

1. SD1: SD1 is the standard deviation of the Poincaré plot perpendicular to line-of-identity and semi-minor axis (half of the shortest diameter) of the fitted ellipse, see Figure 3.8 (green vector). SD1 is an estimation of short-term HRV that describes parasympathetic activity since it represents the deviation of heart rate from the line-of-identity (constant heart rate) and shows how the heart rate is indifferent regarding the present level of drowsiness [135].
2. SD2: SD2 is the standard deviation of the Poincaré plot along the line-of-identity and semi-major axis (half of the largest diameter) of the fitted ellipse, see Figure 3.8 (red vector). SD2 is an estimation of long-term HRV that describes sympathetic activity since SD2 is along the line-of-identity and shows how the heart rate changes with changes in driver drowsiness [135].
3. SD1/SD2: SD1/SD2 is the ratio of SD1 to SD2 that describes the ratio of short-term to long-term HRV and the ratio between the strength of parasympathetic and sympathetic activities [135].

Other features that have been proposed by previous studies [38, 136, 137] are also extracted from RR intervals. These features are:

1. MeanRR: It presents the mean values of the time intervals between every two consecutive R-peaks [138].

$$MeanRR = \frac{1}{N_R - 1} \sum_{i=1}^{N_R - 1} RR_{i+1}, \tag{3.15}$$

where N_R is the number of heart beats in the sliding window and RR_{i+1} is equal to $R_{i+1} - R_i$.

2. SDRR: This feature represents the standard deviation of time points that the detected R-peaks are occurred in them [138].

$$SDRR = \sqrt{\frac{1}{N_R}\sum_{i=1}^{N_R}(R_i - \bar{R})^2}, \tag{3.16}$$

where $\bar{R}$ is the average of the time points corresponding to the R-peaks in every sliding window [139].

3. SDSD: This feature represents the standard deviation of RR intervals [139].

$$SDSD = \sqrt{\frac{1}{N_R - 1}\sum_{i=1}^{N_R}(RR_{i+1} - MeanRR)^2}, \tag{3.17}$$

4. RMSSD: This feature calculates the root mean square of the consecutive RR intervals' differences [36].

$$RMSSD = \sqrt{\frac{1}{N_R - 1}\sum_{i=1}^{N_R-1}(RR_{i+1} - RR_i)^2}, \tag{3.18}$$

5. pRR50: This feature measures the ratio of the number of R-peaks that differ more than 50 ms from their next R-peak divided by the total number of R-peaks in every sliding window [36].

$$pRR50 = \frac{RR50_{count}}{N_R}, \tag{3.19}$$

6. VLF: This feature presents the power in the very-low-frequency ranges of 0.003-0.04 Hz of the RR interval time series [38]. In order to calculate this feature and also the LF and HF, the PSD of the RR intervals are computed using Lomb-Scargle periodogram method [140, 141] in every sliding window.

7. LF: This feature presents the power in the low-frequency ranges of 0.04-0.15 Hz of the RR interval time series [38].

8. HF: This feature presents the power in the high-frequency ranges of 0.15-0.40 Hz of the RR interval time series [38].

9. LF/HF: This feature is the ratio of LF divided by HF [38].

The extracted features from RR intervals are listed in Table 3.2.

3.3.3 Extracted Features from Facial-based Data

Facial-based data that are employed for driver drowsiness classification are eyelid opening and pupil diameter signals. These signal are measured for both of right and left eyes and also for the average values of both eyes with the sampling frequency of 100 Hz. The average and standard deviation of the measurement quality of these signals over all performed driving tests are presented in Table 3.3. In this Table, Q_{EL} and Q_{ER} are respectively the average quality of the left and right eyelid signal. The Q_{PL} and Q_{PR} are respectively the average quality of the left and right pupil signal. Finally, Q_{EM} and Q_{PM} are respectively the quality of the average eyelids and pupils of the left and right eyes. As this Table presents, Q_{EM} and Q_{PM} are higher than the corresponding values for only the left or right eyes. Therefore, the average eyelid

Table 3.2: Extracted features from RR intervals derived from ECG signals

Feature	**Description**
Mean RR	Mean of R-to-R (RR) intervals
SDRR	Standard deviation of RR intervals; estimate of overall HRV
SDSD	Standard deviation of differences between adjacent RR intervals; estimate of short-term HRV; describes parasympathetic activity
RMSSD	Square root of the mean of the sum of the squares of differences between adjacent RR intervals; estimate of short term HRV; describes parasympathetic activity
pRR50	Number of pairs of adjacent RR intervals differing by more than 50 ms divided by the total number of all RR intervals
SD1	Standard deviation of the Poincaré plot perpendicular to line-of-identity; semi-minor axis of the fitted ellipse; estimate of short-term HRV; describes parasympathetic activity
SD2	Standard deviation of the Poincaré plot along the line-of-identity; semi-major axis of the fitted ellipse; estimate of long-term HRV; describes sympathetic activity
SD1/SD2	Ratio of SD1 to SD2; describes the ratio of short-term to long-term HRV; describes the ratio between parasympathetic and sympathetic activity
VLF	Power in the very-low-frequency range (0.003–0.04 Hz); describe the oscillations from the heart's intrinsic nervous system and sympathetic nervous system
LF	Power in the low-frequency range (0.04–0.15 Hz); estimate of long-term HRV; reflects both sympathetic and parasympathetic activity
HF	Power in the high-frequency range (0.15–0.4 Hz); estimate of short-term HRV; describes parasympathetic activity
LF/HF	Ratio between LF and HF range powers; describes the ratio of long-term to short-term HRV; describes the ratio between sympathetic and parasympathetic activity

signal (EM) and the average pupil diameter (PM) are utilized in this thesis to classify driver drowsiness using facial-based data. Figure 3.9 and Figure 3.10 show parts of the eyelid opening signal and pupil diameter, respectively.

Table 3.3: The average (Avg.) and standard deviation (Std.) of the measurement quality of the eyelid and pupil diameters signals during all of the driving tests.

–	Q_{EM}	Q_{EL}	Q_{ER}	Q_{PM}	Q_{PL}	Q_{PR}
Avg. (Std.)	**0.78** (0.26)	0.68 (0.32)	0.75 (0.34)	**0.55** (0.21)	0.40 (0.20)	0.48 (0.25)

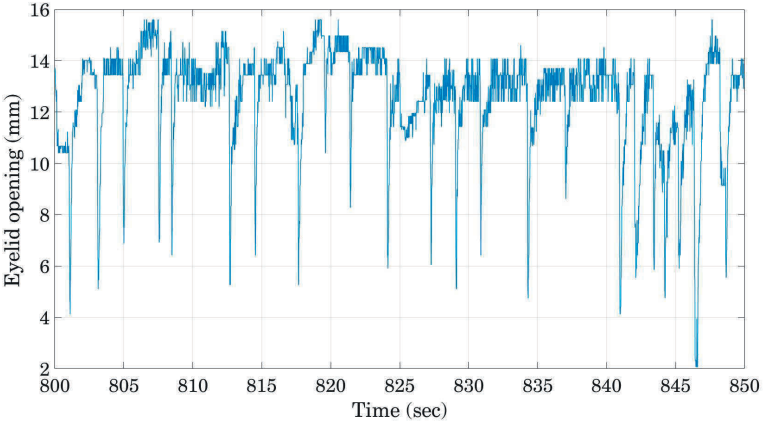


Figure 3.9: A part of the eyelid opening signal (mean value for both eyes) in a rested-automated driving test. This signal is collected as the raw output of the SmartEye eye-tracker system.

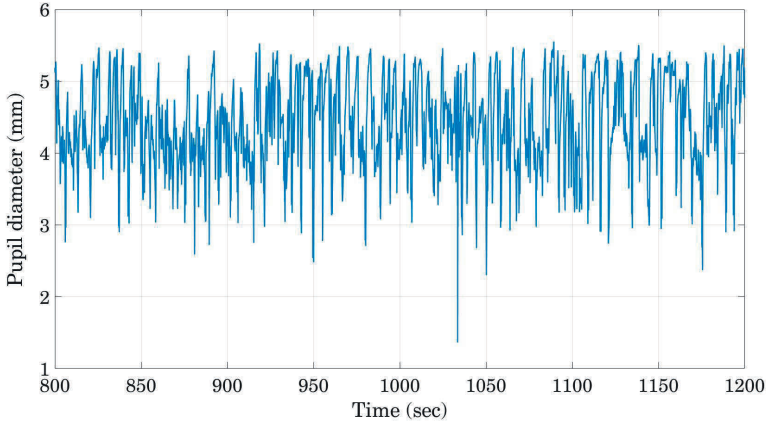


Figure 3.10: A part of the pupil diameter (mean value for both eyes) in a rested-automated driving test. This signal is collected as the raw output of the SmartEye eye-tracker system.

Extracted features from facial-based data are listed as follows:

1. PERcentage of eyelid CLOSure (PERCLOS): PERCLOS is defined as the proportion of the time window that eyes are at least 80% closed in every minute [142]. This feature is calculated using the eyelid opening signal. Figure 3.11 presents a sample eye blink in a driving test that is smoothed using a spline smoother. The maximum eyelid opening signal in this driving test is 17.5 mm; consequently, if the eyelid opening amplitude is less than 3.5 mm (20% of the maximum value), eyes are at least 80% closed.

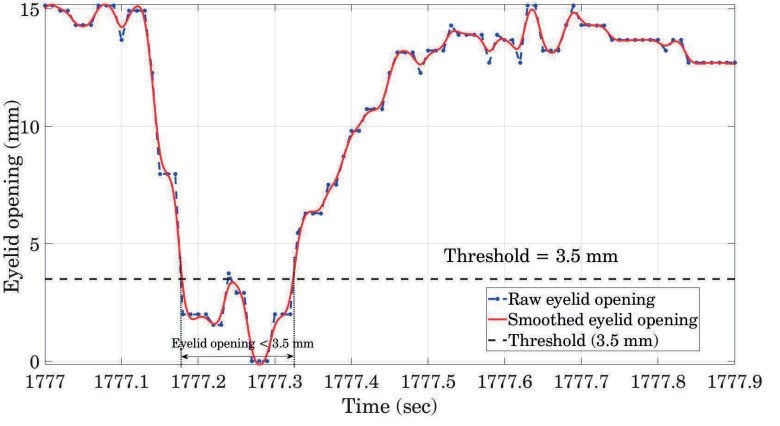
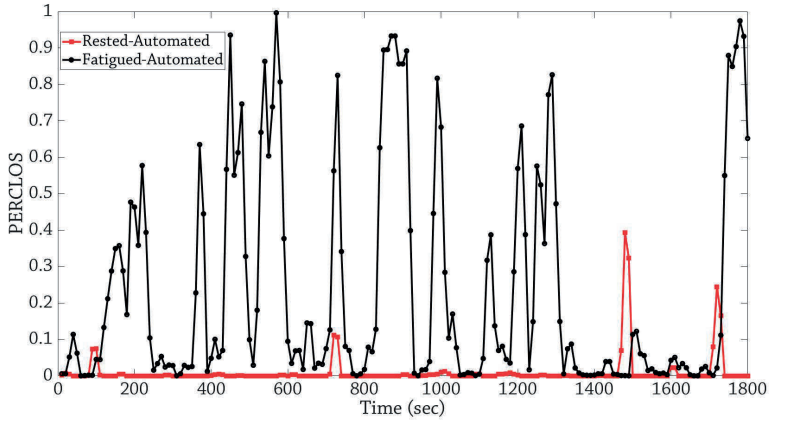
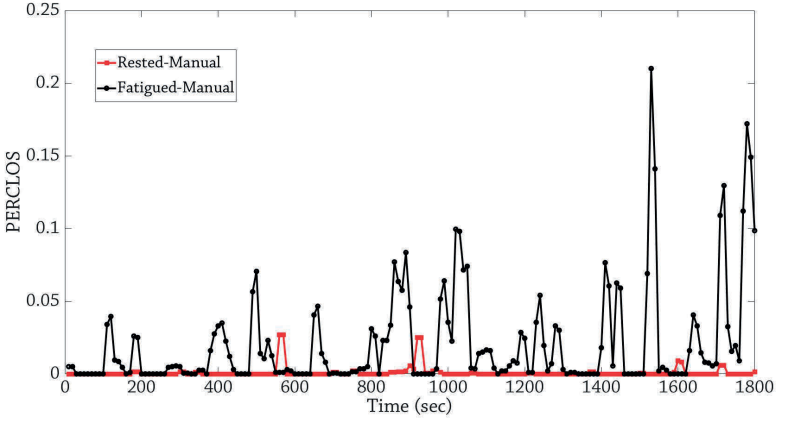


Figure 3.11: A sample eye blink that has been smoothed with a spline smoother and defined threshold (here is 3.5 mm) to calculate the PERCLOS feature.

It has been reported that larger PERCLOS values are associated with the higher levels of drowsiness [142, 143, 144]. Calculated PERCLOS outputs for four different driving modes of one driver are shown in Figure 3.12. A sliding window with the length of 20 s and overlap of 10 s (50%) is used to calculate the PERCLOS. As this Figure shows, PERCLOS increases up to 0.9 in the fatigued-automated test and up to 0.22 in the fatigued-manual test and this means that the driver is drowsy in the fatigued mode tests. This Figure also shows that the PERCLOS is very low (less than 0.05) during the rested-manual test. PERCLOS is less than 0.1 during the first 24 minutes of the rested-automated test and it increases to about 0.3 several times in the last six minutes of the test.



(a)



(b)

Figure 3.12: Calculated PERCLOS in the automated tests (a) and manual tests (b). A sliding window with a 20 s length and a 10 s overlap has been exploited to calculate PERCLOS from eyelid opening data.

2. Blinking frequency: Blinking frequency is defined as the number of detected blinks in every sliding window divided by the length of the window in seconds. Figure 3.13 presents the detected blinks in a part of the eyelid opening signal.

Table 3.4: Used signals from every data source and number of extracted features from each data source.

Data source	Used Signals	No. of extracted features
Vehicle-based	Lateral acceleration, lateral deviation form road center line, steering wheel angle, speed, and yaw rate	64
ECG signals	RR intervals	12
Facial-based	Eyelid opening and pupil diameter	6
–	Sum	82

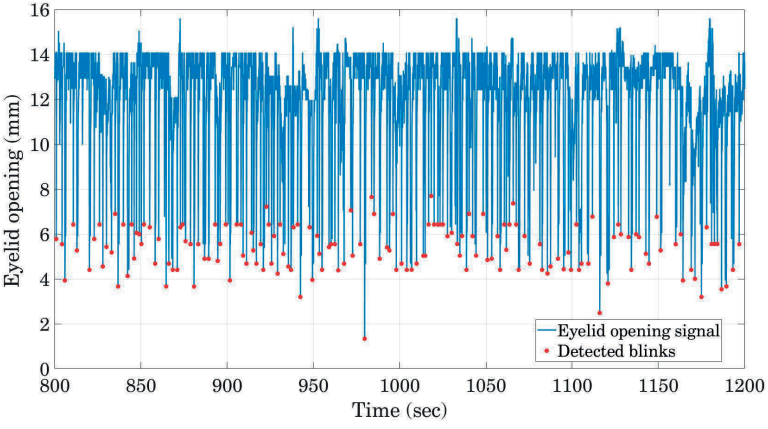


Figure 3.13: A part of the eyelid opening signal and detected blinks to calculate the blinking frequency.

3. Mean of pupil diameter
4. Standard deviation of pupil diameter
5. Mean of eyelid opening
6. Standard deviation of eyelid opening

Table 3.4 shows the used signals and the number of the extracted features from all three sources of the data (vehicle-based data, ECG signals, and facial-based data). As this Table shows, in overall, 82 features are extracted from the data of every driving test (vehicle-based features are only extracted in the manual driving tests). Some of these features might be non-informative or redundant to the driver drowsiness classes. The next section describes the feature selection process to select a subset of these features that have strong relationship with driver drowsiness levels.

3.4 Feature Selection

Feature selection is a process to select a subset of features from the multidimensional data space to enhance the accuracy of classifiers, decrease the computational burden in the classification process, and a better understanding of the data set structure in machine learning applications [145]. The real-world data set may have redundant and dependent features that cannot provide additional information about the class labels. Moreover, some of the features usually are not correlated to the class labels and can produce bias noise for the classifiers and reduce its classification performance [146]. In this thesis, Neighbourhood Component Analysis (NCA) is employed as a feature selection method that is described in the next subsection.

3.4.1 Neighbourhood Component Analysis for Feature Selection

Neighbourhood component analysis (NCA) is a non-parametric feature selection algorithm that is applied to maximize classification accuracy. This method helps to learn a linear transformation of input features to maximize the accuracy of the k-nearest neighbors in the transformed space. The feature weights are adjusted such that the leave-one-out probability of correct classification is maximized [147].

Assume that a training set with n samples is given by Equation (3.20).

$$S = \{(x_i, y_i), i = 1, 2, ..., n\}, \tag{3.20}$$

where x_i is p-dimensional feature vector, $y_i = \{1, 2, ...c\}$ are the class labels, and c is the number of classes. The goal is to build a classifier that maps the input feature vector x to its correct class label y. In the NCA, a randomized classifier is considered which (1) randomly chooses a sample from S, ref(x) as the reference point for x and (2) assigns a label to the input sample x same as the label of ref(x). Since the reference point is selected randomly, all of the training samples in S have some probability to be selected as the reference point. The probability of selecting x_j as the reference point for x is shown as P(ref(x) = $x_j|S$) which is higher if x_j is closer to the x. In order to select the reference point, the following distance function is used [148].

$$d_w(x_i, x_j) = \sum_{r=1}^{p} w_r^2 |x_{ir} - x_{jr}|, \tag{3.21}$$

where w_r is the vector of feature weights. In the NCA, it has been assumed that

$$P(\text{ref}(x) = x_j|S) \propto \kappa\left(d_w(x, x_j)\right), \tag{3.22}$$

that means the probability of the selecting x_j as the reference point for x is proportional to the kernelized distance between x and x_j where $\kappa(.)$ is the kernel function and defined as

$$\kappa(z) = exp\left(\frac{-z}{\sigma}\right), \sigma > 0, \tag{3.23}$$

where σ is the width of the kernel function [148]. Using different σ changes the probability of each training sample being picked to be the reference point for x. If σ is very small, only the nearest samples to the x have a high probability. On the other hand, if σ is very large, all of the points have approximately the same probability. If all of the input feature set is scaled to have the same range, it is recommended to set σ equal to 1. As the sum of P(ref(x) = $x_j|S$) over all samples of S should be equal to 1, the Equation (3.24) can be defined [149].

$$P(\text{ref}(x) = x_j|S) = \frac{\kappa\left(d_w(x, x_j)\right)}{\sum_{j=1}^{n} \kappa\left(d_w(x, x_j)\right)}, \tag{3.24}$$

The leave-one-out performance of this classifier is the final output that is predicting the label of x_i using the data in S^{-i} (the S excluding the (x_i, y_i)). The Equation (3.25) presents the probability of selecting point x_j as the reference point for x_i [149].

$$p_{ij} = P(\text{ref}(x_i) = x_j | S^{-i}) = \frac{\kappa(d_w(x_i, x_j))}{\sum_{j=1, j\neq i}^{n} \kappa(d_w(x_i, x_j))}, \tag{3.25}$$

The average leave-one-out probability of correct classification, p_i is defined by Equation (3.26). This Equation shows the probability of correctly classifying the x_i using S^{-i} [149].

$$p_i = \sum_{j=1, j\neq i}^{n} P(\text{ref}(x_i) = x_j | S^{-i}) I(y_i = y_j) = \sum_{j=1, j\neq i}^{n} p_{ij} y_{ij}, \tag{3.26}$$

where $y_{ij} = I(y_i = y_j)$ is

$$\begin{cases} 1 & \text{if } y_i = y_j, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, the average leave-one-out probability of correct classification can be defined as

$$h(w) = \frac{1}{n} \sum_{i=1}^{n} p_i, \tag{3.27}$$

where the $h(w)$ depends on the weight vector w [149]. The aim of the NCA is to maximize the $h(w)$ with respect to the w, but to reduce the risk of overfitting in classification, the regularized objective function is used instead [149]. This function is defined as

$$F(w) = h(w) - \lambda \sum_{r=1}^{p} w_r^2 = \frac{1}{n} \sum_{i=1}^{n} F_i(w), \tag{3.28}$$

where λ is the regularization parameter that makes many of the weight in w equal to 0 and $F_i(w)$ is

$$\sum_{j=1, j\neq i}^{n} p_{ij} y_{ij} - \lambda \sum_{r=1}^{p} w_r^2,$$

If the λ is given, the optimal weight vector w is provided by solving the following minimization problem [149].

$$\hat{w} = \underset{w}{\text{argmin}}\, f(w) = \underset{w}{\text{argmin}} \frac{1}{n} \sum_{i=1}^{n} f_i(w), \tag{3.29}$$

where $f(w) = -F(w)$ and $f_i(w) = -F_i(w)$. After applying some justification to the Equation (3.18), it can be rewritten as Equation (3.30) [149].

$$\hat{w} = \underset{w}{\text{argmin}} \{ \frac{1}{n} \sum_{i}^{n} \sum_{j=1, j\neq i}^{n} p_{ij} l(y_i, y_j) + \lambda \sum_{r=1}^{p} w_r^2 \}, \tag{3.30}$$

where $l(y_i, y_j)$ is

$$\begin{cases} 0 & \text{if } y_i = y_j, \\ 1 & \text{otherwise.} \end{cases}$$

This optimization problem was solved using a gradient descent optimizer in [150]. A threshold is set and only feature vectors whose corresponding weights are higher than this threshold are

Table 3.5: Distribution of the drowsiness levels in the feature set extracted from the data of **manual** driving tests. The feature set is imbalanced, and the alert and extremely drowsy classes are its majority and minority classes, respectively.

Reference Drowsiness level	Number of samples	Percentage of every level
Alert	12,142	63.88%
Moderately drowsy	4,610	25.94%
Extremely drowsy	1,019	5.73%
Sum	17,771	-

Table 3.6: Distribution of the drowsiness levels in the feature set extracted from the data of **automated** driving tests. The feature set is imbalanced data set that alert and extremely drowsy classes are its majority and minority classes, respectively.

Reference Drowsiness level	Number of samples	Percentage of every level
Alert	9,476	55.75%
Moderately drowsy	5,274	31.03%
Extremely drowsy	2,248	13.22%
Sum	16998	-

selected since features with small weights are not informative enough to influence the classification performance. Moreover, they can be redundant and decrease the classification accuracy, and increase the computation complexity.

3.5 Classification of Driver Drowsiness

3.5.1 Imbalanced Dataset Issue

To extract features from input data sources, a sliding window with a length of 20 sec and an overlap of 10 sec between every pair of adjacent windows is exploited. The video observations are also interpolated to the sampling frequency of 100 Hz with this assumption that driver's vigilance is constant until the next change in the video observation is reported. The same sliding window is also applied to the video observations. The most frequent drowsiness level (mode) in every sliding window is used as the final label for drowsiness classification. Thus, the driver's drowsiness level is being classified every 10 sec (except the first window that needs 20 sec of the data). The extracted feature set is an imbalanced data set since the number of data samples that belong to the extremely drowsy class is much less than two other classes. The issue of the imbalanced dataset is inherent in drowsy driving classification problems since long periods of extreme drowsiness will lead to accidents. Table 3.5 and Table 3.6 present the distribution of the observed samples per drowsiness level in the manual and automated driving tests, respectively. As Table 3.5 shows, in the manual driving test, only 5.73% of the samples belong to the class of extremely level of drowsiness while 63.55% of the samples are in the alert class. This means that minority and majority classes are extremely drowsy and alert, respectively. Table 3.6 shows that the imbalanced data set issue is reduced in the automated driving tests since departing the lane is avoided by the lane keeping assist function, allowing the vehicle to keep the lane even for a sleeping driver. This means that the possibility of feeling drowsy in the automated tests was higher than in manual tests.

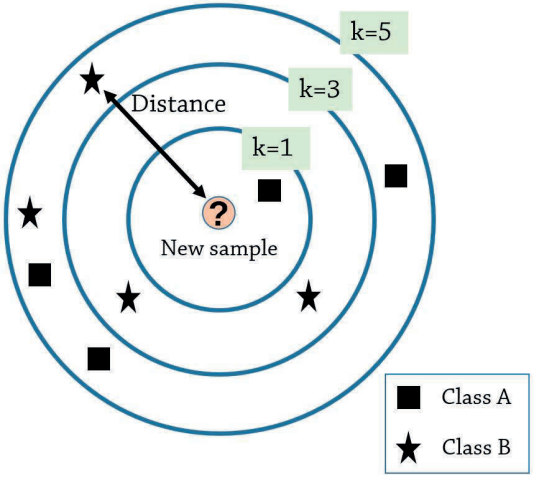


Figure 3.14: Illustration of the k-NN algorithm for a binary classification problem, adapted from [155]

Classification of an imbalanced data set can lead to biased results inclined to the majority class. To solve this issue, several techniques have been proposed, including resampling the data set (oversampling the minority class or/and undersampling the majority class) [151, 152] and class-weighted classification [153]. The following subsection explains the used classifiers and employed methods to mitigate the issue of the imbalanced data set.

3.5.2 Classifiers

Two commonly used and simple classifiers have been exploited to classify the three observed drowsiness levels in drivers: K-nearest neighbors and random forest. These classification methods have been briefly explained as follows.

1. K-nearest neighbors (KNN): KNN that has been firstly introduced in [154] is a non-parametric method widely used for classification and regression problems. In the classification problems, the class of each feature sample is assigned based on the class distribution of its k samples in its neighborhood. In other words, the new sample will be assigned to a class that is most common among its k nearest neighbors. Figure 3.14 illustrates the KNN algorithm for a binary classification problem [155]. As this Figure shows, there are two classes: (1) Class A (square), and (2) Class B (star), and the new sample (shown inside the orange circle by a question mark) should be classified as Class A or Class B. If $k = 1$, the new sample is classified in class A since its nearest sample is a square (Class A). If $k = 3$, the new sample belongs to Class B since there are two samples from Class A and only one sample from Class B in the neighbourhood of the new sample. If $k = 5$, the new sample is classified to Class A since there are more samples from this class in the neighbourhood of the new sample.

   This method has two hyperparameters: the number of neighbors in the feature space (k) and the distance metric to compute the distance between the feature of interest and any other samples in the feature space, e.g., Minkowski distance, Manhattan distance, and Euclidean distance [156]. Moreover, the prior distribution of the training samples is assumed to be uniform here to reduce the imbalanced dataset issue for driver drowsiness classification.

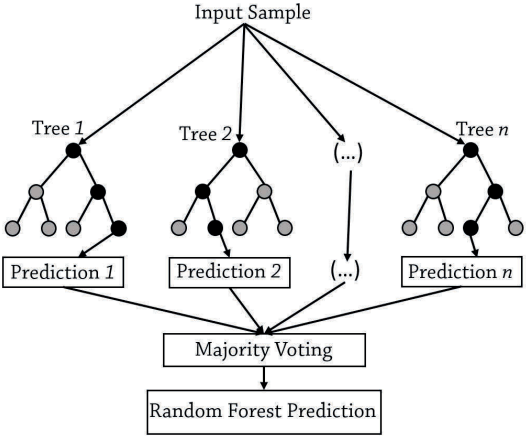


Figure 3.15: Illustration of the random forest algorithm, adapted from [161].

2. Random forest: Random forest is a tree-based classifier or regressor that overcomes the limitations of the decision trees, including loss of generalization accuracy on unseen data and suboptimal accuracy on training data [157]. This method combines the output of multiple decision trees using a majority voting (in classification problems) or averaging (in regression problems) methodology to produce the final output. Figure 3.15 presents the random forest classification problem. Here, the Random Under Sampling Boosting (RUSBoost) algorithm [158] is used in the structure of random forest to handle the issue of the imbalanced data set. This algorithm employs the combination of Random Under-Sampling (RUS) [159], and the standard Adaptive Boosting (AdaBoost) [160] to model the minority samples by random elimination of the majority class samples.

To evaluate the used classifiers, the elements of the confusion matrices of the test dataset are calculated. These matrices provide four different values that are calculated for every drowsiness level:

1. True-Negative (TN): Number of samples that do not belong to the specific class (for example alert) and are also classified in any of the two other classes (moderately drowsy or extremely drowsy) by the classifier.

2. True-Positive (TP): Number of samples that belong to the specific class (for example alert) and are also correctly classified in the same class.

3. False-Negative (FN): Number of samples that belong to the specific class (for example alert) but are wrongly classified in any of the two other classes (moderately drowsy or extremely drowsy) by the classifier.

4. False-Positive (FP): Number of samples that do not belong to the specific class (for example alert) but are wrongly classified in the same class.

These four values are used to calculate the four different metrics for every level of driver drowsiness:

1. Specificity (true negative rate): specificity is the ratio of TN divided by the sum of TN and FP.

2. Sensitivity (true positive rate): Sensitivity is the ratio of TP divided by the sum of TP and FN.

3. Precision (positive predictive value): Precision is the ratio of TP divided by the sum of TP and FP.

4. F1-score: F1-score is the harmonic mean of precision and sensitivity. Harmonic mean (m_h) of two scalar variables of x_1 and x_2 is computed as follows:

$$m_h = \frac{2x_1x_2}{x_1 + x_2} \tag{3.31}$$

3.6 Results

This section describes the results of the drowsiness classification when different data sources are used as input to the classifiers. The performance of the system is reported when the following feature sets are utilized as input for classifiers:

1. Only vehicle-based features (vehicle-based data are only used for driver drowsiness classification in the manual driving tests),

2. Only ECG-based features,

3. Only facial-based features,

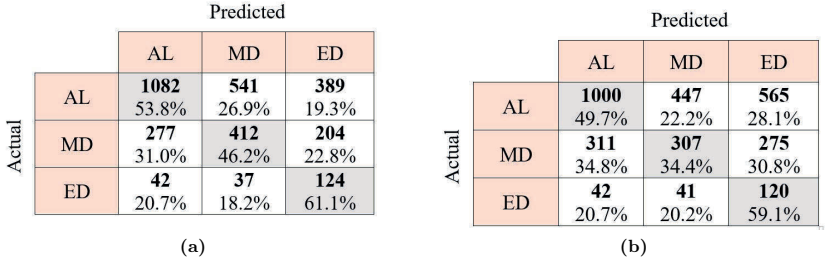
4. All of the extracted features from different data sources (no feature selection),

5. Only selected features by NCA from all of the extracted features.

This procedure helps to investigate the influence of each data source on the performance of the driver drowsiness classification system separately. For evaluation of the designed system, every data set is separated randomly into two sets: training (80% of the samples) and test (20% of the samples). In order to evaluate the performance of the system to classify different levels of drowsiness, balanced accuracy is utilized. Balanced accuracy is the average of the accuracies over three drowsiness levels.

3.6.1 Results of Drowsiness Classification in the Manual Driving Tests

Confusion matrices of the KNN and random forest classifiers applied to the (1) vehicle-based features, (2) ECG-based features, (3) facial-based features, (4) all of the extracted features (no feature selection), and (5) only selected features by NCA method in the manual tests are provided in Figure 3.16 to Figure 3.20, respectively. The grey elements in these matrices represent the number of the correctly classified input features. Accordingly, the percentage numbers written in these elements show correct classification accuracy for every specific drowsiness level.



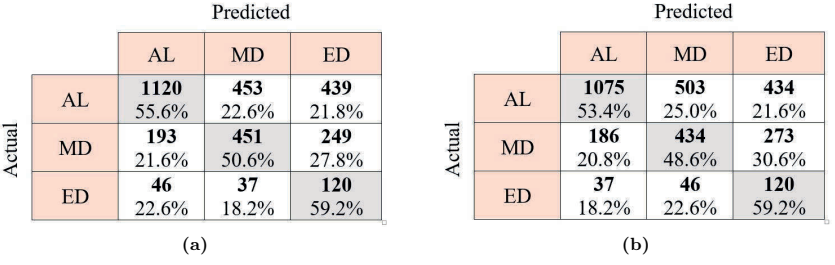
(a)

Actual \ Predicted	AL	MD	ED
AL	**1082** 53.8%	**541** 26.9%	**389** 19.3%
MD	**277** 31.0%	**412** 46.2%	**204** 22.8%
ED	**42** 20.7%	**37** 18.2%	**124** 61.1%

(b)

Actual \ Predicted	AL	MD	ED
AL	**1000** 49.7%	**447** 22.2%	**565** 28.1%
MD	**311** 34.8%	**307** 34.4%	**275** 30.8%
ED	**42** 20.7%	**41** 20.2%	**120** 59.1%

Figure 3.16: Confusion matrices of random forest (a) and KNN (b) classifiers in the **manual** tests when only **vehicle-based** features are used. Gray elements represents the true-positive classified data samples. AL: Alert, MD: Moderately Drowsy, and ED: Extremely Drowsy.



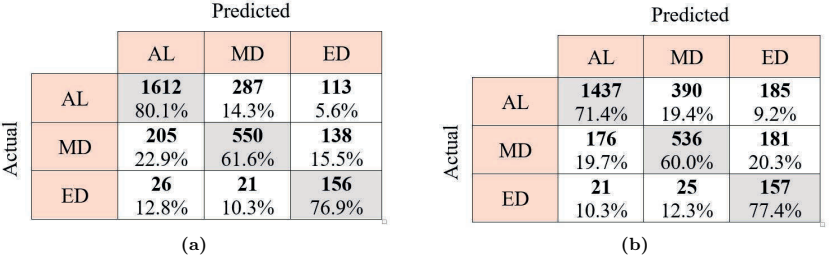
(a)

Actual \ Predicted	AL	MD	ED
AL	**1120** 55.6%	**453** 22.6%	**439** 21.8%
MD	**193** 21.6%	**451** 50.6%	**249** 27.8%
ED	**46** 22.6%	**37** 18.2%	**120** 59.2%

(b)

Actual \ Predicted	AL	MD	ED
AL	**1075** 53.4%	**503** 25.0%	**434** 21.6%
MD	**186** 20.8%	**434** 48.6%	**273** 30.6%
ED	**37** 18.2%	**46** 22.6%	**120** 59.2%

Figure 3.17: Confusion matrices of random forest (a) and KNN (b) classifiers in the **manual** tests when only **ECG-based** features are used.



(a)

Actual \ Predicted	AL	MD	ED
AL	**1612** 80.1%	**287** 14.3%	**113** 5.6%
MD	**205** 22.9%	**550** 61.6%	**138** 15.5%
ED	**26** 12.8%	**21** 10.3%	**156** 76.9%

(b)

Actual \ Predicted	AL	MD	ED
AL	**1437** 71.4%	**390** 19.4%	**185** 9.2%
MD	**176** 19.7%	**536** 60.0%	**181** 20.3%
ED	**21** 10.3%	**25** 12.3%	**157** 77.4%

Figure 3.18: Confusion matrices of random forest (a) and KNN (b) classifiers in the **manual** tests when only **facial-based** features are used.

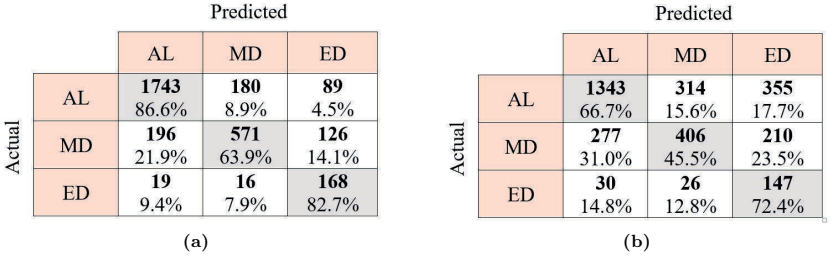


Figure 3.19: Confusion matrices of random forest (a) and KNN (b) classifiers in the **manual** tests when **all** of the extracted features (no feature selection) are used.

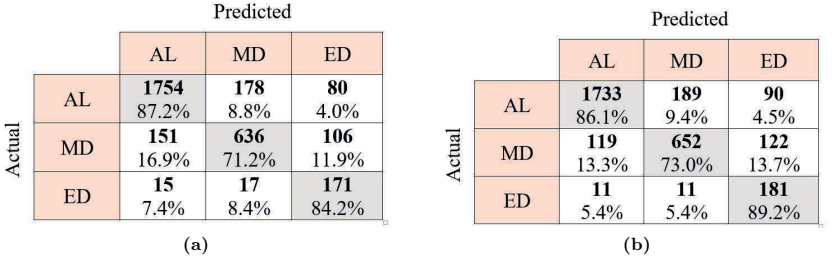


Figure 3.20: Confusion matrices of random forest (a) and KNN (b) classifiers in the **manual** tests when only selected features by **NCA** method are used.

As these figures show, the classification accuracy for classification of moderately drowsy is lower than the accuracy of the two other classes, regardless of the used feature set and classifier. This means that the driver drowsiness's binary classification might help increase the drowsiness detection accuracy; however, it will not detect the transition from alertness to the extreme level of drowsiness. The NCA-selected features achieve the highest classification accuracy for all three classes. The highest classification accuracy for alert, moderately drowsy, and extremely drowsy are respectively 87.2%, 73.0%, and 89.2%.

Figure 3.21 shows the balanced accuracies of driver drowsiness classification systems designed using extracted features in the manual tests. This Figure shows that the best-balanced accuracy occurs when only selected features from the NCA method are used as input features (80.80% by using random forest and 82.77% using KNN). Therefore, NCA removes redundant or irrelevant features and improves the performance of the system. If only one source of the data is used, facial-based features outperform the ECG-based and vehicle-based feature sets. Suppose all of the extracted features without feature selection are utilized, and the random forest classifier is applied to the features. In that case, balanced accuracy is higher than when only the facial-based feature set is exploited.

The accuracy of the KNN classifier trained by only facial-based features is higher than the accuracy of the KNN classifier trained by all of the extracted features. However, the trained random forest using all features outperforms the trained random forest by only facial-based data. These comparisons also show that the random forest classifier is more robust that KNN classifier for dealing with noisy or irrelevant features. Figure 3.22 shows the distribution of the NCA feature weights calculated based on the extracted features from manual driving tests. The weight threshold of the NCA is chosen as 0.2 empirically, and only features with higher weights

are utilized for the classification.

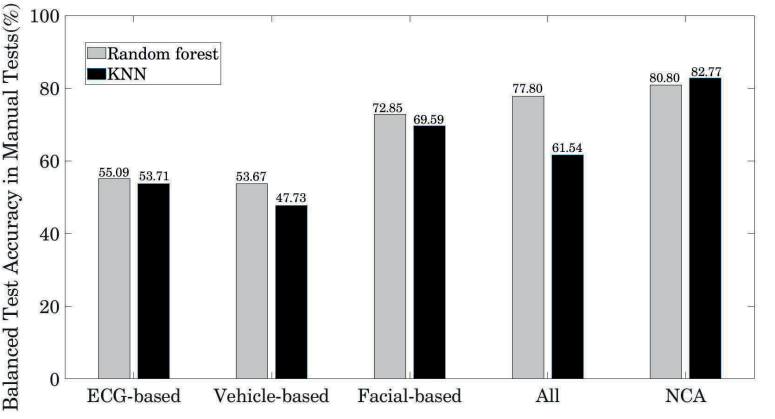


Figure 3.21: Balanced accuracy of driver drowsiness detection systems in **manual** tests when different features sets are utilized.

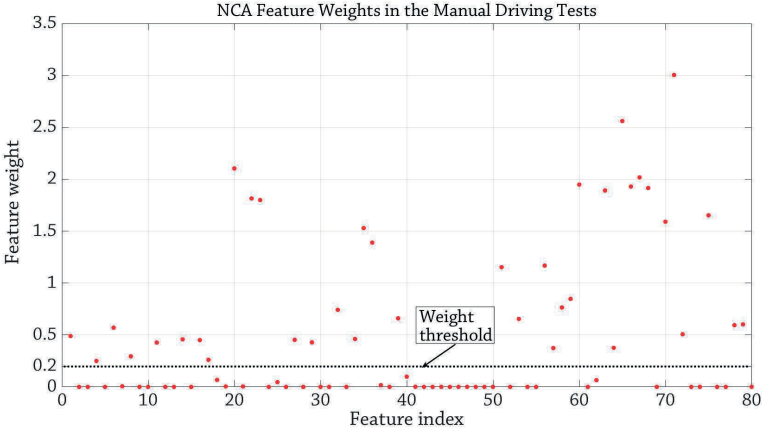


Figure 3.22: NCA feature weights in the **manual** driving tests. A weight threshold for the NCA is set at 0.2 and only features with higher weights are selected to be used in the classification.

For the sake of brevity, classification metrics including specificity, sensitivity, precision, and F1-score [162] are calculated only for the best classifier (KNN trained by NCA-selected features) and presented in Table 3.7. As this Table presents, the precision value for the extremely drowsy level is low. This has occurred because the number of TP is low for the extremely drowsy class. However, the sensitivity and specificity metrics are 0.89 and 0.93, respectively. This means that the low precision for the extremely drowsy class is caused by the low number of samples in this class.

Three features that have the highest features weight in the manual tests are (1) PERCLOS, from the facial-based feature set, (2) LFrel, from the ECG-based feature set, and (3) Blinking

frequency, from the facial-based feature set. A box plot of these three features for three classes of driver drowsiness is presented in Figure 3.23. The PERCLOS is higher in the extremely drowsy class than in the two other classes. The median of PERCLOS in the extremely drowsy class is about 9%, and it is smaller than 2% in the alert class. The same relationship is valid for the blinking frequency feature, where the median of this feature in the extremely drowsy class and alert class is about 20 and 6 blinks per minute, respectively.

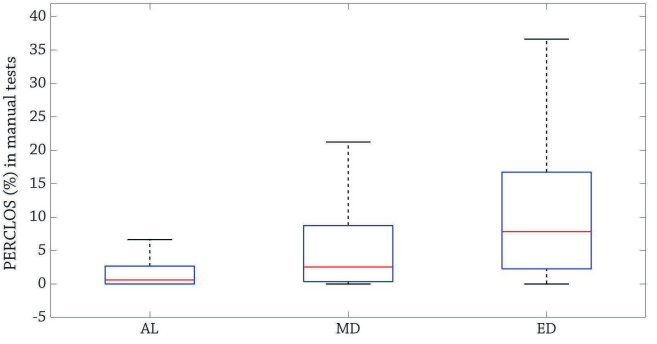
Table 3.7: Classification metrics for the best classifier (KNN trained by NCA-selected features) in the **manual** driving tests. Spe.: specificity; Sen.: sensitivity; Pre.: precision; F1S: F1-score.

-	Spe.	Sen.	Pre.	F1S.
AL	0.88	0.86	0.95	0.90
MD	0.91	0.73	0.76	0.74
ED	0.93	0.89	0.46	0.61

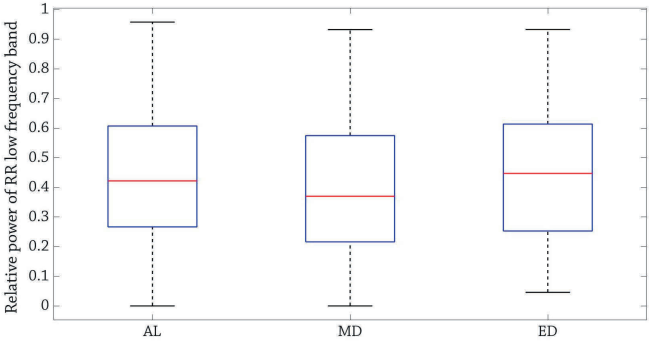
Table 3.8 presents the selected features by the NCA method in the manual driving tests. As this Table shows, 38 features out of 82 features are selected by NCA that 27 of them belong to the vehicle-based, 6 of them belong to the facial-based, and 5 belong to the ECG-based features. This Table shows that all of the facial-based features and approximate entropies of vehicle-based data except steering wheel angle are selected by NCA. Five selected features from ECG signals are also MeanRR, SDRR, pRR50, LF, and HF.

Table 3.8: Selected features by using NCA method in the **manual** driving tests.

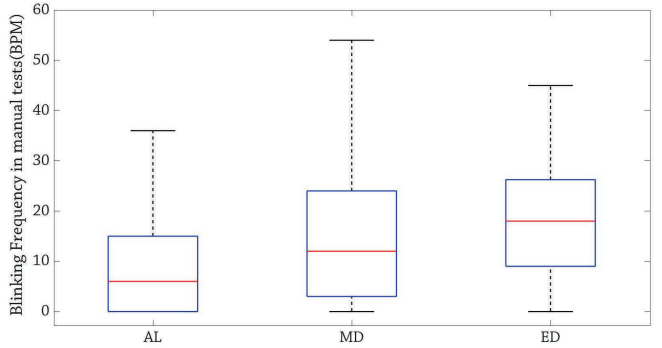
Signal	Selected features by NCA in the manual tests	No. Selected features
Steering wheel angle	First quartile, third quartile, standard deviation, zero crossing rate and NMRHOLD	5
Speed	Skewness, energy, first quartile, second quartile, third quartile, standard deviation and approximate entropy	7
Yaw rate	Skewness, first quatile, second quartile, third quartile, standard deviation, zero crossing rate, approximate entropy	7
Lateral acceleration	Approximate entropy	1
Lane deviation	Skewness, first quartile, third quartile, standard deviation, zero crossing rate, spectral entropy and approximate entropy	7
Eyelid opening	PERCLOS, blinking frequency, mean of eyelid opening, mean of pupil diameter, standard deviation of eyelid opening and standard deviation of pupil diameter	6
ECG	MeanRR, SDRR, pRR50, LF and HF	5
-	-	Sum = 38



(a)



(b)



(c)

Figure 3.23: Box plot of three features that have the highest weights calculated by NCA in the **manual** tests: (a) PERCLOS, (b) LFrel and (c) Blinking frequency; AL= Alert; MD = Moderately drowsy; ED = Extremely drowsy.

3.6.2 Results of Drowsiness Classification in the Automated Driving Tests

Confusion matrices of the KNN and random forest classifiers applied to the (1) ECG-based features, (2) facial-based features, (3) all of the extracted features (no feature selection), and (4) only selected features by NCA method in the automated tests are provided in Figure 3.24 to Figure 3.27, respectively. As these Figures presents, the maximum accuracy of the alert class is provided by the random forest classifier trained by NCA-selected features (84.9%). The maximum classification accuracy of the moderately drowsy class is also provided by the same classifier and same feature set (69.8%). However, the maximum classification of the extremely drowsy class is provided by using both ECG-based and facial-based features together and without applying feature selection (80.7%).

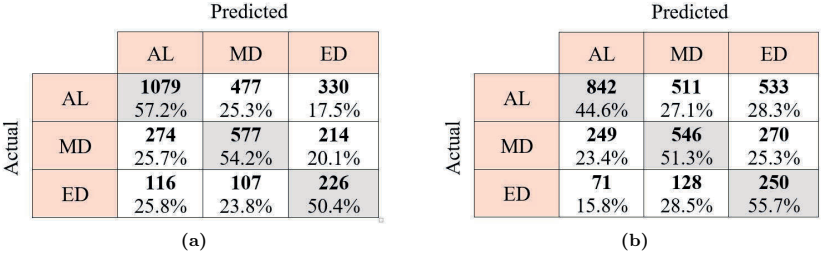


Figure 3.24: Confusion matrices of random forest (a) and KNN (b) classifiers in the **automated** tests when only ECG-based features are used.

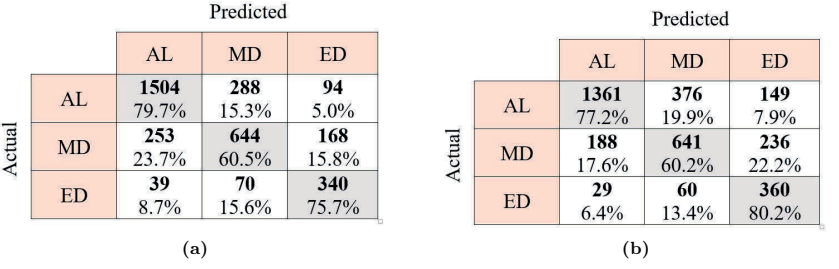


Figure 3.25: Confusion matrices of random forest (a) and KNN (b) classifiers in the **automated** tests when only facial-based features are used.

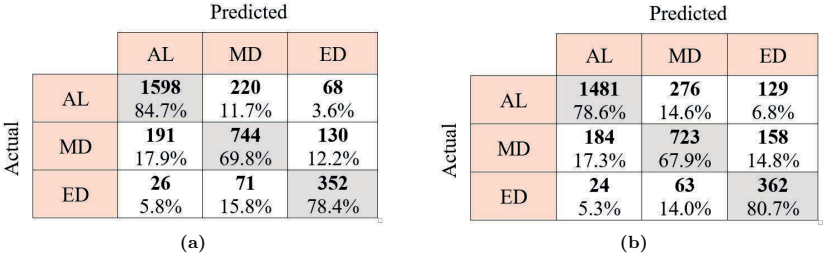


Figure 3.26: Confusion matrices of random forest (a) and KNN (b) classifiers in the **automated** tests when all of the extracted features (no feature selection) are used.

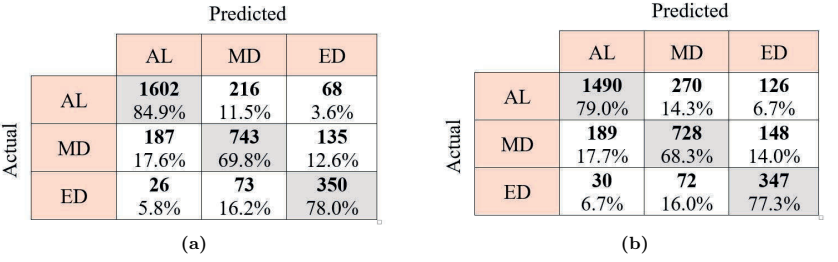


Figure 3.27: Confusion matrices of random forest (a) and KNN (b) classifiers in the **automated** tests when only selected features by NCA method are used.

Balanced accuracies of the designed systems to classify the drowsiness in the automated driving tests have been presented in Figure 3.28. Since drivers insert no input to the vehicle during these tests, vehicle-based data cannot be used to detect the driver drowsiness. As Figure 3.28 shows, the maximum balanced accuracy is obtained by random forest (77.66%) when both of the ECG-based and facial-based feature sets(All) are used. Trained classifiers by only selected features by NCA also return approximately similar accuracies by using fewer features. Comparing the balanced accuracies in the automated and manual driving tests shows the influence of vehicle-based data on drowsiness classification; The best-achieved classification accuracy for the manual tests is 82.77% (see Figure 3.21) while the best accuracy is 77.66% for the automated tests (5.11% lower).

The same feature weight threshold value (0.2) has been used for selecting the features in the automated tests. Selected features by NCA in the automated tests have been presented in Table 3.9. Based on this Table, 12 out of 16 features are selected equally from both feature sets of ECG-based features and facial-based features. It is worth mentioning that the extracted facial-based features are also selected for the automated test as they have already been selected in the manual tests. Therefore, facial-based data are informative to classify driver drowsiness regardless of the driving mode. Selected ECG-based features are MeanRR, SDRR, SDSD, RMSSD, pRR50, and HF. Comparison between the NCA-selected ECG-based features in the manual and automated driving tests shows that the MeanRR, SDRR, pRR50, and HF are selected in both driving modes. This means that the time domain features extracted from ECG signals together with the parasympathetic activity of the HRV (HF) are informative for driver drowsiness classification regardless of the driving mode. The distribution of the feature weights computed using NCA is shown in Figure 3.29.

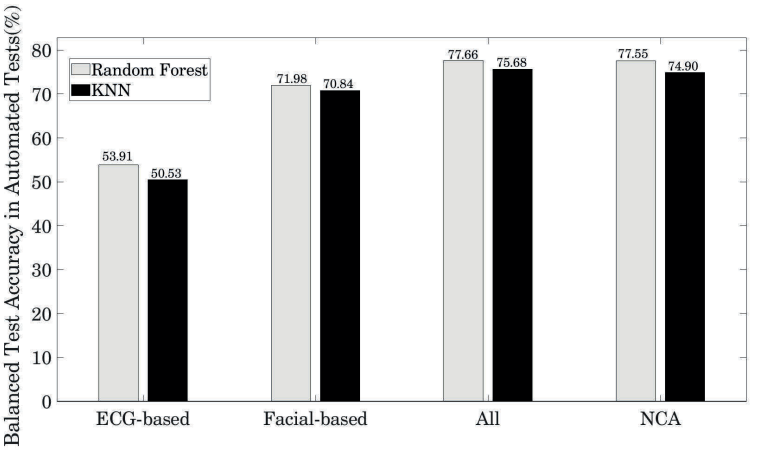


Figure 3.28: Balanced accuracy of driver drowsiness detection systems in **automated** tests when different features sets are utilized.

Table 3.9: Selected features by using NCA method in the **automated** driving tests.

Signal	**Selected features by NCA in the automated tests**	**No. Selected features**
Eyelid opening	PERCLOS, blinking frequency, mean of eyelid opening, mean of pupil diameter, standard deviation of eyelid opening and standard deviation of pupil diameter	6
ECG	MeanRR, SDRR, SDSD, RMSSD, pRR50, and HF	6
-	-	Sum = 12

Four different classification metrics, including specificity, sensitivity, precision, and F1-score, are provided for the random forest trained by all extracted features from automated driving tests. These metrics are provided by Table 3.10.

Table 3.10: Classification metrics for the best classifier (random forest trained by all extracted features) in the **automated** driving tests. Spe.: specificity; Sen.: sensitivity; Pre.: precision; F1S: F1-score.

-	**Spe.**	**Sen.**	**Pre.**	**F1S.**
AL	0.86	0.85	0.88	0.86
MD	0.87	0.70	0.72	0.71
ED	0.93	0.78	0.64	0.70

3.7 Discussion

This chapter presents a workflow to use traditional machine learning to classify driver drowsiness levels by data fusion of three different data sources, including vehicle-based data, ECG-based

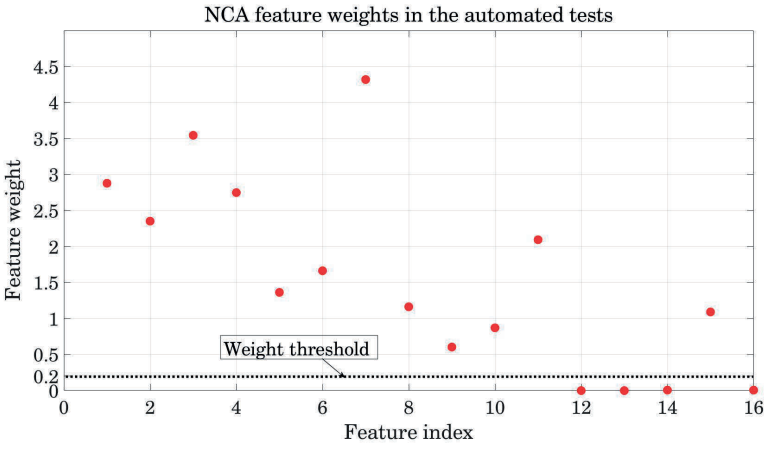


Figure 3.29: NCA feature weights in the **automated** driving tests. Weight threshold on the NCA is set on 0.2 and only features that their corresponding weights are higher than this threshold are selected to be used for classification.

data, and facial-based data. Overall, 82 features are extracted from these data sources, and NCA is employed as a feature selector to distinguish the only informative features and remove redundant or noisy features. The K-nearest neighbors and random forest are utilized as two frequently and simple classifiers to classify the level of drive drowsiness into three classes: alert, moderately drowsy, and extremely drowsy. Results show that, in the manual driving, 38 out of 80 features are informative and relevant to the drowsiness levels are trained K-nearest Neighbors using only selected feature obtains the maximum balanced accuracy, about 83%. In the automated driving mode, only ECG-based and facial-based features (in overall 16 features) are used as inputs to the classifier. Feature selector selects 12 out of 16 features, and classification results show that classification using all features (16 features) achieves approximately similar accuracy when only selected features (12 features) are utilized. However, the feature selector reduces the computation complexity of the classification problems.

The main advantages of the proposed methods in this chapter are as follows:

1. Evaluation of the performance of the three various data sources including vehicle-based, facial-based and ECG-based data and their data fusion through a feature selection method and classification of driver drowsiness into three different levels.

2. Classification of driver drowsiness in both modes of manual and automated driving. Previous research mainly concentrated on the driver drowsiness detection only in manual driving tests.

3. The proposed workflow provides informative features from every data source regarding driver drowsiness classification. This helps to design a system with optimal computation power. Results also show that facial-based data are informative for both manual and automated driving tests. Data fusion of different feature sets outperforms the accuracy of the classifiers trained on every individual feature set.

To compare the results of the proposed methodology, Table 3.11 provides the results of some recent similar works. The number of the classified levels of drowsiness in every study is also provided in this Table. As this Table presents, the highest classification accuracy is provided by

[163] (94.6%). Samiee et al. [163] employed the data fusion of steering angle, lateral deviation, and eye blinking using three ANN classifiers, each trained using one of the inputs. The number of drowsiness classes was 2 (alert or drowsy) in that study. Between studies that classified drowsiness into three classes Li et al. (b) [164] obtains the highest accuracy (88.0%). In that study, approximate entropy features of steering wheel angle and yaw angle were used as inputs to an ANN classifier with two hidden layers (six neurons in each hidden layer). The proposed method by this chapter (KKN trained by NCA-selected features) provides less accuracy (82.8%); however, its computational complexity is lower since KNN has a more straightforward structure than neural networks.

In order to improve the achieved results of this chapter, the following ideas are proposed:

1. Lateral velocity of the vehicle can be estimated using vehicle dynamics model and already measured vehicle-based data. This velocity can help to increase the results of drowsiness classification in the manual mode.

2. Extracted features from other biosignals such as respiration and EEG channels can also be employed for drowsiness classification. However, EEG channels need to be thoroughly preprocessed; otherwise, muscle noise and movement artifacts might decrease the classification performance.

3. Other classification methodologies such as multilayer perceptron neural network and fuzzy logic-based classification can be utilized. In the fuzzy logic-based method, fuzzy rules help to investigate the relationships between extracted features and drowsiness levels.

4. In this chapter, RUSBoost is used to mitigate the effect of imbalanced data issues. Oversampling or undersampling methods such as adaptive synthetic (ADASYN) sampling [152] and Synthetic Minority Oversampling Technique (SMOTE) [166] can also be employed to resample the feature set by increasing the number of minority class' samples, decreasing the majority class' samples or a combination of them.

Table 3.11: A comparison between proposed method in this chapter and some recent studies. The classification results of the manual tests are used for this comparison. DL is the drowsiness levels in classification and Acc. is the accuracy.

Study	**Method**	**Inputs**	**DL**	**Acc.**
McDonald et al. [27]	Random forest classifier	Steering angle	2	79.0%
Samiee et al. [163]	Weighted output of three ANNs, each trained on one input	Steering angle, lateral deviation, and eye blinking	2	94.6%
Wang and Xu [22]	Multilevel ordered logit (MOL) modeling using driver behaviour, and eye features metrics	Steering angle, lateral displacement, speed, eye blinking, and pupil diameter	3	68.4%
Li et al. (a) [12]	Warping distance between linearized approximate entropy in sliding windows	Steering wheel angle	2	78.0%
Li et al. (b) [164]	ANN classifier trained by approximate entropy features of inputs	Steering wheel angle and yaw angle	3	88.0%
Barua et al. [109]	SVM classifier with 10-fold cross validation trained by extracted features from input data	EEG, EOG and contextual information such as time of day	3	79.0%
Vicente et al. [39]	LDA classifier trained by frequency domain features of inputs	HRV and respiration signals	2	78.5%
Awais et al. [1]	SVM classifier trained by extracted features from inputs	O2 EEG channel and HRV	2	80.9%
Moujahid et al. [165]	SVM classifier trained by extracted features using Pyramid-Multi Level (PML) face representation	Driver's face images	2	80.0%
Present study	KNN trained by NCA-selected features	HRV, eyelid, pupil diameter, and five vehicle-based data	3	82.8%

Chapter 4

Driver Drowsiness Classification using Deep Neural Networks

4.1 Introduction

Deep networks refer to the networks that utilize multiple layers to extract features and classify the input data. In these networks, each layer processes the output of its previous layer [167, 168]. In recent years, by improving the computational power of hardware, availability of big data sets and newly developed training algorithms, the application of deep neural networks is increasing. In deep networks, feature extraction, feature selection, and model construction are integrated into one module via end-to-end optimization methods [169]. These networks can outperform traditional machine learning methods, however, they are prone to get overfitted or stuck in the local minima during the training process. This issue increases the generalization error which is measured by using the test dataset. To avoid overfitting, these networks need a huge amount of data to get trained. Moreover, their hyperparameters should also be optimized to find the specific combination of them that provides the best performance among others.

Deep neural network techniques in image processing were applied to design a real-time driver drowsiness detection system using facial-based information [170] and real-time categorization of driver's gaze zone [171]. Deep Convolutional Neural Networks (CNNs) were exploited for automatic sleep stage scoring with single-channel EEG [172]. In [173] the input EEG signals were transferred into time-frequency domain images by applying short-term-Fourier-transform (STFT) then decomposed into five EEG sub-bands. These images were used as inputs to deep neural networks and the extracted features were inserted into a Softmax classifier [174] to detect drowsiness.

This chapter focuses on two data sources, eyelid movements and ECG signals, to improve driver drowsiness classification accuracy by applying deep neural networks. To achieve this goal, two CNNs are employed to be trained on the wavelet scalogram images of ECG and eyelid movement data separately and an ensemble of deep networks is proposed for manual and automated driving modes to enhance every individual network. The next section briefly describes the structure of the CNN and used layers in its architecture.

4.2 Convolutional Neural Network (CNN)

Convolutional neural networks have been frequently used as a method to learn features from raw input data in various applications including medical image classification [175], human activity recognition [176] and automatic speech recognition [177]. These networks create models that are invariant to the transformation of the input data. In the CNN, convolution layers and pooling operations are stacked to build a feature learning block. Convolution layers employ convolution

filters to extract features from input data (such as images or time series). The subsequent pooling layers extract the significant features by using a fixed-length sliding window (pooling size) over the output of convolution layers by pooling operations. Simple features, such as lines, edges, and corners of the input images are extracted in the first (initial) convolution layers, while the more complicated (abstract) features are extracted using deeper layers.

In order to present the processes in a convolution layer, a simple example is provided. Assume that the input is a 5×5 image (5 pixels in width and 5 pixels in height) and the convolution filter is 3×3 square as shown in Figure 4.1. The elements of the convolution filter are trainable parameters that will be updated during the training. The input image is zero-padded by adding zero columns (rows) to the left and right (top and bottom) of the input image. This makes the convolution output that is called activation map have the same dimension as the input image. The operation of convolution between the filter and input image is presented in Figure 4.2. In order to this operation, convolution filter starts to move from the top-left pixel and move to the right with a predefined horizontal stride (here is 1) till it parses the entire width. Then, it moves down with a predefined vertical stride (hers is 1). This process repeats till the whole input image is covered. The element-wise multiplication of filter and its covered pixels in the input image is calculated and all the outputs are added together to compute elements of the activation map. For example, the three first operations for the given input image and convolution filter of Figure 4.1 are presented in Figure 4.2.

0	0	0	0	0	0	0
0	60	113	56	139	85	0
0	73	121	54	84	128	0
0	131	99	70	129	127	0
0	80	57	115	69	134	0
0	104	126	123	95	130	0
0	0	0	0	0	0	0

0	-1	0
-1	5	-1
0	-1	0

Figure 4.1: Input image padded by zero columns and rows (left) and used convolution filter (right)

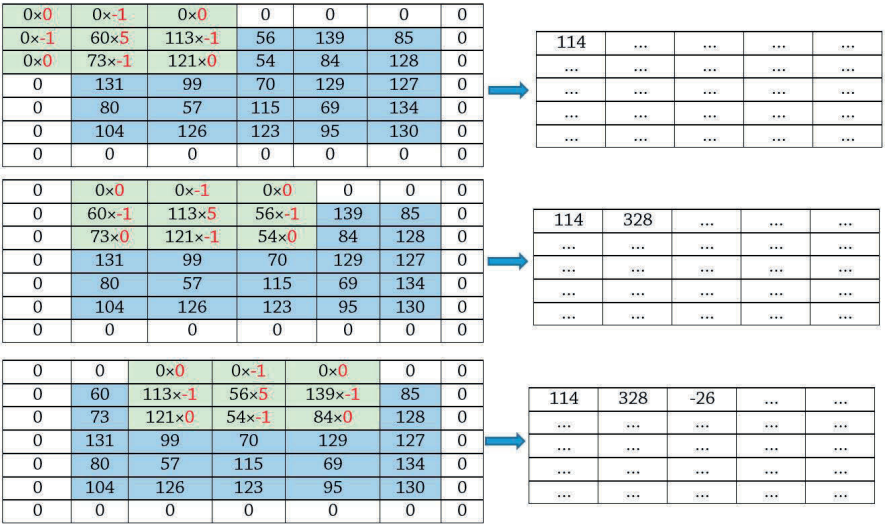


Figure 4.2: Convolution operation in the convolution layers. The input image is zero-padded to make the activation map have the same size as input image.

The next step after calculation of the activation map is inserting it to an activation function. Different activation functions have been used for convolution neural networks such as sigmoid, tangent hyperbolic, and ReLU. A detailed list of these functions has been provided in [178]. In this chapter, the ReLU activation function is used. Equation (4.1) presents this function [178].

$$ReLU(x) = \max(0, x) \tag{4.1}$$

where x is the every element of the calculated activation map. As this Equation shows, ReLU returns the 0 if the input is negative, otherwise, it returns the same input. Figure 4.3 shows the completed activation map of the Figure 4.2 and the output of the ReLU function applied to this activation map.

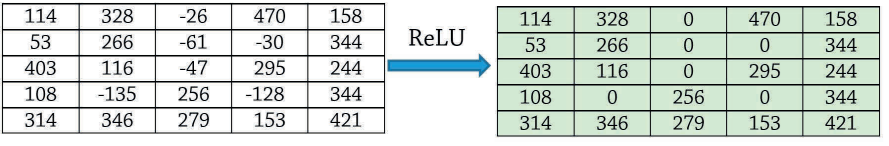


Figure 4.3: Operation of ReLU activation function on the input activation map.

After applying the ReLU activation function, pooling operation is applied to sub-sample its output and select the most important information. Frequently used pooling operations consist of max-pooling and average-pooling. Max-pooling selects the maximum value of the pooling size over the output of the ReLU function as the most essential features, while average-pooling provides the average value of the pooling size as the output value [179]. In this study, the max-pooling is employed as the pooling method. The size of the pooling window is considered to be 2×2 that has stride of 2 for both of the horizontal and vertical direction. Figure 4.4 shows the output of the max-pooling operation.

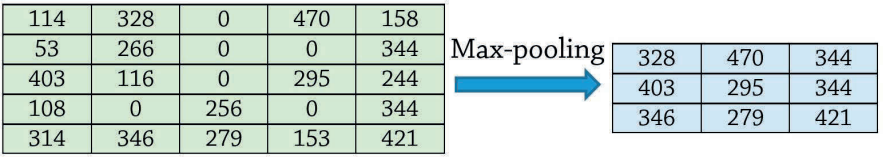


Figure 4.4: Max-pooling operation on the output of ReLU activation function.

During the training process of the deep neural networks, one batch of the input data is used in every iteration to update the parameters of the network. Thus, every activation filter is applied to the every input sample in the batch. For example, if the batch size is considered to be 32, every activation filter generates 32 activation maps as output in every iteration. The distribution of the inputs of every layer changes during the training process since the network parameters (weights and biases) are being updated. This issue is called internal covariate shift and can slow down the training process [180]. Batch Normalization (BN) layers are employed to tackle this issue and reduce the number of necessary iterations to obtain an acceptable performance. Moreover, these layers help to make the deep network more robust to different weight initialization methods and train with larger learning rates [180, 181]. The equations of BN layers are provided in Equation (4.2) to Equation (4.5).

$$\mu_B = \frac{1}{M}\sum_{i=1}^{M}(x_i), \tag{4.2}$$

$$\sigma_B^2 = \frac{1}{M}\sum_{i=1}^{M}(x_i - \mu_B)^2, \tag{4.3}$$

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}, \tag{4.4}$$

$$BN(x_i) = \gamma\hat{x}_i + \beta, \tag{4.5}$$

where M is the batch size (number of activation maps outputted from every convolution filter), x_i is the i-th activation map, μ_B and σ_B^2 are respectively mean and variance of the i-th activation map, ϵ is a small positive constant (usually $1 \times 10^{-}5$) to avoid numerical instability. Finally, γ and β are trainable parameters of the BN layer to shift and scale the standardized input, respectively. Thus, every convolution filter adds two more trainable parameters to the network by using the batch normalization layer.

Convolutional neural networks are mostly employed in image processing applications. In the present application, to generate images from times series signals measured in the driving tests, the continuous wavelet scalogram [182] is utilized. By applying this method, the one-dimensional signals are transformed into the two-dimensional images in the time-frequency domain and can be used as inputs to the CNNs for driver drowsiness classification. The next section explains the process of calculating the wavelet scalogram and training of the deep CNNs.

4.3 Driver Drowsiness Classification using CNN Trained by Wavelet Scalograms of Input Data

4.3.1 Wavelet Scalogram

Wavelet analysis calculates the correlation (similarity) between the input signal and a given wavelet function $\psi(t)$. Unlike Fourier transform, wavelet analysis provides a multi-resolution time-frequency output under this assumption that low frequencies last for the whole duration in the input signal and high frequencies appear in different time points as short events.

A function $\psi(t)$ is considered as a wavelet if the two following conditions are satisfied: (1) The energy of this function should be finite [183];

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 \, dt < \infty, \tag{4.6}$$

(2) If the $\hat{\psi}(f)$ is the Fourier transform of the function $\psi(t)$, the following condition must be satisfied (known as admissibility condition) [183];

$$C_{\psi} = \int_{0}^{\infty} \frac{|\hat{\psi}(f)|^2}{f} \, df < \infty. \tag{4.7}$$

Therefore, the wavelet should have no zero frequency component ($\hat{\psi}(0) = 0$) and this implies that the mean of the wavelet $\psi(t)$ must be zero.

The wavelet function can be scaled and translated by using two real-valued parameters of $s > 0$ and u, respectively and generate a wavelet filter-bank of $\psi_{u,s}$ [182].

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi \left(\frac{t-u}{s} \right). \tag{4.8}$$

By using the transformed wavelet, Continuous Wavelet Transform (CWT) of input signal $x(t)$ at time u and scale s can be calculated as

$$X_{WT}(u,s) = \int_{-\infty}^{\infty} x(t) \psi^*_{u,s}(t) \, dt, \tag{4.9}$$

where $x(t)$ is the every sliding window of the eyelid and ECG data, $\psi^*(t)$ is the complex conjugate of $\psi(t)$ and $X_{WT}(u,s)$ provides the the frequency contents of $x(t)$ corresponding to the time u and the scale s [183]. By using the two parameters of u and s, it is possible to investigate the input signal $x(t)$ in two domains of time and frequency simultaneously where resolution of time and frequency depends on the value of the scale parameter s. Therefore, CWT provides the time-frequency decomposition of $x(t)$ in the time-frequency plane. This method can be more useful than other methods such as Short-Time-Fourier-Transform (STFT) when investigating the non-stationary signals since it provides a higher time resolution in the higher frequencies (lower scales s) while the time and frequency resolution is constant in STFT. The scalogram of $x(t)$ in any positive scale is calculated as the norm of $X_{WT}(u,s)$ [184]

$$\mathcal{S}(s) = ||X_{WT}(u,s)|| = \left(\int_{-\infty}^{\infty} |X_{WT}(u,s)|^2 \, du \right)^{\frac{1}{2}}. \tag{4.10}$$

This Equation calculates the energy of X_{WT} at a scale s. Therefore, using scalogram the most representative and significant scales (frequencies) in the signal can be found.

The wavelet scalogram is used here to transform the time-series ECG and eyelid opening signals to their time-frequency domains. The Morse wavelet [185] is employed to compute the wavelet transform for the input signals and is formulated in the frequency domain. The Fourier transform of this wavelet is [185]

$$\Psi_{\beta,\gamma}(\omega) = U(\omega) a_{\beta,\gamma} \omega^{\beta} e^{-\omega^{\gamma}}, \tag{4.11}$$

where $U(\omega)$ is the unit step, β is used as the decay of compactness parameter and γ presents the symmetry of the Morse wavelet around its centre point. The parameter $a_{\beta,\gamma}$ is also a normalizing constant. More details about the parameters of Morse wavelet and their effects on the wavelet shape is explained in [185]. The next subsection explains the application of wavelet transformation to generate the time-frequency images of ECG and eyelid signals.

4.3.2 Scalogram Calculation of Eyelid and ECG signals

Electrocardiogram and eyelid signals are segmented into sliding time windows with a length of 10 sec and an overlap of 5 sec between every two adjacent windows. Table 4.1 and Table 4.2 provide the number of data samples that belong to every level of driver drowsiness in manual and automated driving conditions, respectively. The eyelid and ECG signals are respectively collected with the sampling frequencies of 100 Hz and 500 Hz and are synchronized using the described method in subsection 3.2.1.

Table 4.1: Number of data samples belonged to each class after applying sliding windows to generate the scalogrms in the **manual** driving mode.

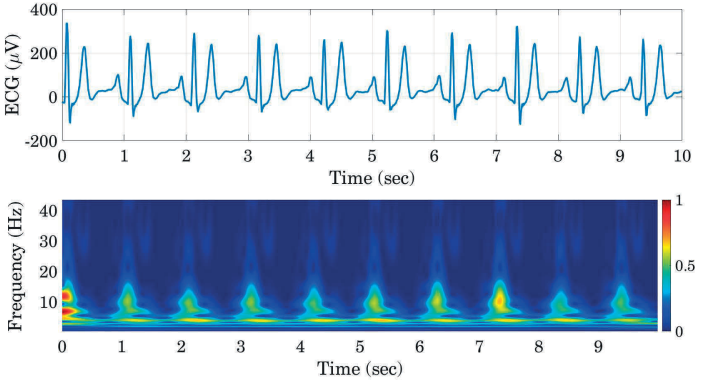
Reference Drowsiness Level	**Number of Samples**	**Percentage**
Alert	23722	67.38%
Moderately drowsy	9371	26.62%
Extremely drowsy	2111	6.00%

Table 4.2: Number of data samples belonged to each class after applying sliding windows to generate the scalogrms in the **automated** driving mode.

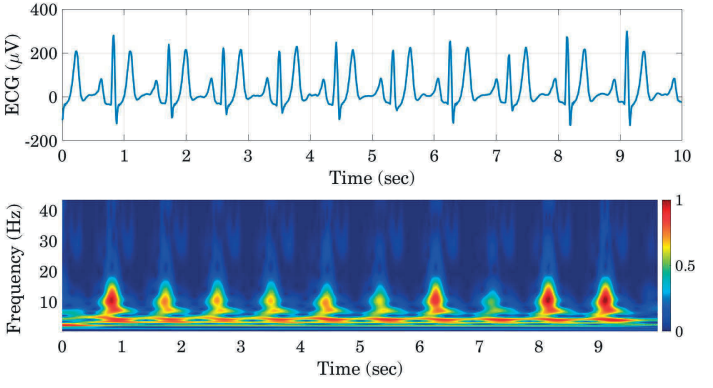
Reference Drowsiness Level	**Number of Samples**	**Percentage**
Alert	19508	56.33%
Moderately drowsy	10699	30.89%
Extremely drowsy	4427	12.78%

To calculate the scalograms images of the ECG and eyelid signals Morse wavelet ($\gamma = 3, \beta = 20$) is employed. Figure 4.5 and Figure 4.6 show examples of ECG and eyelid signals and their corresponding scalogram images for every level of driver drowsiness in a fatigued-automated test, respectively.

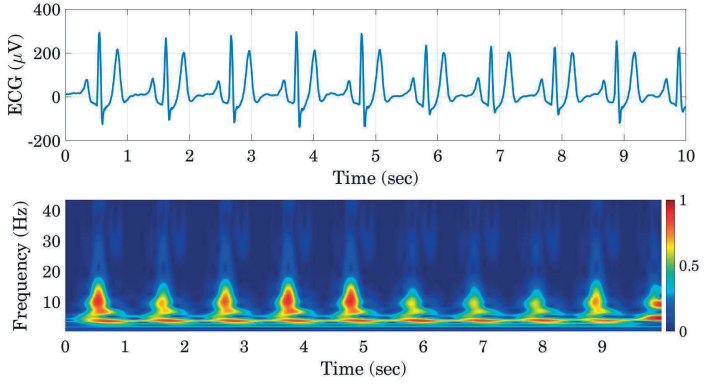
The generated RGB scalogram images are resized to 224×224 pixels and transformed to the grayscale images to reduce the computational complexity of the deep network training. Figure 4.7 shows a sample of the grayscale resized ECG and eyelid scalogram images. A deep CNN is trained for driver drowsiness classification using each of ECG and eyelid scalogram images separately and finally, an ensemble voting strategy has been applied to the output of every deep CNN to outperform every individual network. The next subsection explains the architecture of the used CNN, optimization of its hyperparameters and ensemble voting method.



(a)

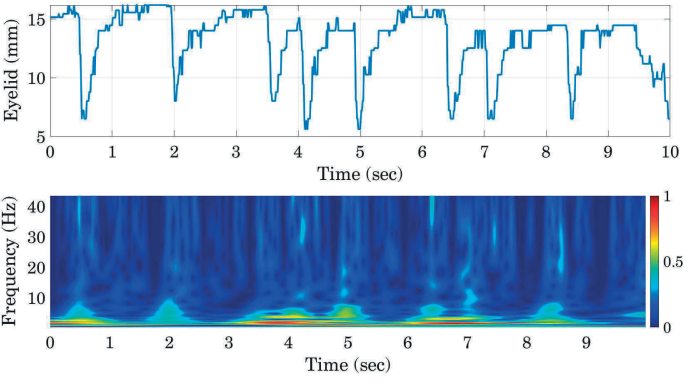


(b)

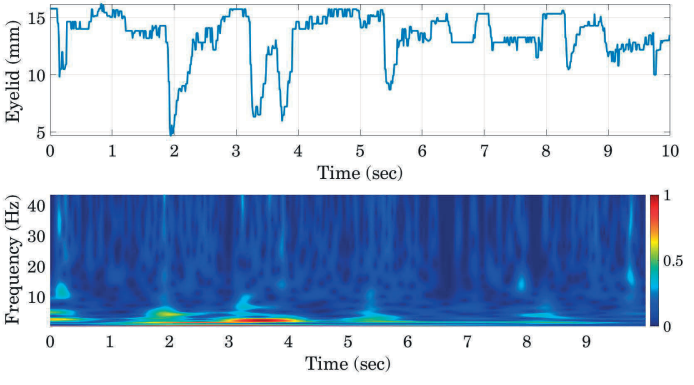


(c)

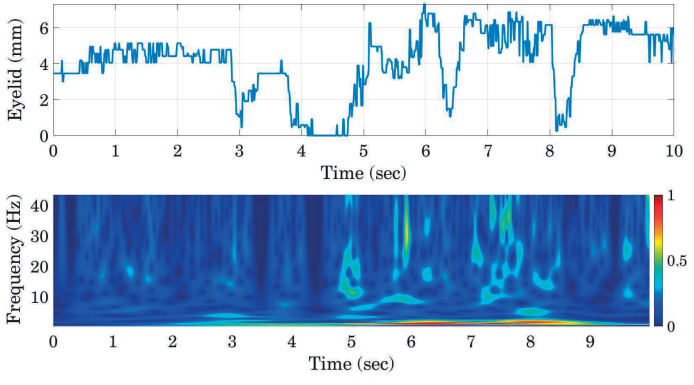
Figure 4.5: Samples of ECG signal and their corresponding scalograms for the alert (a), moderately drowsy (b), and extremely drowsy (c) classes in a fatigued-automated test.



(a)

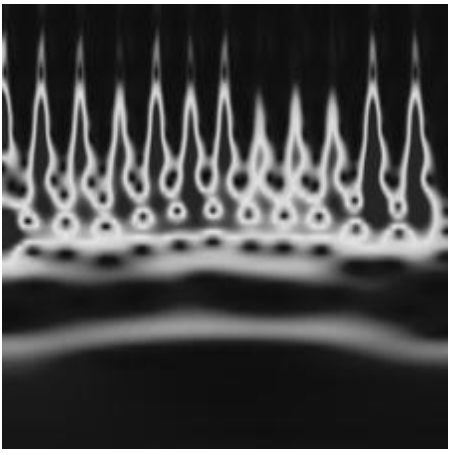


(b)



(c)

Figure 4.6: Samples of eyelid signal and their corresponding scalograms for the Alert (a), Moderately drowsy (b), and Extremely drowsy (c) classes in a fatigued-automated test.



(a)

(b)

Figure 4.7: Samples of grayscale resized images (224×224) of the ECG (a) and eyelid (b) scalogram image.

4.3.3 Architecture of Deep CNNs and Optimization of their Hyperparameters

To prepare the image data set for training the deep CNN, first, the input images are normalized to have zero-mean and unit variance. Then, the entire dataset is split randomly into the training set, validation set, and test set, with 80% of the observations in the training set, 10% in the validation set, and 10% in the test set. The input images are split into these data sets in a way that distribution of the images is the same for three data sets and the percentage of the classes is approximately the same as Table 4.1 for manual tests and Table 4.2 for automated tests.

The utilized deep CNN is composed of five convolutional blocks and one fully connected block in its hidden layer. Convolution and fully connected blocks are presented in Figure 4.8, where Conv, BN, ReLU, Max Pool, and FC are convolution layers, batch normalization layer,

ReLU activation function, max-pooling layer and fully connected layer, respectively. The hidden layer is followed by the output layer that is constructed using an FC layer, soft-max layer, and weighted classification layer (Weight). The number of neurons in the fully connected layer of the output layer is equal to the number of classes (here three). The Weight layer is also employed to mitigate the data imbalance issue. Figure 4.9 presents the architecture of the deep CNN, where five convolution blocks are followed by one fully connected block. Moreover, one Dropout layer is also added after convolution blocks to reduce the possibility of overfitting or getting stuck in the local minima during the training process. The dropout layer temporarily eliminates some neurons with a predefined probability, along with all of their input and output connections [186].

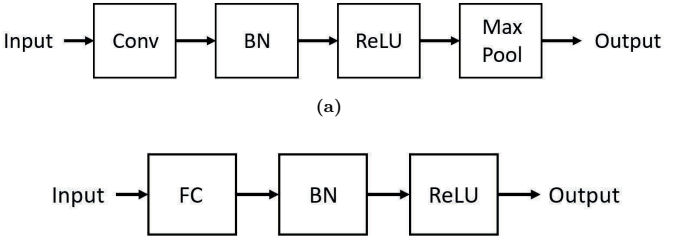


Figure 4.8: Convolution (a) and fully connected (b) blocks that are used to construct the deep CNN.

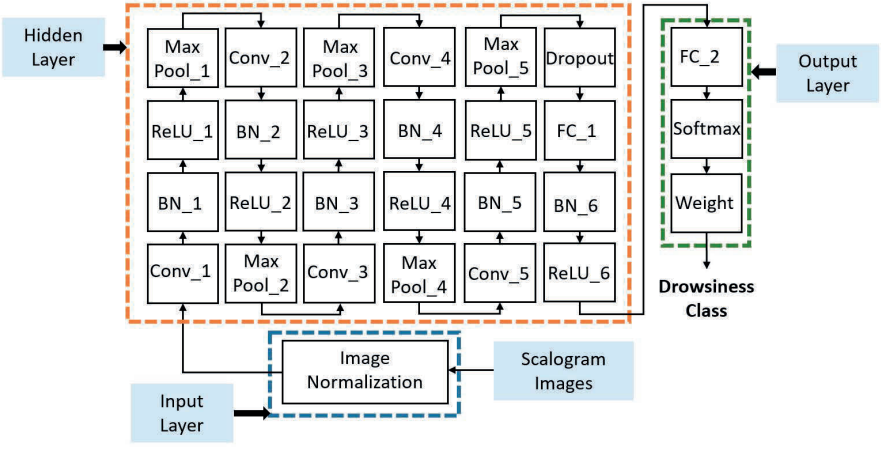


Figure 4.9: The architecture of the deep CNN to classify ECG and eyelid scalogram images into three classes of drowsiness.

The weight of every class is calculated by using Equation (4.12).

$$W_i = \frac{N_c}{C_i \sum_{i=1}^{N_c} \frac{1}{C_i}}, \tag{4.12}$$

where N_c is the number of classes (here three), C_i is the number of data samples that belong to the i-th class, and finally, W_i is the calculated weight for i-th class. By applying the Equation 4.12 to the data samples that belong to the drowsiness classes in the manual and automated

modes (presented in Table 4.1 and Table 4.2), the corresponding weights for every class are computed. Table 4.3 provides these weights. As this Table shows, the class weights of the extremely drowsy class are higher in both manual and automated mode tests. By using these weights, the misclassification error of the drowsiness classes increases in comparison to the alert class. Therefore, if the network classifies a drowsy sample in the alert class wrongly, it results in a large misclassification error that would have a significant influence on the optimization process and reduce the occurrence possibility of this classification error.

Table 4.3: Class weights of the different drowsiness classes used in the deep CNNs to alleviate the imbalanced data set issue. (AL: alert; MD: moderately drowsy; ED: extremely drowsy)

-	**Manual**	**Automated**
AL	0.203	0.415
MD	0.514	0.757
ED	2.283	1.828

Deep neural networks have multiple hyperparameters such as learning rate, regularization, and number of neurons that can influence on the network performance. Finding a proper combination of these hyperparamters is a major task in the field of deep learning [187].

Here, in order to optimize the hyperparameters of the designed network, the Bayesian optimization method [188] is applied. This method has the capability of reasoning about the iterations' performance before they are carried out. Therefore, less number of iterations is needed to provide the optimal hyperparameter combination in comparison to other hyperparameter optimization methods. Moreover, previous works show that this method achieves better performance on the test data set than other methods such as grid search and random search [189]. Four different hyperparameters have been considered to be optimized using Bayesian optimization method including:

- Initial learning rate: This hyperparameter is used in the structure of the optimization method to update the learnable parameters of the network (such as weights in fully connected layers and elements in convolution filters). In order to avoid overfitting, the initial learning rate is usually decreased during performing the last epochs [190].

- L2 regularization: During the training process, weights will become larger to handle the feature extraction from input data and provide better performance. However, large weights can destabilize the network. In this case, only a small variation in the input data will make large differences in the output results while the small variations can be originated from statistical noise in the training data. Consequently, the deep network will be overfitted to the training data if its parameters are too large. The L2 regularization is a method to penalize the large weights by adding a penalization term to the loss function. Here, the cross-entropy loss is employed for performing the drowsiness classification and the L2 regularization term is added to this function [191].

- Dropout probability: The more dropout probability value means that it is more probable to turn off a specific neuron and its all input and output connections. This helps to reduce the network complexity and consequently avoid the overfitting issue.

- Number of filters in convolution layers: This hyperparameter determines the network complexity and its capability to solve the classification problem. A too large number of neurons will lead to a huge computational load and increase the possibility of overfitting. On the other hand, if the number of neurons is too small the network will be stuck in the local minima and it will cause the underfitting issue. Here, it has been assumed that the

number of filters in Conv1 to Conv5 and the number of neurons in FC1 are equal and only one hyperparameter is defined to find their optimal values.

Table 4.4 presents the specified search space for each of these hyperparamters that have been optimized using Bayesian optimization method.

Table 4.4: Defined hyperparameters of Deep CNN to be optimized using Bayesian optimizer.

Index	Hyperparameter	Search space
H_1	Initial learning rate	$[5 \times 10^{-5} - 0.001]$
H_2	Dropout probability	$[0.2 - 0.4]$
H_3	L2 regularization	$[10^{-8} - 10^{-2}]$
H_4	Number of filters in convolution layers (Conv1 to Conv5) and neurons in fully connected layer (FC1) [Integer value]	$[30 - 60]$

ADAptive Moment estimation (ADAM) optimizer [192] is employed to train the parameters of the designed deep CNNs. Maximum number of epochs is empirically considered to be 15 and a schedule for learning rate is utilized that multiples the initial learning rate by 0.1 after 12 epochs to alleviate the overfitting issue in the last training epochs. Literature proposed that small min-batches with the size of 4 to 32 input images provide better performance on the test data set and more stable training process. In [193] was proposed that the mini-batch size of 16 provides better performance than other sizes. Here, the size of the mini-batch is also defined to be constant and equal to 16. The training process is conducted on a system with CPU and GPU types of Intel Core™i7-782HQ and NVIDIA™Quadro M2200, respectively. One deep CNN is trained for each of the ECG and eyelid scalograms in the manual and automated driving modes separately. Finally, an methodology is employed to classify the level of driver drowsiness based on the ensemble of the trained networks to improve the performance of every individual network. The next subsection explains the ensemble method used here.

4.3.4 Ensemble Learning of Deep CNNs

In order to implement ensemble learning, the soft voting method [194] is exploited here. Figure 4.10 shows the pipeline of the ensemble learning. This method predicts the output class by using the calculated prediction probability of every class label obtained by two individual networks. The final output of the ensemble assigned to the class that has the highest posterior probability sum. By using this approach, every CNN that can provides a better discrimination between drowsiness classes will be more significant than another CNN for estimating the final drowsiness class to the input image. Equation 4.13 presents the soft voting ensemble method formula [195].

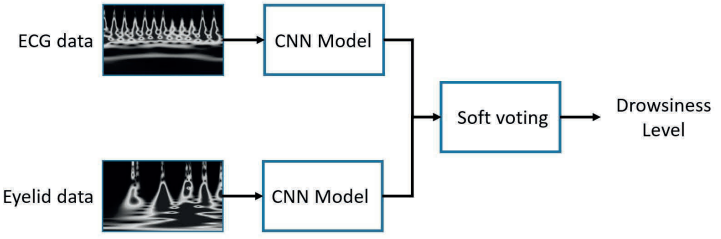


Figure 4.10: Ensemble learning applied on the trained deep CNNs.

$$\hat{C}_i = \underset{k}{\operatorname{argmax}} \sum_{j=1}^{M} p_{i,k}^{j} \tag{4.13}$$

where $p_{i,k}^{j}$, and M presents the posterior probability of the j-th network for predicting the i-th data sample in k-th class calculated using Softmax classifier and number of networks, respectively. For example, assume that the trained CNN by eyelid data and trained CNN by ECG data provides the following probabilities P_{Eyelid} and P_{ECG}, respectively.

$$P_{Eyelid} = (0.3, 0.3, 0.4)\,; P_{ECG} = (0.3, 0.5, 0.2)\,, \tag{4.14}$$

In these vectors the first, second and third elements show the probability that the input image belong to the alert, moderately drowsy, and extremely drowsy, respectively. Based on these probabilities, trained CNNs by eyelid and ECG classify the input image into the extremely drowsy and moderately drowsy classes, respectively. By using the Equation (4.13), the two vectors of Equation (4.14) are added and the result is $(0.6, 0.8, 0.6)$. Therefore, the input image is classified into the moderately drowsy class.

4.4 Results

This section first explains the results of Bayesian hyperparameter optimization of deep CNNs for each driving mode and the used input scalogrm image types. Then, it describes results of the driver drowsiness classification using deep CNNs. The results are provided separately for the two driving modes, manual and automated driving. For each of these modes, the results of the classifiers are presented when either ECG scalograms or eyelid scalogram are utilized as input images, along with their ensemble classification results using soft majority voting.

4.4.1 Results of Bayesian Hyperparameter Optimization of Deep CNNs

As presented in Table 4.4, four hyperparameters are considered to be optimized in the structure of the deep CNNs. Table 4.5 provides the optimized values of the hyperparameter of deep CNNs in different driving modes and by inputting different types of scalogram images (ECG and eyelid) to the deep CNNs. As this Table shows, number of filters in the convolution layers and neurons in the fully connected layer (presents by the hyperparameter H_4) is higher in the automated driving mode for both ECG and eyelid scalogram images. Therefore, computational cost is higher in the automated tests to classify driver drowsiness using the proposed deep CNNs. The L2 regularization value (presents by the hyperparameter H_3) is much higher in the manual tests than in automated tests. Thus, deep CNN needs larger parameters to classify the driver drowsiness in the manual tests. The dropout probability (presented by the hyperparameter H_2) of the trained deep CNN for the ECG signals in the automated tests is higher than other

designed deep CNNs for the manual tests and for the eyelid data in the automated tests. The number of neurons is also higher for the deep CNN trained by the ECG signals for automated driving. Therefore, its network is wider than other networks and the dropout probability of the networks should also be higher to turn off more neurons and avoid overfitting.

Table 4.5: Optimized values of hyperparameters in different driving modes and by inputting different scalogram types to the deep CNNs. Hyperparameters H_1 to H_4 are defined in Table 4.4.

Driving mode	**Input scalogram type**	H_1	H_2	H_3	H_4
Manual	ECG	0.0001	0.202	0.0015	42
Manual	Eyelid	0.0003	0.287	0.008	49
Automated	ECG	0.0002	0.321	2.74×10^{-8}	60
Automated	Eyelid	0.0005	0.251	4.81×10^{-6}	56

4.4.2 Results of Driver Drowsiness Classification in the Manual Driving Tests

In order to provide a performance comparison for each of the trained CNNs and their ensemble in the manual driving tests, the confusion matrices of the test data sets are presented in Figure 4.11. In this Figure, the diagonal elements (in gray) provide the number and percentage of the input images that are correctly classified in different classes of drowsiness, according to the ground truth classification from the video observations. While, non-diagonal cell presents the number of samples that are misclassified. This Figure shows that all of the diagonal cells of ensemble model's confusion matrix are higher than the same values in the two individual deep CNNs. Therefore, the ensemble method provides better accuracies for all drowsiness classes in comparison to the individual models. However, the differences are rather small which indicates that in case of loss of one signal the misclassification error is still small. This an important result since the quality of eyelid data depends on the light conditions and the quality of ECG signals can also be degraded because of the driver's body motions.

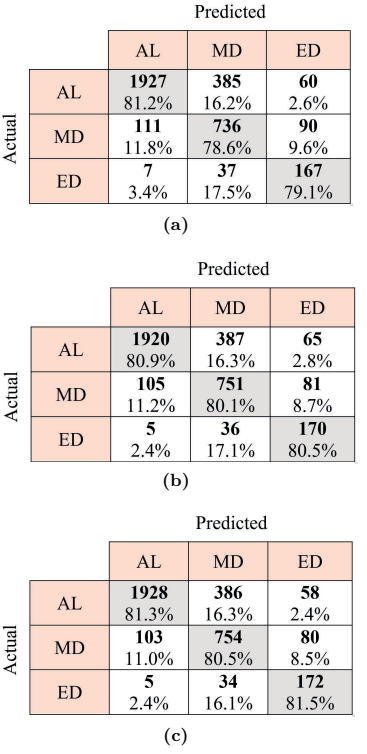


Figure 4.11: Confusion matrices of deep CNNs for driver drowsiness classification in the **manual** driving tests using only ECG scalograms (a), only eyelid scalograms (b), and ensemble of ECG and eyelid scalograms (c).

Sensitivity, specificity, precision, and F1-score of the test data set are presented in Table 4.6 as classification performance metrics. The sensitivity metrics measure the proportion of the actual observations in each class that has been classified correctly. Based on this Table, the sensitivity value for all of the drowsiness classes are the same and equal to 0.81. The specificity explains the proportion of the data samples that does not belong to an specific class and are correctly classified into two other classes. As Table 4.6 presents, the moderately drowsy class and extremely drowsy class provide the lowest and highest values for specificity metric in the manual driving tests, respectively. Therefore, the moderately drowsy class is the most difficult class of drowsiness to be classified correctly using the developed classification method.

The specific accuracy for every individual drowsiness class and their balanced accuracy are presented in Table 4.7. As this Table presents, the ensemble method outperforms the classification accuracy of every individual deep CNN for classification of every drowsiness class and their corresponding balanced accuracy. Moreover, comparison between results of the deep CNN classification method with the results of the traditional machine learning classifiers (presented in Figure 3.21) shows that the proposed deep CNNs significantly outperform the traditional machine learning methods for both facial-based and ECG-based data in the manual driving tests (see section 4.5).

Table 4.6: Classification metrics for the applied ensemble learning on deep CNNs trained by ECG and eyelid scalogram images in the **manual** driving tests. Spe.: specificity; Sen.: sensitivity; Pre.: precision; F1S: F1-score.

-	Spe.	Sen.	Pre.	F1S.
AL	0.91	0.81	0.95	0.87
MD	0.84	0.81	0.64	0.71
ED	0.96	0.81	0.55	0.65

Table 4.7: Accuracy of different classes of drowsiness by using different scalogram images as input to the deep CNN and their ensemble in the **manual** driving tests. AL Acc.: Accuracy of the Alert class; MD Acc.: Accuracy of the Moderately Drowsy class; ED Acc.: Accuracy of the Extremely Drowsy class; Bal. Acc.: Balanced Accuracy.

-	ECG	Eyelid	Ensemble
AL Acc. %	81.2	80.9	**81.3**
MD Acc. %	78.6	80.1	**80.5**
ED Acc. %	79.1	80.5	**81.5**
Bal. Acc. %	79.6	80.5	**81.1**

4.4.3 Results of Driver Drowsiness Classification in the Automated Driving Tests

Confusion matrices of the deep CNNs trained by ECG and eyelid scalogram images generated in the automated driving tests and their ensemble are presented in Figure 4.12. The gray diagonal cells of the matrices present the correctly classified data samples of the test data set. According to this Table, all of these gray cells are higher in the confusion matrix of the ensemble method. Thus, the ensembling method returns a better performance than two deep CNNs. However, their difference is also small for the automated tests.

Table 4.8 provides the classification metrics of the deep ensemble method in the automated driving tests. Based on this Table, the specificity and sensitivity metrics of the moderately drowsy class are lower than the two other classes. Extremely drowsy class also has the highest sensitivity and specificity metrics. Table 4.9 also presents the classification accuracies of the different classes and their balanced accuracy. Based on this Table, the moderately drowsy level has the lowest accuracy therefore, it is the most difficult level to be classified correctly using the deep CNN models.

Comparison between the classification results of the two driving modes also shows that the proposed method performs better in the manual driving mode since the accuracy of the moderately drowsy class in the manual driving tests is about 5.5% higher than in the automated ones (see Table 4.7). This is a significant result since higher classification accuracy of the moderately drowsy class is crucial to warn the driver before the transition to the extreme level of drowsiness or even micro-sleeps. According to this result, detection of moderately drowsy level is harder in the automated driving. This can be caused by two major reasons: (1) lower eyelid quality in the automated tests since the participants insert no input to the vehicle and it can be monotonous for them to only look toward the test track for the whole time of test (30 min). Therefore, they look to the other directions and it will reduce the quality of the collected eyelid, (2) in the automated mode, participants have no obligation or stress to keep the vehicle in the right lane during the test. This makes the participants more relaxed therefore the heart rate stays more stable in the test and it will be harder to distinguish the transition between

Table 4.8: Classification metrics for the applied ensemble learning on deep CNNs trained by ECG and eyelid scalogram images in the **automated** driving tests. Spe.: specificity; Sen.: sensitivity; Pre.: precision; F1S: F1-score.

-	Spe.	Sen.	Pre.	F1S.
AL	0.90	0.82	0.91	0.86
MD	0.86	0.75	0.70	0.72
ED	0.93	0.83	0.65	0.73

alertness and extremely drowsy level.

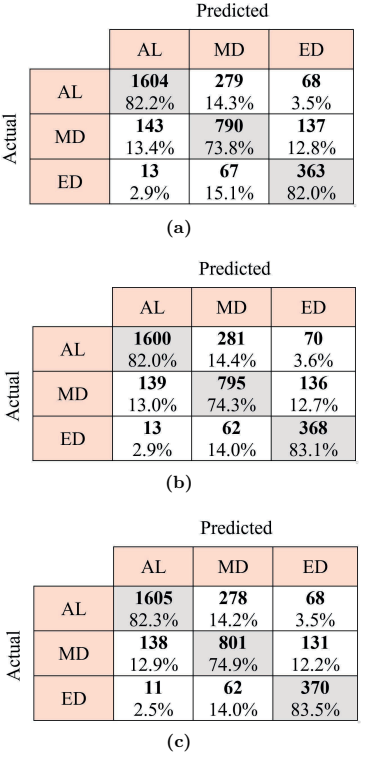


Figure 4.12: Confusion matrices of deep CNNs for driver drowsiness classification in the **automated** driving tests using only ECG scalograms (a), only eyelid scalograms (b), and ensemble of ECG and eyelid scalograms (c).

Table 4.9: Accuracy of different classes of drowsiness by using different scalogram images as input to the deep CNN and their ensemble in the **automated** driving tests. AL Acc.: Accuracy of the Alert class; MD Acc.: Accuracy of the Moderately Drowsy class; ED Acc.: Accuracy of the Extremely Drowsy class; Bal. Acc.: Balanced Accuracy.

-	ECG	Eyelid	Ensemble
AL Acc. %	82.2	82.0	**82.3**
MD Acc. %	73.8	74.3	**74.9**
ED Acc. %	82.0	83.1	**83.5**
Bal. Acc. %	79.3	79.8	**80.2**

4.5 Discussion

This chapter proposed a deep learning strategy to classify driver drowsiness in two driving modes of manual and automated. First, ECG and eyelid data are segmented by using a sliding time window. Then, wavelet scalogram images of every window of the data are generated and images are transformed into grayscale images and resized to 224×224. One deep CNN is trained for every type of scalogram image (EEG and eyelid). The Bayesian optimization method is also utilized to select the optimal hyperparameter set for every deep CNN. Finally, an ensemble method based on a soft voting strategy is employed to outperform every individual trained network.

A comparison between the proposed deep CNNs in this chapter with traditional machine learning methodologies used in Chapter 3 is provided by Table 4.10. As this Table shows, the deep CNNs could significantly outperform the traditional classifiers in all of the drowsiness classes and consequently in the balanced accuracy of classification. For example, the balanced accuracy of the random forest applied to the ECG-based features (RF-ECG) is 55.1% while the same accuracy for the deep CNN applied to the ECG signal is 79.6% (improvement by about 24.5%). Deep CNN considers the frequency contents of the entire ECG sub-waves, whereas the traditional classifiers are trained only based on the heart rate variability data derived from the R-peaks in the ECG signal. This Table also represents that the differences between every individual Deep CNNs and their ensemble are rather small, which indicates that in case of signal loss of eyelid or ECG data, the classification error would be small. This is a significant result since eyelid data depends on face detection performance and light condition. On the other hand, contactless precise ECG measurement requires innovative but also low-cost sensors whose robustness and accuracy are still unknown.

Table 4.10: Comparison between the accuracies of the traditional classifiers (Chapter 3) and deep CNNs applied to the ECG and facial-based data in the manual driving tests. For the KNN and random forest (RF) classifiers extracted features from ECG and facial-based data (eyelid opening and pupil diameter signals) are used as input features.

-	**AL. Acc.**	**MD. Acc.**	**ED. Acc.**	**Bal. Acc.**
RF-ECG	55.6	50.6	59.2	55.1
RF-facial	80.1	61.6	76.9	72.8
KNN-ECG	53.4	48.6	59.2	53.7
KNN-facial	71.4	60.0	77.4	69.6
CNN-ECG	81.2	78.6	79.1	79.6
CNN-facial	80.9	80.1	80.5	80.5
CNN-ensemble	81.3	80.5	81.5	81.1

The following tasks can be investigated to improve the results of the proposed method in this chapter:

1. The transition between drowsiness classes (alert, moderately drowsy, and extremely drowsy) has tricky dynamics when relying on the ground truth derived based on the video observations. These observations depend on the expert knowledge and significance level of drowsiness signs shown by individual drivers. This limitation might be alleviated by using other ground truths methods such as labelling the drowsiness-related sub-waves in the EEG signals.

2. The combination of CNN and Recurrent Neural Network (RNN) such as Long-Short Term Memory (LSTM) and Gated Recurrent Units (GRU) might also improve the performance of the proposed method. Recent studies show that this combination leads to outperform the individual CNN in image classification [196, 197].

3. The proposed methods in this Chapter and Chapter 3 develop generic driver drowsiness classification systems that consider no driver-specific differences. Only two hours of data is available for every driver that might not be sufficient to train a driver-specific deep network. In order to build driver-specific system, the transfer learning methodology [198] can be employed. Using this method, deep CNNs can be trained using a training set that contains the entire data set except the data of the one specific driver. By using this method, the amount of data collected from each driver that is required to build a driver-specific system can be reduced.

Chapter 5

Driver Drowsiness Estimation using EEG Signals with a Dynamical Encoder-Decoder Modeling Framework

5.1 Introduction

Neural activities collected using EEG electrodes are widely exploited to classify and predict the different levels of driver drowsiness. For example, in designing a driver drowsiness detection system, Chaotic features including Higuchi and Petrosian fractal dimensions and the logarithm of energy were extracted from EEG signals in [199] and results showed that the neural network classifier trained by these features obtains the accuracy of about 84% for detection of driver drowsiness. Extracted features from a single-channel EEG signal using wavelet packet transform were used in [200] and results showed that this method obtained an accuracy of about 85% for driver drowsiness detection. Budak et al. [201] also proposed the ensemble majority voting of three deep networks that were trained using different EEG features to classify the vigilance state into two classes: awake and drowsy. On the contrary, a support vector machine-based posterior probabilistic model was proposed in [202] that used the power of Theta, Alpha, and Beta sub-bands of EEG data and transformed the drowsiness level to any value between 0 and 1.

Previous studies mainly concentrated on designing a classification method to construct a decision boundary that discriminates the levels of driver's vigilance e.g., alert or drowsy in binary classification. In this chapter, a new generative framework is proposed to track the driver drowsiness in real-time. The PERcentage of Eyelid CLOSure (PERCLOS) variable is assumed to define the actual level of drowsiness where PERCLOS can represents the driver drowsiness in a real-time manner and it is independent of video observations performed by expert raters. Previous studies have shown that the higher PERCLOS values strongly correlate with drowsiness signs in drivers where drowsiness was also associated with lane deviation in the road [203, 204].

The first goal of this chapter is to design a user-specific interpretable framework to estimate driver drowsiness in real-time based on EEG features. Finding the specific EEG features that are statistically significant and consistent across different drivers is the second goal of this chapter. An encoding-decoding framework is employed to accomplish these goals by estimating the PERCLOS using extracted features from EEG signals. This framework is used since it has been successfully exploited in different applications such as extracting multi-dimensional auditory and visual stimulus-response correlations [205], reconstructing natural images using Bayesian

decoder [206], and decoding hidden cognitive states [207]. Here, the proposed encoding-decoding framework consists of three key steps: 1) a model that characterizes the dynamics of PERCLOS, 2) the encoder model that characterizes how neural features encode the PERCLOS, and 3) the decoder model that estimates the PERCLOS using selected EEG features by the encoder model. In the next sections, the preprocessing of EEG signals and the structure of the encoder-decoder modeling framework is explained.

5.2 Preprocessing of EEG signals

The data of 18 driving sessions from 13 drivers (5 females and 8 males; age 44.5±18.8 years) that represent a wide range of PERCLOS and also acceptable quality for eyelid data are used to estimate the level of drowsiness in drivers. EEG signals are collected using a g.Nautilus® (research version) [115] device with gel electrodes, and the sampling frequency of 500 Hz. Eight electrodes are used to collect EEG signals including Cz, Fz, T7, T8, C3, C4, PO7, and PO8. Two EOG channels are also connected vertically to the lower and upper areas of the right eye. The EOG signal is calculated as the difference between the two EOG electrodes to measure the eye movement activities. This represents redundant information to confirm the SmartEye eyelid opening measurement. Positions of these electrodes in the 10-20 system are marked by red circles in Figure 2.4.

EEG signals are contaminated by various noise sources including eye movement, eye blinking, and high-frequency muscle activities (originated from head movement, yawning, etc.). To eliminate these artifacts, Independent Component Analysis (ICA) technique is employed [208]. ICA decomposes the raw EEG signals into independent components, where the source of each component can be identified using its scalp topography. The denoised EEG signals are reconstructed by eliminating those components that are not originated from brain lobes. More details about the theory of the ICA method is presented in [209].

Here, EEGLAB Matlab toolbox [210] is used to preprocess the EEG raw data. EEG channels are first decomposed to their independent components, then the ICLabel tool [211] implemented in the EEGLAB is used to check which component is originated by brain. ICLabel provides a probability for every component that shows where it is originated from. This probability is calculated based on the spectral analysis and scalp topography map of every component. For example, Figure 5.1 presents the map topography and assigned label to independent components of the EEG raw data in a rested-automated driving test. Based on this Figure, independent components of number 1 to number 3 are generated by eye movement activities and must be removed to preprocess the EEG channels. Figure 5.2 shows a part of the raw and preprocessed Fz channel in the same rested-automated test. As this Figure shows, large EEG amplitudes that are generated by eye movements or blinks are removed in the preprocessed signal while other samples of the signals are barely influenced by the applied ICA.

5.3 PERCLOS Neural Encoder Model

PERCLOS is assumed as a stochastic process characterized by a positive random variable, with a range of 0 to 1. The PERCLOS temporal dynamic over-time is defined by the following state transition process presented in Equation (5.1).

$$x_k = 0.5(1 + \tanh(ax_{k-1} + b + \epsilon_{k-1})), \tag{5.1}$$

where x_k is the PERCLOS value at k-th time interval, $\{a, b\} \in \mathbb{R}$ are free parameters and $\epsilon \in \mathbb{R}$ is a zero-mean Gaussian noise with the variance of σ_ϵ^2; $\epsilon_k \sim \mathcal{N}(0, \sigma_\epsilon^2)$. It is assumed that $Y_k = [y_k^1, y_k^1, ..., y_k^C]$ is a $C \times N$ matrix of EEG features at k-th time interval where C is the number of EEG features, N is the number of samples from each feature and y_k^i $(i = 1, 2, ...,)$

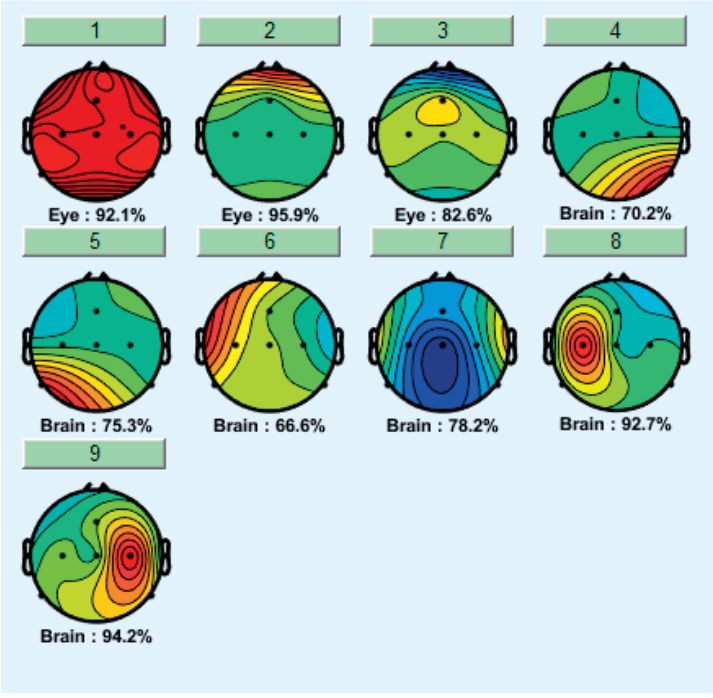


Figure 5.1: Scalp topography maps and assigned labels to every independent component of the EEG signals in a rested-automated test.

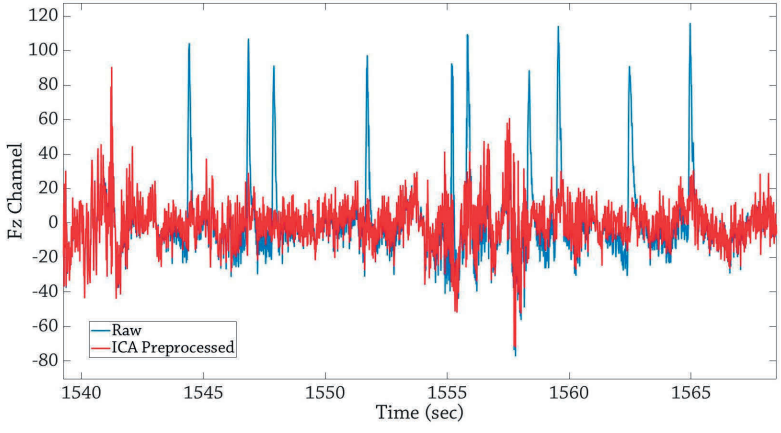


Figure 5.2: A part of the Raw and preprocessed Fz channel after removing eye movement components in a rested-automated test.

is a $1 \times N$ vector represents the extracted features from i-th EEG channel. Moreover, It is assumed that all EEG features are independent of each other for a given PERCLOS x_k. The conditional distribution of each feature is presented by Equation (5.2).

$$y_k^i | x_k \sim f(x_k; \theta_i); i = 1, 2, ..., C \tag{5.2}$$

where $f(.)$ defined the conditional distribution and θ_i is the set of parameters for each i-th EEG channel. Equation (5.1) and Equation (5.2) define the dynamical encoder model, characterizing how changes in EEG features over time encode the PERCLOS progression. In the modeling of the EEG features, it is assumed that the conditional distribution of each feature given PERCLOS follows a normal distribution. The mean of the distribution is defined as a linear function of the PERCLOS and the distribution variance is constant over time. This distribution is defined by Equation (5.3).

$$y_k^i | x_k \sim \mathcal{N}(\alpha^i x_k + \beta^i, \sigma_{vi}^2), \tag{5.3}$$

where α^i, β^i and σ_{vi}^2 are slope parameter, intercept parameter, and the variance of observed noise.

5.3.1 PERCLOS Decoder Model

Bayesian filtering [212] is employed here to estimate PERCLOS from neural data. The Bayesian filter is a recursive technique that can be performed by calculating two equations per time index: one-step perdition and update. The goal of this method is to estimate the PERCLOS values using EEG-based features, therefore Equation (5.1) and Equation (5.3) are used as state process and observation equations of the Bayesian filtering, respectively. Figure 5.3 shows the structure of the Bayesian filtering that has three main steps: 1) using Chapman-Kolmogorov equation [213] to calculate the one-step prediction of the state, 2) using the relationship between every EEG feature given PERCLOS data to compute the likelihood function, and 3) applying the Bayes' rule to update the one-step prediction using the calculated likelihood.

For $k = 1,2, \dots, K$

- **Compute the prediction of the state** (Computation of the prediction distribution of the state x_k by the Chapman–Kolmogorov equation)
$$P(x_k | y_{1:k-1}) = \int P(x_k | x_{k-1}) P(x_{k-1} | y_{1:k-1})\, dx_{k-1}$$
- **Compute the likelihood function:** Computation of the likelihood using measurement equation
$$L(Y_k; x_k) = \prod_{m=1}^{C} N(y_k^m; \alpha^m x_k + \beta^m, \sigma_m^2)$$
- **Update the prediction step:** Computation of the posterior distribution of the state x_k given the measurement y_k at time step k
$$P(x_k | y_{1:k}) = \frac{1}{Z_k} L(Y_k; x_k) P(x_k | y_{1:k-1})$$
$$Z_k = \int L(Y_k; x_k) P(x_k | y_{1:k-1})\, dx_k$$

Figure 5.3: Three main steps of the Bayesian filtering that is used as the decoder to estimate the PERCLOS using EEG features [214].

5.3.2 Model Identification

This subsection explains how to identify the unknown parameters of the encoder model presented in Equation (5.1) and Equation (5.3).

Estimation of state transition process' parameters: The state transition process defined in Equation (5.1) can be rewritten by Equation (5.4). Now, the equation becomes a linear function of a and b when PERCLOS values are known:

$$ax_{k-1} + b + \epsilon_{k-1} = \text{arctanh}(2x_k - 1); k = 1, 2, ..., K \tag{5.4}$$

where K is the total number of PERCLOS sample calculated in all of the driving tests, a, b, and noise variance σ_ϵ^2 are also estimated using the Least Square (LS) technique.

In the LS, it is assumed that $z_k = \text{arctanh}(2x_k - 1)$ is the input for the regression problem:

$$ax_{k-1} + b + \epsilon_{k-1} = z_k; k = 1, 2, ..., K \tag{5.5}$$

In the LS, the Residual Sum of Square (RSS) error, as presented in Equation (5.6), is minimized to obtain the parameters of a and b:

$$RSS_x = \sum_{k=1}^{K} \epsilon_{k-1}^2 = \sum_{k=1}^{K} (z_k - ax_{k-1} - b)^2, \tag{5.6}$$

In order to minimize RSS_x, the derivatives of the RSS_x w.r.t two parameters of a and b are set to be zero:

$$\frac{\partial RSS_x}{\partial a} = 0; \frac{\partial RSS_x}{\partial b} = 0 \tag{5.7}$$

After applying the Equation (5.7) and simplification of the results, the parameters of a and b are calculated as presented in Equation (5.8) [215].

$$a = \frac{\sum_{k=1}^{K} \left(x_{k-1} - \mu_{k-1}^x\right)(z_k - \mu_z)}{\sum_{k=1}^{K} \left(x_{k-1} - \mu_{k-1}^x\right)^2}; b = \mu_z - a\mu_{k-1}^x, \tag{5.8}$$

where $\mu_{k-1}^x = \frac{1}{K-1}\sum_{k=1}^{K-1} x_k$ and $\mu_z = \frac{1}{K}\sum_{k=1}^{K} z_k$ are the PERCLOS mean and input mean, respectively. The noise variance is also calculated as the variance of $\epsilon_{k-1} = z_k - ax_{k-1} - b$ $(k = 1, 2, ..., K)$.

Estimation of observation equation's parameters: Linear regression is also used to identify the parameters of the observation equation per EEG feature (α^i, β^i and σ_{vi}^2). The Equation (5.3) can be rewritten as Equation (5.9).

$$y_k^i = \alpha^i x_k + \beta^i + \nu_k^i; \nu_k^i \sim \mathcal{N}(0, \sigma_{\nu i}^2) \tag{5.9}$$

The RSS error (RSS_y) is calculated for each EEG features by using the Equation (5.10).

$$RSS_y = \sum_{k=1}^{K} \left(y_k^i - \alpha^i x_k - \beta^i\right)^2, \tag{5.10}$$

After minimization of RSS_y for every EEG feature, the unknown parameters of α^i and β^i are determined using Equation (5.11) [215].

$$\alpha^i = \frac{\sum_{k=1}^{K} (x_k - \mu_k^x)\left(y_k^i - \mu_y^i\right)}{\sum_{k=1}^{K} \left(x_{k-1} - \mu_k^x\right)^2}; \beta^i = \mu_y^i - \alpha\mu_k^x, \tag{5.11}$$

where $\mu_k^x = \frac{1}{K}\sum_{k=1}^{K} x_k$ and $\mu_y^i = \frac{1}{K}\sum_{k=1}^{K} y_k^i$.

5.3.3 Model Selection for the Observation Model

This subsection describes the procedure used to select the observation model for EEG features. Though all the neural features can be used in the decoding step, it is more practical to pick a subset of features that shows strong encoding properties. This process helps to build a more robust decoder model by excluding those features which lack reliable and consistent predictive power. With the independence assumption of the neural features, the statistical significance of encoding power of each feature is investigated by checking the distribution of ν_k^i in Equation (5.9) by using the one-sample t-test. The null hypothesis is that ν_k^i comes from a normal distribution with zero-mean and unknown variance. Therefore, a t-test per each neural feature has been applied and checked whether the null hypothesis is rejected or not. The p-value of 0.05 has is set as the significance level and only those features that their corresponding p-value is higher than this level are selected. This subset of features is then used in the decoding step. Therefore, in the decoding step, only a subset of neural features has been picked, whose statistical significance is in a favour of being included in the encoder model feature set.

5.4 Application of the Proposed Methodology

This section firstly describes how neural features are extracted from EEG signals then it explains use of the encoder-decoder pipeline to build the PERCLOS predictor model. According to the literature, the preprocessed EEG data of each channel is decomposed into four sub-bands using band-pass filtering including delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), and beta (12-30 Hz) [50]. For instance, the sub-bands of the Cz channel along with the original EEG signal are presented in Figure 5.4. Statistical features of these sub-bands have been employed in different applications to reduce the dimensionality of the EEG data while significant information is retained during feature extraction [216].

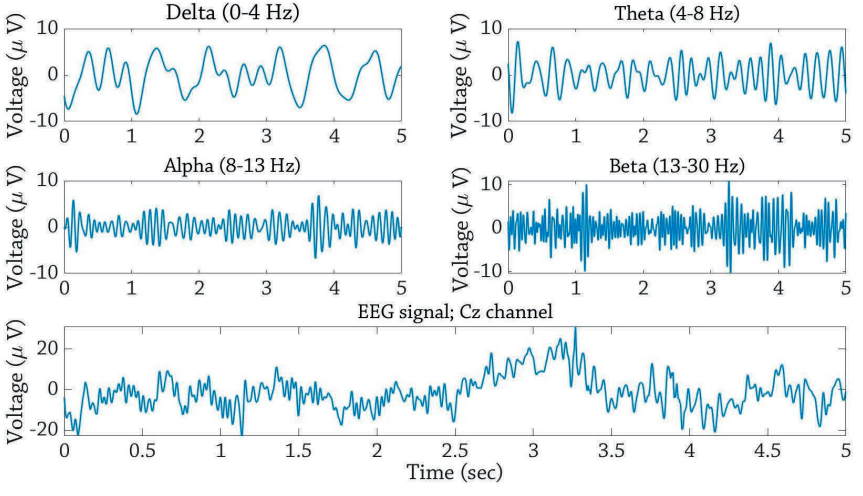


Figure 5.4: Cz channel and its subbands

Extracted features from EEG channels are presented in Table 5.1. Overall, 50 features are extracted from each of the eight EEG channels and one EOG channel that results in 450 EEG features for each driving test. The same sliding time window that is used to calculate the

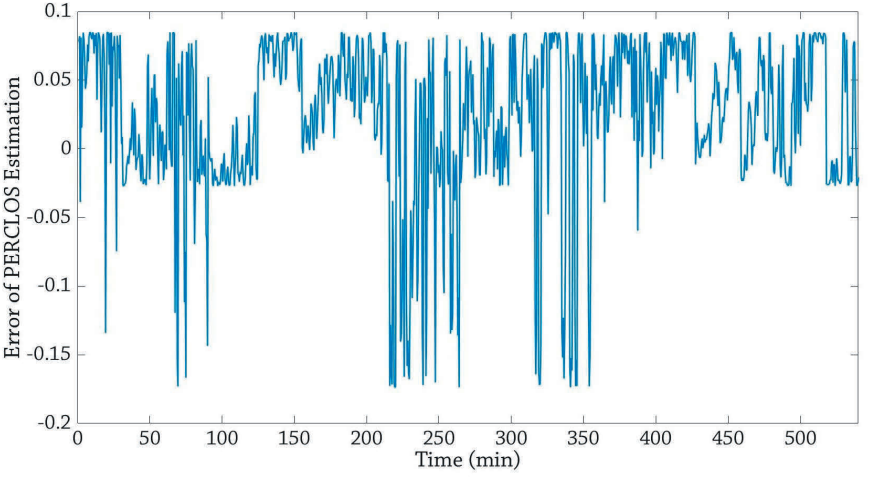


Figure 5.5: The performance of the LS method to estimate the dynamics of the **actual** PERCLOS data.

PERCLOS (1-minute length with a 30-second overlap) has also been exploited for EEG feature extraction.

To find the state transition process parameters, PERCLOS data across all users are concatenated (resulted in approximately 540 minutes of driving) and the LS method is used to estimate a, b, and σ_ϵ^2 parameters. It is assumed that all users have a reasonably similar state transition process. Table 5.2 represents the estimated parameters for PERCLOS dynamical model defined by Equation (5.1). Figure 5.5 shows the PERCLOS residual error of the outputted result from the LS method using the estimated parameters of the state transition equation. As Figure 5 shows, the absolute value of the residual error in some parts is about two to three times larger than other parts. After checking the actual PERCLOS values, it is realized that these parts are either associated with very high (approximately one) or very low (approximately zero) actual PERCLOS values that show the states of completely alert and extremely drowsy situations, respectively. However, it is aimed to model the dynamic transition between these two states (completely alert and extremely drowsy). The root mean squares error (RMSE) between actual PERCLOS and modeled PERCLOS is 0.061. This result suggests that the proposed state transition process can reasonably capture the PERCLOS dynamics with acceptable performance.

In the encoder model, a subset of EEG features has been selected using the model selection approach. Note that the decoder model combines two sources of information at different temporal scales in the prediction of PERCLOS: (1) long-term information that is carried by the state process and (2) instantaneous information carried by neural activity about PERCLOS. These two sources of information are combined through Bayesian Filtering in the estimation of PERCLOS.

Using the proposed modeling framework, a user-specific encoder and decoder model of PERCLOS is built. It is only assumed that PERCLOS temporal dynamics across users share the same characteristics. Given this model, it is possible that a neural feature might be positively correlated with PERCLOS in one participant and negatively correlated in another one. Whilst this might provide a more accurate prediction given the specificity of the model to a specific user, it is possible to search for possible neural biomarkers which are showing consistent encoding properties across participants. Therefore, encoding step helps to find possible biomarkers

Table 5.1: Extracted features from EEG channels; Std.: standard deviation, Min.: minimum, Max.: maximum, Mean En.: mean energy, Log. En. Ent.: log energy entropy, Sh. Ent.: Shannon entropy, Hjorth Mob.: Hjorth mobility.

Index	Feature	Index	Feature
1	Mean of Delta	26	Skewness of Theta
2	Mean of Theta	27	Skewness of Alpha
3	Mean of Alpha	28	Skewness of Beta
4	Mean of Beta	29	Kurtosis of Delta
5	Std. of Delta	30	Kurtosis of Theta
6	Std. of Theta	31	Kurtosis of Alpha
7	Std. of Alpha	32	Kurtosis of Beta
8	Std. of Beta	33	Log En. Ent. of Delta
9	Min. of Delta	34	Log En. Ent. of Theta
10	Min. of Theta	35	Log En. Ent. of Alpha
11	Min. of Alpha	36	Log En. Ent. of Beta
12	Min. of Beta	37	Sh. Ent. of Delta
13	Max. of Delta	38	Sh. Ent. of Theta
14	Max. of Theta	39	Sh. Ent. of Alpha
15	Max. of Alpha	40	Sh. Ent. of Beta
16	Max. of Beta	41	Hjorth Mob. Delta
17	Median of Delta	42	Hjorth Mob. Theta
18	Median of Theta	43	Hjorth Mob. Alpha
19	Median of Alpha	44	Hjorth Mob. Beta
20	Median of Beta	45	Power of Delta (P_δ)
21	Mean En. of Delta	46	Power of Theta (P_θ)
22	Mean En. of Theta	47	Power of Alpha (P_α)
23	Mean En. of Alpha	48	Power of Beta (P_β)
24	Mean En. of Beta	49	P_β/P_α
25	Skewness of Delta	50	$P_\beta/(P_\alpha + P_\theta)$

Table 5.2: State transition process model parameters

Parameter	Meaning	Value
σ_ϵ^2	Noise variance of state transition process model	0.03
a	The slope of the linear regression in Equation 5.4	3.93
b	The intercept of the linear regression in Equation 5.4	-1.79

that are representing PERCLOS changes consistently across users.

5.5 Results

In this section, the modeling results of the proposed encoder-decoder framework in the estimation of PERCLOS is discussed and the two following questions are answered:

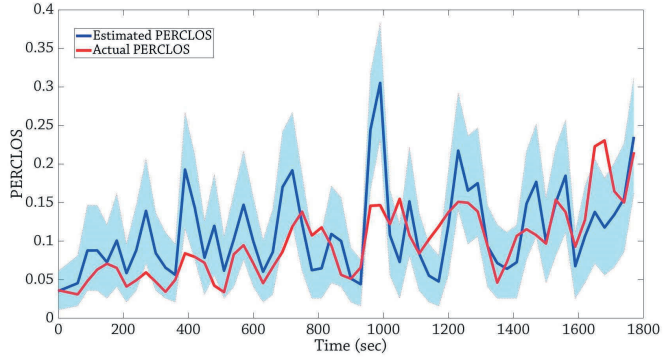
1. How accurate is the estimated PERCLOS in different driving tests?
2. Are there any neural biomarkers representing PERCLOS independent of the driving test? If yes, how are these biomarkers correlated with PERCLOS?

The data set has been randomly separated into two data sets: train and test. The training set contains 15 tests where three driving tests with ID = 6, 9, and 15 have been selected to make a test dataset. The test dataset has not been involved for selecting the biomarkers. Neural biomarkers are found using training dataset and are employed to estimate the PERCLOS in the test dataset. Figure 5.6 shows the decoding results of the test dataset. These results suggest that the proposed framework reasonably traces the drowsiness level presented by the actual PERCLOS data. Figure 5.6 also presents the upper and lower bounds of the 95% confidence interval of the Bayesian estimation. These bounds are utilized to calculate the High Probability Density (HPD) percentage [217]. The HPD presents the percentage of the data samples per driving test where the actual PERCLOS falls in the 95% confidence interval of the estimated one.

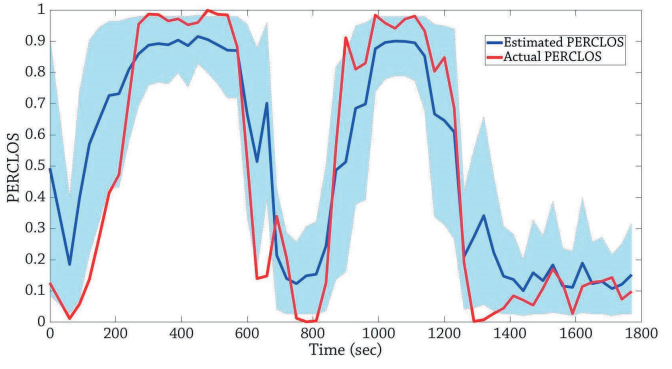
Figure 5.7 shows the RMSE and HPD percentage performance metrics of our proposed method for PERCLOS estimation. It shows that the average RMSE and average HPD percentage are 0.117 and 62.5%, respectively. To investigate the performance of the proposed method in the different levels of driver drowsiness, PERCLOS has been separated into four intervals: 0-0.25, 0.25-0.5, 0.5-0.75, and 0.75-1. The average RMSE and HPD percentage of each one of these intervals during all of the driving tests are presented in Figure 5.8. According to this Figure, this average HPD percentage and RMSE are increasing and decreasing as PERCLOS grows, respectively. Therefore, this model obtains better performance in the higher actual PERCLOS (moderate and extreme levels of driver drowsiness), which in practice is more important to detect the driver drowsiness than states with low PERCLOS values.

To find consistent features that can be neural biomarkers, the EEG features that present a strong correlation to PERCLOS based on their corresponding p-values are investigated. According to the results of the encoder model, 28 EEG features are presented in the selected feature sets for all of the 18 driving tests. Therefore, they can be used as biomarkers to study drowsiness independent of drivers. These biomarkers are presented in Table 5.3 that include skewness of Alpha (for all EEG channels), Delta power (for all EEG channels), Theta power (for all EEG channels except Cz and P08), Hjorth mobility of Delta (for T8, P08, and EOG channels).

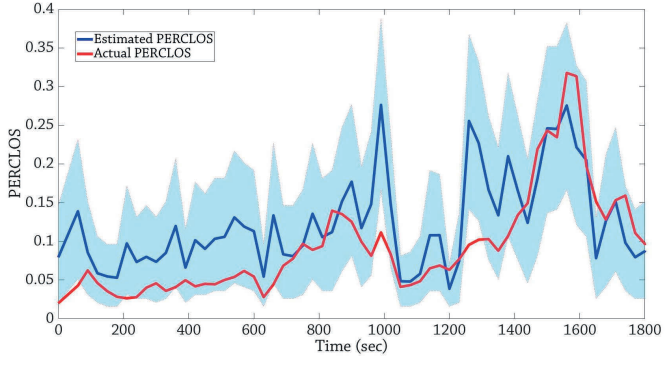
Figure 5.9 also shows the average correlation coefficient between every biomarker and PERCLOS in all driving tests. As this Figure shows, the consistent Delta powers and Theta powers are positively correlated with PERCLOS in all EEG channels. This result is in accordance with the established studies that report increases in the Theta and Delta powers as indicators of drowsiness [218, 219]. The skewness of Alpha in all EEG channels except T8 and P07 is also positively correlated with PERCLOS while Hjorth Mobility of the Delta in T8 and P08 are negatively and in EOG channel positively correlated with PERCLOS. Therefore, the proposed framework establishes biomarkers that have consistent relationships with PERCLOS and these neural biomarkers should be extracted from EEG signals to estimate the drowsiness independently of the drivers and driving conditions.



(a)



(b)



(c)

Figure 5.6: Decoding results in three driving tests with ID=6 (a), ID=9 (b), and ID=15 (c) for estimation of PERCLOS using selected EEG features. Light blue shaded areas show the 95% confidence interval of the estimated PERCLOS. The result suggests a strong correspondence between measured PERCLOS and estimated one.

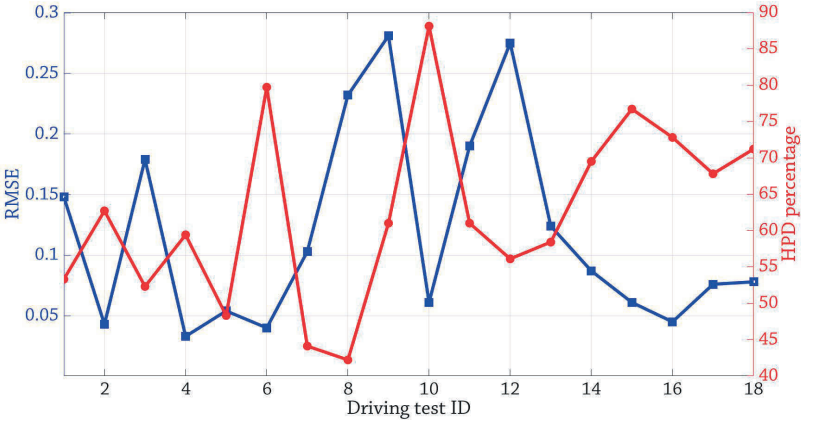


Figure 5.7: RMSE and HPD% metrics to evaluate the performance of the proposed encoding-decoding framework. The average RMSE and average HPD percentage across different driving tests are 0.117 and 62.5%, respectively.

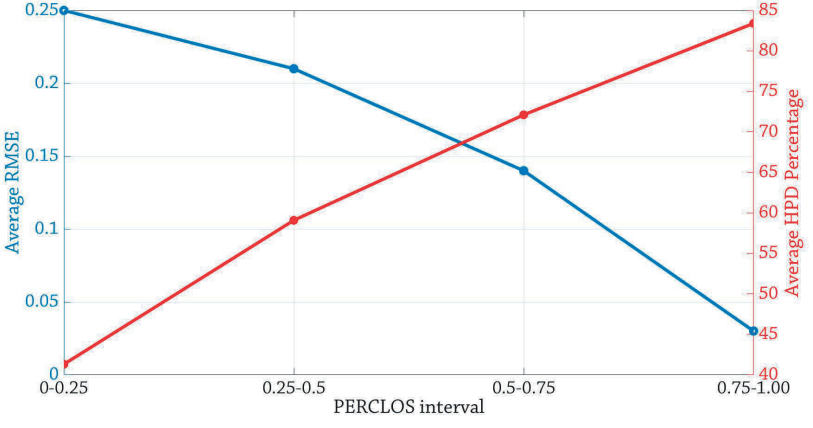


Figure 5.8: The average RMSE and HPD percentage in different PERCLOS intervals. This Figure shows that the proposed method is performing better in higher values of PERCLOS that are more important to detect the moderate and extreme levels of driver drowsiness.

Table 5.3: EEG features that are consistently significant ($p - value < 0.05$) across all the 18 studied driving tests. Overall, 28 features are selected by the encoder regardless of driving tests to encode the PERCLOS dynamics. These features include skewness of Alpha (all EEG channels), Delta power (all EEG channels), Theta power (all EEG channels except Cz and P08), Hjorth mobility of Delta (T8, P08, and EOG channels).

Feature	**Channel**	**Number**
Skewness of Alpha	Cz, Fz, T7, T8, C3, C4, P07, P08, EOG	9
Delta power	Cz, Fz, T7, T8, C3, C4, P07, P08, EOG	9
Theta power	Fz, T7, T8, C3, C4, P07, EOG	7
Hjorth Mobility of Delta	T8, P08, EOG	3
–	**Sum**	**28**

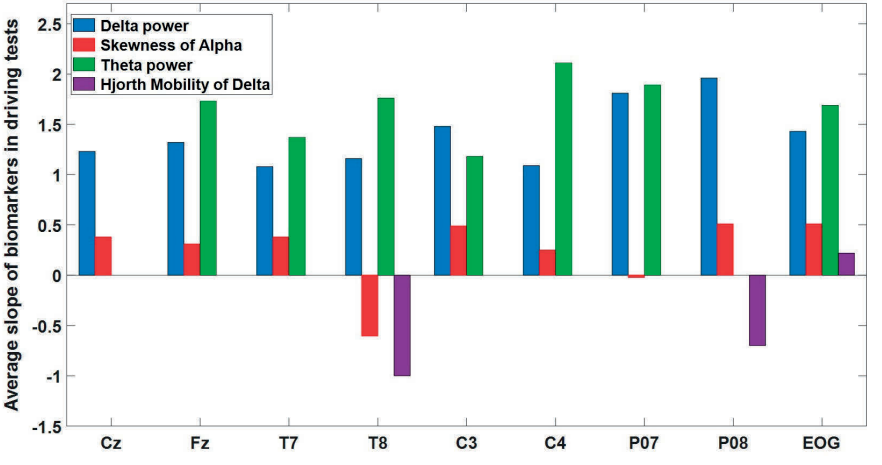


Figure 5.9: The average slope of the biomarkers in the driving tests. Delta and Theta powers of EEG channels are positively correlated with PERCLOS in all of the driving tests.

Figure 5.10 shows the relationships between biomarkers and PERCLOS in the driving tests. According to this Figure, about 73% and 66% of the consistent Theta and Delta powers in the driving tests are positively correlated with PERCLOS, respectively. On the other hand, only 48.8% and 44.4% of the consistent skewness of Alpha and Hjorth mobility of are positively correlated with PERCLOS. Therefore, discovered biomarkers makes a "push-pull mechanism" to estimate the driver drowsiness. In this mechanism, one group of biomarkers that includes Theta and Delta powers are increasing with increasing the level of drowsiness (pushing part of the mechanism), whereas another group of biomarkers that consists of skewness of Alpha and Hjorth mobility of Delta are decreasing with increasing the drowsiness level (pulling part of the mechanism). The interactions between these two parts of the mechanism obtains a satisfying estimation of driver drowsiness associated with the PERCLOS data.

5.6 Discussion

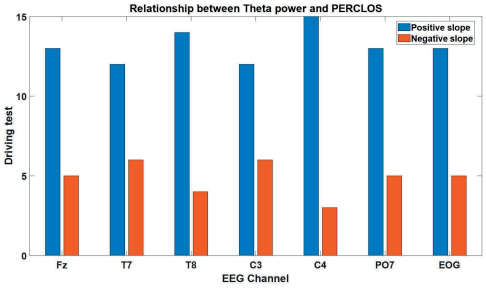
In this research, a new modeling framework using neural activities is developed to provide an instantaneous estimation of the PERCLOS as a widely used estimation of driver drowsiness. The PERCLOS is being considered as a robust correlate of driver drowsiness which is widely studied to assess driver's performance in the different vigilance states [220]. The proposed framework is derived from extensive work in the neuroscience domain where the question was finding the relationship between cognitive state and neural correlates [221]. The framework has two steps: encoder and decoder. When each of these steps is built through a sequential process, the dynamical estimation of PERCLOS is achieved as a function of the selected number of neural features. One of the advantages of this method compared to previously developed methods is that a posterior distribution of PERCLOS at every time point is provided which is a fairly complete measure of PERCLOS. Through this measure, other metrics can be provided to assess the progression of a driver's drowsiness and anticipate the time that the driver can be in a dangerous level of driver drowsiness or even decide about whether the driver drowsiness level is higher above one specific predefined level or not (a predefined threshold might be used to trigger appropriate action or warning). Another advantage of this model is real-time estimation of driver drowsiness that can reduce the risk of accidents caused by drowsy driving.

The principal output of this approach is finding the neural biomarkers for driver drowsiness which has not been extensively studied in the previous works as those works were classifying driver drowsiness as a black-box model [222, 223, 224].

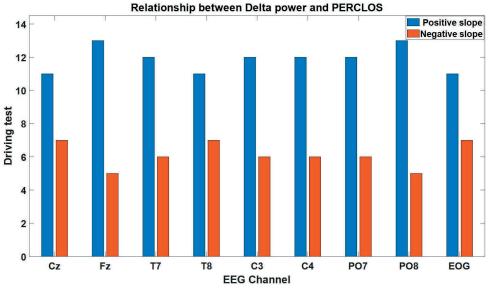
Despite promising results of this method, more research should be conducted to enhance the performance of this method. Some of the challenges that need to be addressed are as follows:

1. Producing a personalized model is possible by including more EEG channels but there is a trade-off between the utility of the device and the number of EEG channels. More physiological information such as ECG and heart rate variability data [38, 39] can be utilized to enhance the model performance rather than increasing the number of the EEG channels.

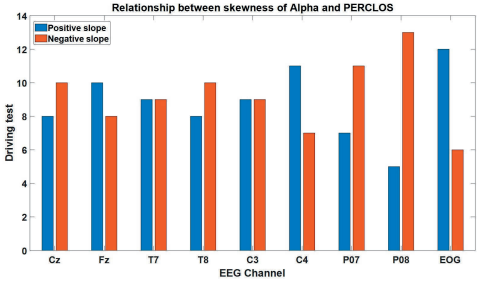
2. Although the proposed method provides a solution to find neural biomarkers, only the fluctuation of nodal frames are studied while more advanced techniques are studying network global dynamics [225]. Employing other features that are presenting the network global dynamics of brain activities like coherence, correlation and mutual information between different EEG channels [226, 227, 228] might improve the performance of the framework. It should be noted that the proposed framework is flexible enough to incorporate those features into our model but the question is which one of them are informative to estimate the driver drowsiness. The proposed framework can also be helpful to select the informative features.



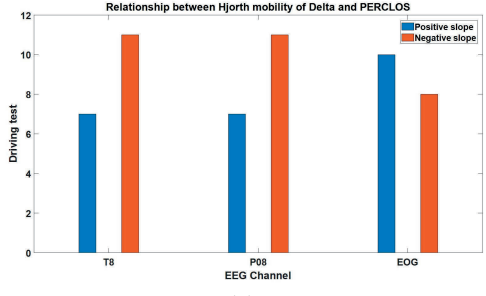
(a)



(b)



(c)



(d)

Figure 5.10: Slope signs of the biomarkers in different driving tests that are statistically significant. According to these results, discovered biomarkers make a 'push-pull mechanism' to estimate the driver drowsiness.

3. This chapter discussed the estimation or decoding capability of the proposed framework but another important application of this method is its prediction capability where the level of drowsiness can be predicted based on the current and previous neural activities. This capability requires to use of a more accurate state transition process that is tuned for every individual driver.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

In this thesis, the problem of driver drowsiness estimation and classification using data fusion of three different data sources was addressed. Three data sources were (1) vehicle-based data, (2) facial-based data, and (3) biosignals. The used data set was collected during the WACHSens project from 92 participants that each one of them performed 4 different driving experiments: rested-automated, rested-manual, fatigued-automated, and fatigued-manual. The ground truth of drowsiness classification was obtained based on the video observations of the drowsiness signs in the simulated driving tests.

One of the main contribution of this work was proposing a methodology using traditional machine learning methods for data fusion that was explained in Chapter 3. Unlike the previous works that mostly classified the drowsiness in two classes of alert and drowsy, the drowsiness was classified here in three classes: alert, moderately drowsy, and extremely drowsy. These three classes provided a better resolution for the progression in driver drowsiness and can help to decrease the false alarm and detect the critical level of drowsiness in a proper time to prevent drowsiness-related crashes. Different statistical features were extracted from every source of the data and neighbourhood component analysis was employed to remove the redundant or irrelevant features. Two different broadly used classifiers including K-nearest neighbours and random forest were used to classify the drowsiness. Results showed that this method achieved the balanced accuracy of about 83% and 75% for drowsiness classification in the manual and automated driving modes, respectively. However, this method did not obtain satisfying performance for classification using only ECG signals because the ECG-based features were extracted from heart rate variability data based on the R-peak detection algorithm however the ECG signals have more sub-waves such as P and T waves that can carry important information regarding the driver drowsiness that was ignored in Chapter 3.

In order to improve the drowsiness classification based on the ECG signal, a deep learning method was proposed in Chapter 4. In this Chapter, two deep convolutional neural networks were built where the Bayesian optimization method was employed to achieve the optimal hyperparameter combination that provides the best performance among others. The input data of these networks were the grayscale wavelet scalogram images extracted from ECG and eyelid signals using the Morse wavelet. Therefore, two types of scalogram images were obtained: eyelid-based and ECG-based. Each one of these types was separately used as input data to the one deep neural network. The weight regularization and dropout layers were used to avoid the overfitting issue during the training process. The batch normalization layers were also placed in the architecture of the deep network to speed up the training process. An ensemble learning method based on the soft voting algorithm was also applied to the outputs of the deep networks on the test data set. Results showed that the deep learning method can significantly outperform the traditional machine learning methods trained by the ECG-based features. According

to the results, the balanced accuracy of the deep convolutional neural network trained by ECG scalogram images was about 81% and 80% in the manual and automated modes, respectively. While these accuracies were about 54% in both of the manual and automated modes for the traditional machine learning methods used in Chapter 3.

In Chapter 5, a neural encoder-decoder modeling framework was proposed to estimate the driver drowsiness using EEG and EOG channels. In this Chapter, the PERCLOS was used as the ground truth for driver drowsiness and the Bayesian filter was estimate PERCLOS using selected EEG features. In the encoder part of the framework, the EEG features that showed significant correlation to the PERCLOS were selected to be used as the input to the Bayesian filter that estimated the actual PERCLOS values. Moreover, this framework could obtain a set of neural features that were significant in all of the driving tests. These EEG features can be helpful to decrease the computational load of the EEG processing methods for driver drowsiness detection.

6.2 Future Work

In order to improve the performance and reliability of the proposed methods in this thesis, the following tasks are suggested:

1. In the driving tests performed in the WACHSens project, drivers were aware of the mode of the driving test (manual or automated) before starting the test. In future research, different test procedures that are more similar to real-world automated driving situations can be investigated. Some control transition from an automated system to the driver can be implemented in the test procedure to study the value of the reaction time in the drivers in different levels of drowsiness of sleep deprivation.

2. An alert system can be designed in future research to warn the driver based on the proposed classification methods in this thesis. The acceptability of this system should be investigated during the simulated driving tests.

3. Deep transfer learning can be used to design a personalized driver drowsiness detection system based on the proposed generic model in this thesis. This method will help to reduce the amount of data needed to design a personalized system.

Bibliography

[1] M. Awais, N. Badruddin, and M. Drieberg, "A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability," *Sensors*, vol. 17, no. 9, 2017. [Online]. Available: https://doi.org/10.3390/s17091991

[2] AAA Foundation for Traffic Safety, "2019 traffic safety culture index (technical report), june 2020," Washington, D.C. [Online]. Available: https://aaafoundation.org/2019-traffic-safety-culture-index/

[3] National Highway Traffic Safety Administration, "Traffic safety facts: 2017 fatal motor vehicle crashes: Overview," 1200 New Jersey Avenue SE., Washington. [Online]. Available: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812603

[4] Agustina Garcés Correa, Lorena Orosco, and Eric Laciar, "Automatic detection of drowsiness in eeg records based on multimodal analysis," *Medical Engineering & Physics*, vol. 36, no. 2, pp. 244–249, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1350453313001690

[5] S. Klauer, V. Neale, T. Dingus, Jeremy Sudweeks, and D. J. Ramsey, "The prevalence of driver fatigue in an urban driving environment : Results from the 100-car naturalistic driving study," 2006.

[6] Fraunhofer-Gesellschaft, "Eyetracker warns against momentary driver drowsiness - press release oktober 12, 2010," 4/14/2021. [Online]. Available: https://www.fraunhofer.de/en/press/research-news/2010/10/eye-tracker-driver-drowsiness.html

[7] M. Gonçalves, R. Amici, R. Lucas, T. Åkerstedt, F. Cirignotta, J. Horne, D. Léger, W. T. McNicholas, M. Partinen, J. Téran-Santos, P. Peigneux, L. Grote, and National Representatives as Study Collaborators, "Sleepiness at the wheel across europe: a survey of 19 countries," *Journal of Sleep Research*, vol. 24, no. 3, pp. 242–253, 2015. [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1111/jsr.12267

[8] T. Inagaki and T. B. Sheridan, "A critique of the sae conditional driving automation definition, and analyses of options for improvement," *Cognition, Technology & Work*, vol. 21, no. 4, pp. 569–578, 2019. [Online]. Available: https://link.springer.com/article/10.1007/s10111-018-0471-5

[9] A. Shahid, K. Wilkinson, S. Marcu, and C. M. Shapiro, "Karolinska sleepiness scale (kss)," in *STOP, THAT and one hundred other sleep scales*, A. Shahid, Ed. New York: Springer, 2012, pp. 209–210. [Online]. Available: https://link.springer.com/chapter/10.1007/978-1-4419-9893-4_47

[10] G. R. Poudel, C. R. H. Innes, and R. D. Jones, "Distinct neural correlates of time-on-task and transient errors during a visuomotor tracking task after sleep restriction," *NeuroImage*, vol. 77, pp. 105–113, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811913003042?via%3Dihub

[11] K. Kaida, M. Takahashi, T. Akerstedt, A. Nakata, Y. Otsuka, T. Haratani, and K. Fukasawa, "Validation of the karolinska sleepiness scale against performance and eeg variables," *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, vol. 117, no. 7, pp. 1574–1581, 2006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1388245706001428

[12] Z. Li, S. E. Li, R. Li, B. Cheng, and J. Shi, "Online detection of driver fatigue using steering wheel angles for real driving conditions," *Sensors (Basel, Switzerland)*, vol. 17, no. 3, 2017. [Online]. Available: https://www.mdpi.com/1424-8220/17/3/495

[13] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American journal of physiology. Heart and circulatory physiology*, vol. 278, no. 6, pp. H2039–49, 2000. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/10843903/

[14] D. J. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series," in *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*, ser. AAAIWS'94. AAAI Press, 1994, pp. 359–370. [Online]. Available: https://dl.acm.org/doi/10.5555/3000850.3000887

[15] Meng Chai, shi-wu Li, wen-cai Sun, meng-zhu Guo, and meng-yuan Huang, "Drowsiness monitoring based on steering wheel status," *Transportation Research Part D: Transport and Environment*, vol. 66, pp. 95–103, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1361920917306582

[16] F. Friedrichs and B. Yang, "Drowsiness monitoring by steering and lane data based features under real driving conditions," in *2010 18th European Signal Processing Conference*, 2010, pp. 209–213. [Online]. Available: https://ieeexplore.ieee.org/document/7096521

[17] J. C. Bezdek, S. K. Chuah, and D. Leep, "Generalized k-nearest neighbor rules," *Fuzzy Sets and Systems*, vol. 18, no. 3, pp. 237–256, 1986. [Online]. Available: https://doi.org/10.1016/0165-0114(86)90004-7

[18] D. Reynolds, "Gaussian mixture models," in *Encyclopedia of biometrics*, S. Z. Li and A. K. Jain, Eds. New York and London: Springer, 2009, pp. 659–663. [Online]. Available: https://link.springer.com/referenceworkentry/10.1007%2F978-0-387-73003-5_196

[19] A. J. Izenman, "Linear discriminant analysis," in *Modern multivariate statistical techniques*, ser. Springer Texts in Statistics, A. J. Izenman, Ed. New York: Springer, 2008, pp. 237–280. [Online]. Available: https://link.springer.com/chapter/10.1007%2F978-0-387-78189-1_8

[20] Daniel Berrar, *Bayes' theorem and naive Bayes classifier*. Amsterdam, Netherlands: Elsevier, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128096338204731?via%3Dihub

[21] D. Ververidis and C. Kotropoulos, Eds., *Emotional Speech Classification Using Gaussian Mixture Models and the Sequential Floating Forward Selection Algorithm*. IEEE, 2005. [Online]. Available: https://ieeexplore.ieee.org/document/1521717

[22] Xuesong Wang and Chuan Xu, "Driver drowsiness detection based on non-intrusive metrics considering individual specifics," *Accident Analysis & Prevention*, vol. 95, pp. 350–357, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0001457515300609

[23] Anthony D. McDonald, John D. Lee, Chris Schwarz, and Timothy L. Brown, "A contextual and temporal algorithm for driver drowsiness detection," *Accident Analysis & Prevention*, vol. 113, pp. 25–37, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0001457518300058

[24] Z. Ghahramani, "Learning dynamic bayesian networks," in *Adaptive processing of sequences and data structures*, ser. Lecture notes in computer science, Lecture notes in artificial intelligence 0302-9743, C. L. Giles and M. Gori, Eds. Berlin and London: Springer, 1998, vol. 1387, pp. 168–197. [Online]. Available: https://link.springer.com/chapter/10.1007/BFb0053999

[25] G. Biau and E. Scornet, "A random forest guided tour," *TEST*, vol. 25, no. 2, pp. 197–227, 2016. [Online]. Available: https://doi.org/10.1007/s11749-016-0481-7

[26] Y. Sun, J. Li, J. Liu, B. Sun, and C. Chow, "An improvement of symbolic aggregate approximation distance measure for time series," *Neurocomputing*, vol. 138, pp. 189–198, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231214002872

[27] Anthony D. McDonald, John D. Lee, Chris Schwarz, and Timothy L. Brown, "Steering in a random forest: Ensemble learning for detecting drowsiness-related lane departures," *Human Factors*, vol. 56, no. 5, pp. 986–998, 2014. [Online]. Available: https://doi.org/10.1177/0018720813515272

[28] Walter W. Wierwille and Lynne A. Ellsworth, "Evaluation of driver drowsiness by trained raters," *Accident Analysis & Prevention*, vol. 26, no. 5, pp. 571–581, 1994. [Online]. Available: https://doi.org/10.1016/0001-4575(94)90019-1

[29] J. Krajewski, D. Sommer, U. Trutschel, D. Edwards, and M. Golz, "Steering wheel behavior based estimation of fatigue," in *Proceedings of the 5th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design : Driving Assessment 2009*. Iowa City, Iowa: University of Iowa, 2009, pp. 118–124. [Online]. Available: https://ir.uiowa.edu/drivingassessment/2009/papers/18/

[30] M. Ingre, T. Akerstedt, 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. [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-2869.2006.00504.x

[31] R. Rodríguez, A. Mexicano, J. Bila, S. Cervantes, and R. Ponce, "Feature extraction of electrocardiogram signals by applying adaptive threshold and principal component analysis," *Journal of Applied Research and Technology*, vol. 13, no. 2, pp. 261–269, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1665642315000103

[32] Y. Kubota, L. Y. Chen, E. A. Whitsel, and A. R. Folsom, "Heart rate variability and lifetime risk of cardiovascular disease: the atherosclerosis risk in communities study," *Annals of epidemiology*, vol. 27, no. 10, pp. 619–625.e2, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S104727971730515X?via%3Dihub

[33] F. Sessa, V. Anna, G. Messina, G. Cibelli, V. Monda, G. Marsala, M. Ruberto, A. Biondi, O. Cascio, G. Bertozzi, D. Pisanelli, F. Maglietta, A. Messina, M. P. Mollica, and M. Salerno, "Heart rate variability as predictive factor for sudden cardiac death," *Aging*, vol. 10, no. 2, pp. 166–177, 2018. [Online]. Available: https://www.aging-us.com/article/101386/text

[34] G. D. Furman, A. Baharav, C. Cahan, and S. Akselrod, "Early detection of falling asleep at the wheel: A heart rate variability approach," in *2008 Computers in Cardiology*, 2008, pp. 1109–1112. [Online]. Available: https://ieeexplore.ieee.org/document/4749240

[35] M. Patel, S.K.L. Lal, D. Kavanagh, and P. Rossiter, "Applying neural network analysis on heart rate variability data to assess driver fatigue," *Expert Syst. Appl.*, vol. 38, no. 6, pp. 7235–7242, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0957417410013916

[36] Muhammad Hendra, Danang Kurniawan, Riza Vina Chrismiantari, Trio Pambudi Utomo, and Nuryani Nuryani, "Drowsiness detection using heart rate variability analysis based on microcontroller unit," *Journal of Physics: Conference Series*, vol. 1153, p. 012047, 2019. [Online]. Available: https://iopscience.iop.org/article/10.1088/1742-6596/1153/1/012047

[37] F. Abtahi, A. Anund, C. Fors, F. Seoane, and K. Lindecrantz, "Association of drivers' sleepiness with heart rate variability: A pilot study with drivers on real roads," in *EMBEC & NBC 2017*, ser. IFMBE proceedings, 1680-0737, H. Eskola, O. Väisänen, J. Viik, and J. Hyttinen, Eds. Singapore: Springer, 2018, vol. 65, pp. 149–152. [Online]. Available: https://link.springer.com/chapter/10.1007%2F978-981-10-5122-7_38

[38] K. Fujiwara, E. Abe, K. Kamata, C. Nakayama, Y. Suzuki, T. Yamakawa, T. Hiraoka, M. Kano, Y. Sumi, F. Masuda, M. Matsuo, and H. Kadotani, "Heart rate variability-based driver drowsiness detection and its validation with eeg," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 6, pp. 1769–1778, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8520803

[39] J. Vicente, P. Laguna, A. Bartra, and R. Bailón, "Drowsiness detection using heart rate variability," *Medical & biological engineering & computing*, vol. 54, no. 6, pp. 927–937, 2016.

[40] N. Rodrígue-Ibáñez, M. A. García-González, M. Fernández-Chimeno, and J. Ramos-Castro, "Drowsiness detection by thoracic effort signal analysis in real driving environments," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, vol. 2011, pp. 6055–6058, 2011.

[41] S.-J. Jung, H.-S. Shin, and W.-Y. Chung, "Driver fatigue and drowsiness monitoring system with embedded electrocardiogram sensor on steering wheel," *IET Intelligent Transport Systems*, vol. 8, no. 1, pp. 43–50, 2014.

[42] G. Li and W.-Y. Chung, "Detection of driver drowsiness using wavelet analysis of heart rate variability and a support vector machine classifier," *Sensors (Basel, Switzerland)*, vol. 13, no. 12, pp. 16 494–16 511, 2013. [Online]. Available: https://www.mdpi.com/1424-8220/13/12/16494

[43] M. H. Zweig and G. Campbell, "Receiver-operating characteristic (roc) plots: a fundamental evaluation tool in clinical medicine," *Clinical Chemistry*, vol. 39, no. 4, pp. 561–577, 1993.

[44] S. Bersimis, S. Psarakis, and J. Panaretos, "Multivariate statistical process control charts: an overview," *Quality and Reliability Engineering International*, vol. 23, no. 5, pp. 517–543, 2007.

[45] Ruben Buendia, Fabio Forcolin, Johan Karlsson, Bengt Arne Sjöqvist, Anna Anund, and Stefan Candefjord, "Deriving heart rate variability indices from cardiac monitoring—an indicator of driver sleepiness," *Traffic Injury Prevention*, vol. 20, no. 3, pp. 249–254, 2019.

[46] Debdeep Sikdar, Rinku Roy, and Manjunatha Mahadevappa, "Epilepsy and seizure characterisation by multifractal analysis of eeg subbands," *Biomedical Signal Processing and Control*, vol. 41, pp. 264–270, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1746809417302896

[47] S. Yaacob, N. I. Z. Muhamad'Arif, P. Krishnan, A. Rasyadan, M. Yaakop, and F. Mohamed, "Early driver drowsiness detection using electroencephalography signals," in *IEEE International Conference on Artificial Intelligence in Engineering & Technology (IICAIET 2020)*. [Piscataway, NJ]: IEEE, 2020, pp. 1–4.

[48] S.-H. Hwang, M. Park, J. Kim, Y. Yun, and J. Son, "Driver drowsiness detection using eeg features," in *HCI International 2018 – posters' extended abstracts. Part III*, ser. Communications in Computer and Information Science, C. Stephanidis, Ed. Cham, Switzerland: Springer, 2018, vol. 852, pp. 367–374.

[49] Y. Cui, Y. Xu, and D. Wu, "Eeg-based driver drowsiness estimation using feature weighted episodic training," *IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 27, no. 11, pp. 2263–2273, 2019.

[50] S. Taran and V. Bajaj, "Drowsiness detection using adaptive hermite decomposition and extreme learning machine for electroencephalogram signals," *IEEE Sensors Journal*, vol. 18, no. 21, pp. 8855–8862, 2018.

[51] Y. K. Alp and O. Arıkan, "Time–frequency analysis of signals using support adaptive hermite–gaussian expansions," *Digital Signal Processing*, vol. 22, no. 6, pp. 1010–1023, 2012.

[52] H. Zeng, H. Wu, and J. Lin, "Qrs classification using adaptive hermite decomposition and radial basis function network," in *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Hong Kong SAR, China, October 29-November 1, 1998*, H. K. Chang and Y. T. Zhang, Eds. Piscataway, NJ: IEEE Service Center, 1999, pp. 147–150.

[53] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. E215–20, 2000.

[54] Z. Ren, R. Li, B. Chen, H. Zhang, Y. Ma, C. Wang, Y. Lin, and Y. Zhang, "Eeg-based driving fatigue detection using a two-level learning hierarchy radial basis function," *Frontiers in neurorobotics*, vol. 15, p. 618408, 2021.

[55] M. Awais, N. Badruddin, and M. Drieberg, "Driver drowsiness detection using eeg power spectrum analysis," in *IEEE Region 10 symposium, 2014*. Piscataway, NJ: IEEE, 2014, pp. 244–247.

[56] J. C. Marshall and G. M. Morriss-Kay, "Editorial introduction: Functional anatomy of the brain," *Journal of Anatomy*, vol. 207, no. 1, pp. 1–2, 2005.

[57] M. A. Garces Correa and E. Laciar, "Noise removal from eeg signals in polisomnographic records applying adaptive filters in cascade," in *Adaptive Filtering by Non-Invasive Vital Signals Monitoring and Diseases Diagnosis*, L. Garcia Morales, Ed. INTECH Open Access Publisher, 2011. [Online]. Available: https://www.intechopen.com/chapters/16119

[58] M. B. Kurt, N. Sezgin, M. Akin, G. Kirbas, and M. Bayram, "The ann-based computing of drowsy level," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 2534–2542, 2009.

[59] C. Lin, C. Chang, B. Lin, S. Hung, C. Chao, and I. Wang, "A real-time wireless brain–computer interface system for drowsiness detection," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 4, no. 4, pp. 214–222, 2010. [Online]. Available: https://doi.org/10.1109/TBCAS.2010.2046415

[60] H. Mark and J. Workman, "The statistics of spectral searches," in *Chemometrics in spectroscopy*, H. Mark and J. Workman, Eds. Place of publication not identified: Elsevier Ltd. : Academic Press, 2018, pp. 507–511.

[61] F. Lin, L. Ko, C. Chuang, T. Su, and C. Lin, "Generalized eeg-based drowsiness prediction system by using a self-organizing neural fuzzy system," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 59, no. 9, pp. 2044–2055, 2012.

[62] M. Lennon, G. Mercier, M. C. Mouchot, and L. Hubert-Moy, "Independent component analysis as a tool for the dimensionality reduction and the representation of hyperspectral images," in *IGARSS 2001*. IEEE, 2001, pp. 2893–2895.

[63] Mariette Awad and Rahul Khanna, "Support vector regression," in *Efficient Learning Machines*. Apress, Berkeley, CA, 2015, pp. 67–80. [Online]. Available: https://link.springer.com/chapter/10.1007/978-1-4302-5990-9_4

[64] S. Sieniutycz, "Complex systems of neural networks," in *Complexity and complex thermo-economic systems*, S. Sieniutycz, Ed. Amsterdam: Elsevier, 2020, pp. 51–84.

[65] S. S. Elanayar V T and Y. C. Shin, "Radial basis function neural network for approximation and estimation of nonlinear stochastic dynamic systems," *IEEE transactions on neural networks*, vol. 5, no. 4, pp. 594–603, 1994.

[66] G. Castellano and A. M. Fanelli, "A self-organizing neural fuzzy inference network," in *IJCNN 2000*, S. Amari, Ed. Los Alamitos, Cailf.: IEEE Computer Society, 2000, pp. 14–19 vol.5. [Online]. Available: https://doi.org/10.1109/IJCNN.2000.861428

[67] Y. Ma, B. Chen, R. Li, C. Wang, J. Wang, Q. She, Z. Luo, and Y. Zhang, "Driving fatigue detection from eeg using a modified pcanet method," *Computational Intelligence and Neuroscience*, vol. 2019, p. 4721863, 2019.

[68] H. Wang, A. Dragomir, N. I. Abbasi, J. Li, N. V. Thakor, and A. Bezerianos, "A novel real-time driving fatigue detection system based on wireless dry eeg," *Cognitive Neurodynamics*, vol. 12, no. 4, pp. 365–376, 2018.

[69] Venkatesh Balasubramanian and K. Adalarasu, "Emg-based analysis of change in muscle activity during simulated driving," *Journal of Bodywork and Movement Therapies*, vol. 11, no. 2, pp. 151–158, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1360859207000034

[70] M. Akin, M. B. Kurt, N. Sezgin, and M. Bayram, "Estimating vigilance level by using eeg and emg signals," *Neural Computing and Applications*, vol. 17, no. 3, pp. 227–236, 2008.

[71] R. Fu and H. Wang, "Detection of driving fatigue by using noncontact emg and ecg signals measurement system," *International journal of neural systems*, vol. 24, no. 3, p. 1450006, 2014.

[72] J. Miettinen, K. Nordhausen, and S. Taskinen, "fica: Fastica algorithms and their improved variants," *The R Journal*, vol. 10, no. 2, p. 148, 2019.

[73] F. J. Massey, "The kolmogorov-smirnov test for goodness of fit," *Journal of the American Statistical Association*, vol. 46, no. 253, p. 68, 1951.

[74] L. Wang, H. Wang, and X. Jiang, "A new method to detect driver fatigue based on emg and ecg collected by portable non-contact sensors," *PROMET*, vol. 29, no. 5, pp. 479–488, 2017.

[75] S. Johansson, C. Ytterberg, B. Back, L. W. Holmqvist, and L. von Koch, "The swedish occupational fatigue inventory in people with multiple sclerosis," *Journal of rehabilitation medicine*, vol. 40, no. 9, pp. 737–743, 2008.

[76] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[77] F. Guede-Fernández, M. Fernández-Chimeno, J. Ramos-Castro, and M. A. García-González, "Driver drowsiness detection based on respiratory signal analysis," *IEEE Access*, vol. 7, pp. 81 826–81 838, 2019.

[78] M.-N. Fiamma, Z. Samara, P. Baconnier, T. Similowski, and C. Straus, "Respiratory inductive plethysmography to assess respiratory variability and complexity in humans," *Respiratory physiology & neurobiology*, vol. 156, no. 2, pp. 234–239, 2007.

[79] S. E. H. Kiashari, A. Nahvi, H. Bakhoda, A. Homayounfard, and M. Tashakori, "Evaluation of driver drowsiness using respiration analysis by thermal imaging on a driving simulator," *Multimedia Tools and Applications*, vol. 79, no. 25, pp. 17 793–17 815, 2020.

[80] José Solaz, José Laparra-Hernández, Daniel Bande, Noelia Rodríguez, Sergio Veleff, José Gerpe, and Enrique Medina, "Drowsiness detection based on the analysis of breathing rate obtained from real-time image recognition," *Transportation Research Procedia*, vol. 14, pp. 3867–3876, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2352146516304793

[81] H. De Rosario, J.S. Solaz, N. Rodriguez, and L.M. Bergasa, "Controlled inducement and measurement of drowsiness in a driving simulator," *IET Intelligent Transport Systems*, vol. 4, no. 4, pp. 280–288(8), 2010. [Online]. Available: https://digital-library.theiet.org/content/journals/10.1049/iet-its.2009.0110

[82] C. Yang, X. Wang, and S. Mao, "Respiration monitoring with rfid in driving environments," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 2, pp. 500–512, 2021.

[83] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 1, pp. 208–220, 2013.

[84] Y. Qiu, G. Zhou, Y. Zhang, and A. Cichocki, "Canonical polyadic decomposition (cpd) of big tensors with low multilinear rank," *Multimedia Tools and Applications*, 2020.

[85] X. Fan, B.-C. Yin, and Y.-F. Sun, "Yawning detection for monitoring driver fatigue," in *International Conference on Machine Learning and Cybernetics, 2007.* Piscataway, NJ: IEEE Service Center, 2007, pp. 664–668.

[86] J. Miao, W. Gao, Y. Chen, and J. Lu, "Gravity-center template based human face feature detection," in *Advances in multimodal interfaces _ ICMI 2000*, ser. Lecture notes in computer science, 0302-9743, T. Tan, Y. Shi, and W. Gao, Eds. Berlin and London: Springer, 2000, vol. 1948, pp. 207–214.

[87] T. S. Lee, "Image representation using 2d gabor wavelets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 10, pp. 959–971, 1996.

[88] C. B. S. Maior, M. J. d. C. Moura, J. M. M. Santana, and I. D. Lins, "Real-time classification for autonomous drowsiness detection using eye aspect ratio," *Expert Syst. Appl.*, vol. 158, p. 113505, 2020.

[89] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Computer Vision and Pattern Recognition 2014,*, pp. 1867–1874. [Online]. Available: https://ieeexplore.ieee.org/document/6909637

[90] Q. Massoz, T. Langohr, C. François, 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, pp. 1–7. [Online]. Available: https://ieeexplore.ieee.org/document/7477715

[91] Q. Massoz, J. G. Verly, and M. van Droogenbroeck, "Multi-timescale drowsiness characterization based on a video of a driver's face," *Sensors (Basel, Switzerland)*, vol. 18, no. 9, 2018.

[92] W. Deng and R. Wu, "Real-time driver-drowsiness detection system using facial features," *IEEE Access*, vol. 7, pp. 118 727–118 738, 2019.

[93] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "High-speed tracking with kernelized correlation filters," *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 3, pp. 583–596, 2015.

[94] M. Kamel and L. Guan, "Histogram equalization utilizing spatial correlation for image enhancement," in *Visual Communications and Image Processing IV*, ser. SPIE Proceedings, W. A. Pearlman, Ed. SPIE, 1989, p. 712.

[95] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 477–482, 2000.

[96] Z. Zhang and J. Zhang, "A new real-time eye tracking based on nonlinear unscented kalman filter for monitoring driver fatigue," *Journal of Control Theory and Applications*, vol. 8, no. 2, pp. 181–188, 2010.

[97] C. P. Papageorgiou, M. Oren, and T. Poggio, "A general framework for object detection," in *Sixth International Conference on Computer Vision*, S. Chandran and U. B. Desai, Eds. New Dehli and London: Narosa, 1998, pp. 555–562.

[98] Y. Chen, X. Zhao, and X. Jia, "Spectral–spatial classification of hyperspectral data based on deep belief network," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 2381–2392, 2015.

[99] 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.

[100] P. Viola and M. J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.

[101] C. Kaufmann, M. Frühwirth, D. Messerschmidt, M. Moser, A. Eichberger, and S. Arefnezhad, "Driving and tiredness: Results of the behaviour observation of a simulator study with special focus on automated driving," *Transactions on Transport Sciences*, vol. 11, no. 2, pp. 51–63, 2020.

[102] S. Arefnezhad, A. Eichberger, M. Fruhwirth, C. Kaufmann, and M. Moser, "Driver drowsiness classification using data fusion of vehicle-based measures and ecg signals," in *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 10/11/2020 - 10/14/2020, pp. 451–456.

[103] S. Arefnezhad, S. Samiee, A. Eichberger, M. Frühwirth, C. Kaufmann, and E. Klotz, "Applying deep neural networks for multi-level classification of driver drowsiness using vehicle-based measures," *Expert Syst. Appl.*, vol. 162, p. 113778, 2020.

[104] S. Arefnezhad, S. Samiee, A. Eichberger, and A. Nahvi, "Driver drowsiness detection based on steering wheel data applying adaptive neuro-fuzzy feature selection," *Sensors (Basel, Switzerland)*, vol. 19, no. 4, 2019.

[105] SENSODRIVE GmbH, "Sensodrive gmbh - robotic und force-feedback," 2021. [Online]. Available: https://www.sensodrive.de/products/force-feedback-steering-wheels.php

[106] AVL, "Avl vsm™ vehicle simulation; release 2020 r1: Highlights of the latest release of our vehicle dynamics simulation tool," 2021. [Online]. Available: https://www.avl.com/-/avl-vsm-vehicle-simulation

[107] "World medical association declaration of helsinki: ethical principles for medical research involving human subjects," *JAMA*, vol. 310, no. 20, pp. 2191–2194, 2013. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/24141714/

[108] Victor Javier Kartsch, Simone Benatti, Pasquale Davide Schiavone, Davide Rossi, and Luca Benini, "A sensor fusion approach for drowsiness detection in wearable ultra-low-power systems," *Information Fusion*, vol. 43, pp. 66–76, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1566253517306942

[109] Shaibal Barua, Mobyen Uddin Ahmed, Christer Ahlström, and Shahina Begum, "Automatic driver sleepiness detection using eeg, eog and contextual information," *Expert Syst. Appl.*, vol. 115, pp. 121–135, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0957417418304792

[110] 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.

[111] Torbjörn Åkerstedt and Mats Gillberg, "Subjective and objective sleepiness in the active individual," *International Journal of Neuroscience*, vol. 52, no. 1-2, pp. 29–37, 1990.

[112] R. T. Wilkinson and D. Houghton, "Field test of arousal: A portable reaction timer with data storage," *Human Factors*, vol. 24, no. 4, pp. 487–493, 1982.

[113] Ahmed Deyab, "Development of an experimental set-up in a driving simulator study for driver drowsiness detection," Bachlor Thesiss, Graz University of Technology, Graz, July 2019.

[114] Smart Eye AB, "Smart eye pro 5.0 user manual," Göteborg, Sweden. [Online]. Available: https://cb3.unl.edu/dbrainlab/wp-content/uploads/sites/2/2013/12/Eye-Tracker-Manual_5.0.pdf

[115] g.tec medical engineering GmbH, "g.nautilus research wearable eeg headset," 2021. [Online]. Available: https://www.gtec.at/product/gnautilus-research/

[116] Joysys, "Chronocord 8-ultra hd hrf," Weiz, Austria, 2021. [Online]. Available: https://joysys.at/en/chronocord/

[117] K. B. Böcker, J. A. van Avermaete, and M. M. van den Berg-Lenssen, "The international 10-20 system revisited: cartesian and spherical co-ordinates," *Brain topography*, vol. 6, no. 3, pp. 231–235, 1994. [Online]. Available: https://doi.org/10.1007/BF01187714

[118] K. Khunti, "Accurate interpretation of the 12-lead ecg electrode placement: A systematic review," *Health Education Journal*, vol. 73, no. 5, pp. 610–623, 2014.

[119] W. Boucsein, D. C. Fowles, S. Grimnes, G. Ben-Shakhar, W. T. roth, M. E. Dawson, and D. L. Filion, "Publication recommendations for electrodermal measurements," *Psychophysiology*, vol. 49, no. 8, pp. 1017–1034, 2012. [Online]. Available: https://doi.org/10.1111/j.1469-8986.2012.01384.x

[120] "Alice 6 respiratory sensor nasal/oral (thermistor)," 5/16/2021. [Online]. Available: https://www.ternimed.de/Alice-6-respiratory-sensor-nasal/oral-Thermistor

[121] Anund, Anna AND Fors, Carina AND Hallvig, David AND Åkerstedt, Torbjörn AND Kecklund, Göran, "Observer rated sleepiness and real road driving: An explorative study," *PLOS ONE*, vol. 8, no. 5, pp. 1–8, 2013.

[122] Y. Wang, H. Yao, and S. Zhao, "Auto-encoder based dimensionality reduction," *Neurocomputing*, vol. 184, pp. 232–242, 2016.

[123] C. M. Wang and W. C. Xiao, "Second-order iir notch filter design and implementation of digital signal processing system," in *Proceedings of the 2nd International Symposium on Computer, Communication, Control and Automation*. Paris, France: Atlantis Press, 2/22/2013 - 4/24/2013.

[124] P. Gupta, K. K. Sharma, and S. D. Joshi, "Baseline wander removal of electrocardiogram signals using multivariate empirical mode decomposition," *Healthcare technology letters*, vol. 2, no. 6, pp. 164–166, 2015. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/26713161

[125] F. Scholkmann, J. Boss, and M. Wolf, "An efficient algorithm for automatic peak detection in noisy periodic and quasi-periodic signals," *Algorithms*, vol. 5, no. 4, pp. 588–603, 2012. [Online]. Available: https://www.mdpi.com/1999-4893/5/4/588

[126] T. Giannakopoulos and A. Pikrakis, "Audio features," in *Introduction to audio analysis*, T. Giannakopoulos and A. Pikrakis, Eds. Amsterdam: Academic Press, 2014, pp. 59–103. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780080993881000042

[127] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proceedings of the National Academy of Sciences*, vol. 88, no. 6, pp. 2297–2301, 1991.

[128] A. LESNE, "Shannon entropy: a rigorous notion at the crossroads between probability, information theory, dynamical systems and statistical physics," *Mathematical Structures in Computer Science*, vol. 24, no. 3, 2014.

[129] H. Helakari, J. Kananen, N. Huotari, L. Raitamaa, T. Tuovinen, V. Borchardt, A. Rasila, V. Raatikainen, T. Starck, T. Hautaniemi, T. Myllylä, O. Tervonen, S. Rytky, T. Keinänen, V. Korhonen, V. Kiviniemi, and H. Ansakorpi, "Spectral entropy indicates electrophysiological and hemodynamic changes in drug-resistant epilepsy - a multimodal mreg study," *NeuroImage: Clinical*, vol. 22, p. 101763, 2019.

[130] L. Gadenz, J. Hashemi, M. H. Shariat, L. Gula, and D. P. Redfearn, "Clinical role of dominant frequency measurements in atrial fibrillation ablation - a systematic review," *Journal of Atrial Fibrillation*, vol. 9, no. 6, p. 1548, 2017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5673341/

[131] King DJ, Mumford DK and Siegmund GP, Ed., *An algorithm for detecting heavy-truck driver fatigue from steering wheel motion*, 1998. [Online]. Available: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.619.5023&rep=rep1&type=pdf

[132] K. Fujiwara, E. Abe, K. Kamata, C. Nakayama, Y. Suzuki, T. Yamakawa, T. Hiraoka, M. Kano, Y. Sumi, F. Masuda, M. Matsuo, and H. Kadotani, "Heart rate variability-based driver drowsiness detection and its validation with eeg," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 6, pp. 1769–1778, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8520803

[133] Rosangela Akemi Hoshi, Carlos Marcelo Pastre, Luiz Carlos Marques Vanderlei, and Moacir Fernandes Godoy, "Poincaré plot indexes of heart rate variability: Relationships with other nonlinear variables," *Autonomic Neuroscience*, vol. 177, no. 2, pp. 271–274, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1566070213001264

[134] A. W. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct least squares fitting of ellipses," in *Proceedings of the 13th international conference on pattern recognition*. IEEE, 1996, pp. 253–257 vol.1.

[135] C.-H. Hsu, M.-Y. Tsai, G.-S. Huang, T.-C. Lin, K.-P. Chen, S.-T. Ho, L.-Y. Shyu, and C.-Y. Li, "Poincaré plot indexes of heart rate variability detect dynamic autonomic modulation during general anesthesia induction," *Acta anaesthesiologica Taiwanica : official journal of the Taiwan Society of Anesthesiologists*, vol. 50, no. 1, pp. 12–18, 2012.

[136] D. Morelli, A. Rossi, M. Cairo, and D. A. Clifton, "Analysis of the impact of interpolation methods of missing rr-intervals caused by motion artifacts on hrv features estimations," *Sensors (Basel, Switzerland)*, vol. 19, no. 14, 2019.

[137] F. Shaffer and J. P. Ginsberg, "An overview of heart rate variability metrics and norms," *Frontiers in public health*, vol. 5, p. 258, 2017.

[138] S. G. Aydin, T. Kaya, and H. Guler, "Heart rate variability (hrv) based feature extraction for congestive heart failure," *International Journal of Computer and Electrical Engineering*, vol. 8, no. 4, pp. 272–279, 2016.

[139] M. K. Moridani, Z. Mahabadi, and N. Javadi, "Heart rate variability features for different stress classification," *Bratislavske lekarske listy*, vol. 121, no. 9, pp. 619–627, 2020.

[140] Ko Keun Kim, Jung Soo Kim, Yong Gyu Lim, and Kwang Suk Park, "The effect of missing rr-interval data on heart rate variability analysis in the frequency domain," *Physiological Measurement*, vol. 30, no. 10, pp. 1039–1050, 2009.

[141] G. D. Clifford and L. Tarassenko, "Quantifying errors in spectral estimates of hrv due to beat replacement and resampling," *IEEE transactions on bio-medical engineering*, vol. 52, no. 4, pp. 630–638, 2005.

[142] R. Hanowski, D. Bowman, A. Alden, W. Wierwille, and R. Carroll, "Perclos+: Development of a robust field measure of driver drowsiness," 2008.

[143] Suhandi Junaedi and Habibullah Akbar, "Driver drowsiness detection based on face feature and perclos," *Journal of Physics: Conference Series*, vol. 1090, p. 012037, 2018.

[144] A. Celecia, K. Figueiredo, M. Vellasco, and R. González, "A portable fuzzy driver drowsiness estimation system," *Sensors (Basel, Switzerland)*, vol. 20, no. 15, 2020.

[145] Rattanawadee Panthong and Anongnart Srivihok, "Wrapper feature subset selection for dimension reduction based on ensemble learning algorithm," *Procedia Computer Science*, vol. 72, pp. 162–169, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877050915035784

[146] Girish Chandrashekar and Ferat Sahin, "A survey on feature selection methods," *Computers & Electrical Engineering*, vol. 40, no. 1, pp. 16–28, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0045790613003066

[147] N. S. Malan and S. Sharma, "Feature selection using regularized neighbourhood component analysis to enhance the classification performance of motor imagery signals," *Computers in Biology and Medicine*, vol. 107, pp. 118–126, 2019.

[148] W. Yang, K. Wang, and W. Zuo, "Neighborhood component feature selection for high-dimensional data," *Journal of Computers*, vol. 7, no. 1, 2012.

[149] J. Goldberger, S. Roweis, G. Hinton, and R. Salakhutdinov, "Neighbourhood components analysis," in *Proceedings of the 17th International Conference on Neural Information Processing Systems*, ser. NIPS'04. Cambridge, MA, USA: MIT Press, 2004, pp. 513–520.

[150] Wei Yang, Kuanquan Wang, and W. Zuo, "Neighborhood component feature selection for high-dimensional data," *J. Comput.*, vol. 7, pp. 161–168, 2012.

[151] R. Blagus and L. Lusa, "Smote for high-dimensional class-imbalanced data," *BMC Bioinformatics*, vol. 14, no. 1, p. 106, 2013.

[152] H. He, Y. Bai, E. A. Garcia, and S. Li, "Adasyn: Adaptive synthetic sampling approach for imbalanced learning," in *Neural Networks, 2008, IJCNN 2008, (IEEE World Congress on Computational Intelligence), IEEE International Joint Conference on.* [Piscataway, N.J.]: IEEE, 2008, pp. 1322–1328.

[153] A. Anand, G. Pugalenthi, G. B. Fogel, and P. N. Suganthan, "An approach for classification of highly imbalanced data using weighting and undersampling," *Amino Acids*, vol. 39, no. 5, pp. 1385–1391, 2010.

[154] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician*, vol. 46, no. 3, pp. 175–185, 1992.

[155] M. Hasan and J.-M. Kim, "Fault detection of a spherical tank using a genetic algorithm-based hybrid feature pool and k-nearest neighbor algorithm," *Energies*, vol. 12, no. 6, p. 991, 2019. [Online]. Available: https://doi.org/10.3390/en12060991

[156] H. A. Abu Alfeilat, A. B. A. Hassanat, O. Lasassmeh, A. S. Tarawneh, M. B. Alhasanat, H. S. Eyal Salman, and V. B. S. Prasath, "Effects of distance measure choice on k-nearest neighbor classifier performance: A review," *Big data*, vol. 7, no. 4, pp. 221–248, 2019.

[157] Tin Kam Ho, "Random decision forests," in *Proceedings of 3rd International Conference on Document Analysis and Recognition*, vol. 1, 1995, pp. 278–282 vol.1.

[158] C. Seiffert, T. M. Khoshgoftaar, J. van Hulse, and A. Napolitano, "Rusboost: Improving classification performance when training data is skewed," in *19th International Conference on Pattern Recognition (ICPR 2008)*. New York: IEEE Press Books, 2008, pp. 1–4.

[159] K. Fujiwara, Y. Huang, K. Hori, K. Nishioji, M. Kobayashi, M. Kamaguchi, and M. Kano, "Over- and under-sampling approach for extremely imbalanced and small minority data problem in health record analysis," *Frontiers in public health*, vol. 8, p. 178, 2020.

[160] T. Chengsheng, L. Huacheng, and X. Bing, "Adaboost typical algorithm and its application research," *MATEC Web of Conferences*, vol. 139, p. 00222, 2017.

[161] S. I. Dimitriadis and D. Liparas, "How random is the random forest? random forest algorithm on the service of structural imaging biomarkers for alzheimer's disease: from alzheimer's disease neuroimaging initiative (adni) database," *Neural regeneration research*, vol. 13, no. 6, pp. 962–970, 2018.

[162] H. M and S. M.N, "A review on evaluation metrics for data classification evaluations," *International Journal of Data Mining & Knowledge Management Process*, vol. 5, no. 2, pp. 01–11, 2015.

[163] S. Samiee, S. Azadi, R. Kazemi, A. Nahvi, and A. Eichberger, "Data fusion to develop a driver drowsiness detection system with robustness to signal loss," *Sensors (Basel, Switzerland)*, vol. 14, no. 9, pp. 17 832–17 847, 2014.

[164] Z. Li, L. Chen, J. Peng, and Y. Wu, "Automatic detection of driver fatigue using driving operation information for transportation safety," *Sensors (Basel, Switzerland)*, vol. 17, no. 6, 2017.

[165] A. Moujahid, F. Dornaika, I. Arganda-Carreras, and J. Reta, "Efficient and compact face descriptor for driver drowsiness detection," *Expert Syst. Appl.*, vol. 168, p. 114334, 2021.

[166] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: Synthetic minority over-sampling technique," *J. Artif. Int. Res.*, vol. 16, no. 1, pp. 321–357, 2002.

[167] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[168] F. J. Ordóñez and D. Roggen, "Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition," *Sensors (Basel, Switzerland)*, vol. 16, no. 1, 2016. [Online]. Available: https://www.mdpi.com/1424-8220/16/1/115

[169] Jinjiang Wang, Yulin Ma, Laibin Zhang, Robert X. Gao, and Dazhong Wu, "Deep learning for smart manufacturing: Methods and applications," *Journal of Manufacturing Systems*, vol. 48, pp. 144–156, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0278612518300037

[170] R. Jabbar, K. Al-Khalifa, M. Kharbeche, W. Alhajyaseen, M. Jafari, and S. Jiang, "Real-time driver drowsiness detection for android application using deep neural networks techniques," *Procedia Computer Science*, vol. 130, pp. 400–407, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877050918304137

[171] In-Ho Choi, Sung Kyung Hong, and Yong-Guk Kim, "Real-time categorization of driver's gaze zone using the deep learning techniques," in *2016 International Conference on Big Data and Smart Computing (BigComp)*, 2016, pp. 143–148.

[172] Arnaud Sors, Stéphane Bonnet, Sébastien Mirek, Laurent Vercueil, and Jean-François Payen, "A convolutional neural network for sleep stage scoring from raw single-channel eeg," *Biomedical Signal Processing and Control*, vol. 42, pp. 107–114, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1746809417302847

[173] Muammer Turkoglu, Omer F. Alcin, Muzaffer Aslan, Adel Al-Zebari, and Abdulkadir Sengur, "Deep rhythm and long short term memory-based drowsiness detection," *Biomedical Signal Processing and Control*, vol. 65, p. 102364, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1746809420304729

[174] M. Jogin, Mohana, M. S. Madhulika, G. D. Divya, R. K. Meghana, and S. Apoorva, "Feature extraction using convolution neural networks (cnn) and deep learning," in *2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*. [Piscataway, NJ]: IEEE, 2018, pp. 2319–2323.

[175] Q. Li, W. Cai, X. Wang, Y. Zhou, D. D. Feng, and M. Chen, "Medical image classification with convolutional neural network," in *2014 13th International Conference on Control Automation Robotics Vision (ICARCV)*, 2014, pp. 844–848.

[176] M. Mario, "Human activity recognition based on single sensor square hv acceleration images and convolutional neural networks," *IEEE Sensors Journal*, vol. 19, no. 4, pp. 1487–1498, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8543606

[177] S. Park, Y. Jeong, and H. S. Kim, "Multiresolution cnn for reverberant speech recognition," in *2017 20th Conference of the Oriental Chapter of the International Coordinating Committee on Speech Databases and Speech I/O Systems and Assessment (O-COCOSDA)*, 2017, pp. 1–4.

[178] Y. Wang, Y. Li, Y. Song, and X. Rong, "The influence of the activation function in a convolution neural network model of facial expression recognition," *Applied Sciences*, vol. 10, no. 5, p. 1897, 2020.

[179] Jinjiang Wang, Yulin Ma, Laibin Zhang, Robert X. Gao, and Dazhong Wu, "Deep learning for smart manufacturing: Methods and applications," *Journal of Manufacturing Systems*, vol. 48, pp. 144–156, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0278612518300037

[180] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proceedings of the 32nd International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, F. Bach and D. Blei, Eds., vol. 37. Lille, France: PMLR, 2015, pp. 448–456. [Online]. Available: http://proceedings.mlr.press/v37/ioffe15.html

[181] Y. Huang and Y. Yu, Eds., *An Internal Covariate Shift Bounding Algorithm for Deep Neural Networks by Unitizing Layers' Outputs*, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9156986

[182] Mallat Stéphane, Ed., *A Wavelet Tour of Signal Processing (Third Edition)*, third edition ed. Boston: Academic Press, 2009. [Online]. Available: https://www.sciencedirect.com/book/9780123743701/a-wavelet-tour-of-signal-processing

[183] H.-y. Feng, J.-p. Wang, Y.-c. Li, and J. Chen, "Wavelet theory and application summarizing," in *Information Computing and Applications*, B. Liu and C. Chai, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 337–343. [Online]. Available: https://doi.org/10.1007/978-3-642-25255-6_43

[184] E. Sejdic, I. Djurovic, and L. Stankovic, "Quantitative performance analysis of scalogram as instantaneous frequency estimator," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3837–3845, 2008.

[185] J. M. Lilly and S. C. Olhede, "Higher-order properties of analytic wavelets," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 146–160, 2009. [Online]. Available: https://ieeexplore.ieee.org/document/4663912

[186] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014. [Online]. Available: https://jmlr.org/papers/v15/srivastava14a.html

[187] I. Ilievski, T. Akhtar, J. Feng, and C. Shoemaker, "Efficient hyperparameter optimization for deep learning algorithms using deterministic rbf surrogates," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, no. 1, 2017. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/10647

[188] Jia Wu, Xiu-Yun Chen, Hao Zhang, Li-Dong Xiong, Hang Lei, and Si-Hao Deng, "Hyperparameter optimization for machine learning models based on bayesian optimizationb," *Journal of Electronic Science and Technology*, vol. 17, no. 1, pp. 26–40, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1674862X19300047

[189] J. Snoek, H. Larochelle, and R. P. Adams, "Practical bayesian optimization of machine learning algorithms," in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 2*, ser. NIPS'12. Red Hook, NY, USA: Curran Associates Inc, 2012, pp. 2951–2959.

[190] K. You, M. Long, J. Wang, and M. I. Jordan, "How does learning rate decay help modern neural networks?" [Online]. Available: https://arxiv.org/pdf/1908.01878

[191] J. Kukačka, V. Golkov, and D. Cremers, "Regularization for deep learning: A taxonomy." [Online]. Available: https://arxiv.org/pdf/1710.10686

[192] Diederik P. Kingma and Jimmy Ba, "Adam: A method for stochastic optimization," *CoRR*, vol. abs/1412.6980, 2015.

[193] D. Masters and C. Luschi, "Revisiting small batch training for deep neural networks." [Online]. Available: https://arxiv.org/pdf/1804.07612

[194] X. Yu, Z. Zhang, L. Wu, W. Pang, H. Chen, Z. Yu, B. Li, and S. Jagannathan, "Deep ensemble learning for human action recognition in still images," *Complexity*, vol. 2020, p. 9428612, 2020.

[195] S. W. A. Sherazi, J.-W. Bae, and J. Y. Lee, "A soft voting ensemble classifier for early prediction and diagnosis of occurrences of major adverse cardiovascular events for stemi and nstemi during 2-year follow-up in patients with acute coronary syndrome," *PLOS ONE*, vol. 16, no. 6, 2021. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0249338

[196] Q. Yin, R. Zhang, and X. Shao, "Cnn and rnn mixed model for image classification," *MATEC Web of Conferences*, vol. 277, p. 02001, 2019.

[197] J. Zhu, H. Chen, and W. Ye, "A hybrid cnn–lstm network for the classification of human activities based on micro-doppler radar," *IEEE Access*, vol. 8, pp. 24 713–24 720, 2020.

[198] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, Eds., *A Comprehensive Survey on Transfer Learning*, vol. 109. IEEE, 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9134370

[199] Z. Mardi, S. N. M. Ashtiani, and M. Mikaili, "Eeg-based drowsiness detection for safe driving using chaotic features and statistical tests," *Journal of medical signals and sensors*, vol. 1, no. 2, pp. 130–137, 2011. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3342623/

[200] Venkata Phanikrishna B and Suchismitha Chinara, "Automatic classification methods for detecting drowsiness using wavelet packet transform extracted time-domain features from single-channel eeg signal," *Journal of Neuroscience Methods*, vol. 347, p. 108927, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165027020303502

[201] U. Budak, V. Bajaj, Y. Akbulut, O. Atila, and A. Sengur, "An effective hybrid model for eeg-based drowsiness detection," *IEEE Sensors Journal*, vol. 19, no. 17, pp. 7624–7631, 2019.

[202] G. Li and W.-Y. Chung, "A context-aware eeg headset system for early detection of driver drowsiness," *Sensors (Basel, Switzerland)*, vol. 15, no. 8, pp. 20 873–20 893, 2015.

[203] A. D. McDonald, C. Schwarz, J. D. Lee, and T. L. Brown, "Real-time detection of drowsiness related lane departures using steering wheel angle," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 56, no. 1, pp. 2201–2205, 2012.

[204] D. Sommer and M. Golz, "Evaluation of perclos based current fatigue monitoring technologies," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, vol. 2010, pp. 4456–4459, 2010.

[205] Jacek P. Dmochowski, Jason J. Ki, Paul DeGuzman, Paul Sajda, and Lucas C. Parra, "Extracting multidimensional stimulus-response correlations using hybrid encoding-decoding of neural activity," *NeuroImage*, vol. 180, pp. 134–146, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811917304299

[206] T. Naselaris, R. J. Prenger, K. N. Kay, M. Oliver, and J. L. Gallant, "Bayesian reconstruction of natural images from human brain activity," *Neuron*, vol. 63, no. 6, pp. 902–915, 2009.

[207] A. Yousefi, I. Basu, A. C. Paulk, N. Peled, E. N. Eskandar, D. D. Dougherty, S. S. Cash, A. S. Widge, and U. T. Eden, "Decoding hidden cognitive states from behavior and physiology using a bayesian approach," *Neural computation*, vol. 31, no. 9, pp. 1751–1788, 2019.

[208] R. Vigario, J. Sarela, V. Jousmiki, M. Hamalainen, and E. Oja, "Independent component approach to the analysis of eeg and meg recordings," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 5, pp. 589–593, 2000.

[209] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural Networks*, vol. 13, no. 4, pp. 411–430, 2000. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0893608000000265

[210] A. Delorme and S. Makeig, "Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, no. 1, pp. 9–21, 2004.

[211] L. Pion-Tonachini, K. Kreutz-Delgado, and S. Makeig, "Iclabel: An automated electroencephalographic independent component classifier, dataset, and website," *NeuroImage*, vol. 198, pp. 181–197, 2019.

[212] S. Särkkä, *Bayesian filtering and smoothing*, ser. Institute of mathematical statistics textbooks. Cambridge, United Kingdom: Cambridge University Press, 2013, vol. 3.

[213] N. G. van KAMPEN, "Chapter iv - markov processes," in *Stochastic Processes in Physics and Chemistry (Third Edition)*, ser. North-Holland Personal Library, N. G. van KAMPEN, Ed. Amsterdam: Elsevier, 2007, pp. 73–95. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780444529657500076

[214] H. Fang, N. Tian, Y. Wang, M. Zhou, and M. A. Haile, "Nonlinear bayesian estimation: from kalman filtering to a broader horizon," *IEEE/CAA Journal of Automatica Sinica*, vol. 5, no. 2, pp. 401–417, 2018.

[215] G. James, D. Witten, T. Hastie, and R. Tibshirani, "Linear regression," in *An introduction to statistical learning*, ser. Springer Texts in Statistics, G. James, D. Witten, T. Hastie, and R. Tibshirani, Eds. New York: Springer, 2013, vol. 103, pp. 59–126. [Online]. Available: https://link.springer.com/chapter/10.1007/978-1-4614-7138-7_3

[216] J. Chen, H. Wang, and C. Hua, "Assessment of driver drowsiness using electroencephalogram signals based on multiple functional brain networks," *International journal of psychophysiology : official journal of the International Organization of Psychophysiology*, vol. 133, pp. 120–130, 2018.

[217] R. C. Srivastava, "Estimation of probability density function based on random number of observations with applications," *International Statistical Review / Revue Internationale de Statistique*, vol. 41, no. 1, p. 77, 1973.

[218] Ibtissem Belakhdar, Walid Kaaniche, Ridha Djemal, and Bouraoui Ouni, "Single-channel-based automatic drowsiness detection architecture with a reduced number of eeg features," *Microprocessors and Microsystems*, vol. 58, pp. 13–23, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0141933117303903

[219] S. Yu, P. Li, H. Lin, E. Rohani, G. Choi, B. Shao, and Q. Wang, "Support vector machine based detection of drowsiness using minimum eeg features," in *2013 International Conference on Social Computing*, 2013, pp. 827–835.

[220] D. Dinges and R. Grace, "Perclos: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance," 1998.

[221] C. R. Holdgraf, J. W. Rieger, C. Micheli, S. Martin, R. T. Knight, and F. E. Theunissen, "Encoding and decoding models in cognitive electrophysiology," *Frontiers in systems neuroscience*, vol. 11, p. 61, 2017.

[222] J. Hu, "Comparison of different features and classifiers for driver fatigue detection based on a single eeg channel," *Computational and mathematical methods in medicine*, vol. 2017, p. 5109530, 2017.

[223] H. Zeng, C. Yang, G. Dai, F. Qin, J. Zhang, and W. Kong, "Eeg classification of driver mental states by deep learning," *Cognitive Neurodynamics*, vol. 12, no. 6, pp. 597–606, 2018.

[224] R. Chai, S. H. Ling, P. P. San, G. R. Naik, T. N. Nguyen, Y. Tran, A. Craig, and H. T. Nguyen, "Improving eeg-based driver fatigue classification using sparse-deep belief networks," *Frontiers in Neuroscience*, vol. 11, p. 103, 2017. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnins.2017.00103

[225] J. Cabral, M. L. Kringelbach, and G. Deco, "Exploring the network dynamics underlying brain activity during rest," *Progress in neurobiology*, vol. 114, pp. 102–131, 2014.

[226] R. Salvador, J. Suckling, C. Schwarzbauer, and E. Bullmore, "Undirected graphs of frequency-dependent functional connectivity in whole brain networks," *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, vol. 360, no. 1457, pp. 937–946, 2005.

[227] S. M. Bowyer, "Coherence a measure of the brain networks: past and present," *Neuropsychiatric Electrophysiology*, vol. 2, no. 1, 2016.

[228] M. A. Guevara and M. Corsi-Cabrera, "Eeg coherence or eeg correlation?" *International journal of psychophysiology : official journal of the International Organization of Psychophysiology*, vol. 23, no. 3, pp. 145–153, 1996.

List of Figures

List of Tables

Monographic Series TU Graz

Reihe Fahrzeugtechnik / Automotive Series

Vol. 1 Haymo Niederkofler

Analyse radselektiv eingreifender Fahrdynamikregelsysteme für die Anwendung in Elektromechanischen Corner-Modulen
2012
ISBN 978-3-85125-220-0

Vol. 2 Andrés Eduardo Rojas Rojas

Passenger Vehicles with In-Wheel Motors
2012
ISBN 978-3-85125-234-7

Vol. 3 Daniel Wallner

Experimental and Numerical Investigations on Brake Squeal
2013
ISBN 978-3-85125-269-9

Vol. 4 Alexander Harrich

CAD-basierte Methoden zur Unterstützung der Karosseriekonstruktion in der Konzeptphase
2015
ISBN 978-3-85125- 420-4

Vol. 5 Cornelia Lex

Maximum Tire-Road Friction Coefficient Estimation
2015
ISBN 978-3-85125-422-8

Vol. 6 Stefan Bernsteiner

Integration of Advanced Driver Assistance Systems on Full-Vehicle Level
2016
ISBN 978-3-85125-468-6

Monographic Series TU Graz

Reihe Fahrzeugtechnik / Automotive Series

Vol. 7 Patrick Rossbacher

Beitrag integrierter Konzeptmodelle zur virtuellen Gesamtfahrzeugentwicklung in der frühen Entwurfsphase
2020
ISBN 978-3-85125-730-4

Vol. 8 Andreas Hackl

Enhanced Tyre Modelling for Vehicle Dynamics Control Systems
2020
ISBN 978-3-85125-728-1

Vol. 9 Christoph Feichtinger

Racebike Dynamics
2021
ISBN 978-3-85125-860-8

Vol. 10 Sadegh Arefnezhad

Evaluation of Driver Performance in Semi-automated Driving by Physiologic, Driver Behaviour and Video-based Sensors
2022
ISBN 978-3-85125-862-2