Characterization and Identification of Distraction During Naturalistic Driving Using Wearable Non-Intrusive Physiological Measure of Galvanic Skin Responses

by

Vikas Rajendra

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science
(Computer and Information Science)
in the University of Michigan-Dearborn
2018

Master's Thesis Committee:

Assistant Professor Omid Dehzangi, Chair
Professor Bruce Maxim
Assistant Professor Anys Bacha

Table of Contents

List of Tables

List of Figures

Abstract

Fatalities due to road accidents are mainly caused by distracted driving. Driving demands continuous attention of the driver. Certain levels of distraction while driving can cause the driver to lose his/her attention which might lead to a fatal accident. Thus, early detection of distraction will help reduce the number of accidents. Several researches have been conducted for automatic detection of driver distraction. Many previous approaches have employed camera-based techniques. However, these methods might detect the distraction rather late to warn the drivers. Although neurophysiological signals using Electroencephalography (EEG) have shown to be another reliable indicator of distraction, EEG signals are very complex, and the technology is intrusive to the drivers, which creates serious doubt for its implementation. In this thesis we investigate a non-intrusive physiological measure-Galvanic Skin Responses (GSR) using a wrist band wearable and conduct an empirical characterization of driver GSR signals during a naturalistic driving experiment. The proposed method is used to evaluate and extract statistical, frequency and time domain features to identify distraction. Also, several data mining techniques such as feature selection, feature-ranking, dimensionality reduction and feature space analysis are performed to generate discriminative bases that reduce the computational complexity for efficient identification of distraction using supervised learning. A signal processing technique: continuous decomposition analysis, exclusive for skin conductance signal was investigated to better understand the behavior of raw signal during cognitive and visual over load from secondary tasks while driving. The proposed driver monitoring and identification system on the edge provided evident results using GSR as a reliable indicator of driver distraction while meeting the requirement of early notification of distraction state to driver.

Chapter 1 Introduction

Distracted driving is one of the major reasons leading to fatalities on road. The most recent report by National Highway Traffic Safety Administration (NHTSA) published in 2017 reports 37461 people (5.6% increase compared to 2015) lost lives to fatal road crashes in 2016 (NHTSA 2013). Although safety improvement in vehicle and on roads were implemented the count of fatalities continues to increase. Accidents on road that causes fatalities are predicted to be the ninth leading cause of death globally and are estimated to become seventh leading cause of death by 2030 (World Health Organization 2015). This increase in fatalities on road exemplifies the need to address this important issue. Preventing fatal accidents on road will help save millions of lives across the world.

Most of the fatalities occur due to diminished driving performance caused by impairment of driver's attention. Attentional impairment includes attributes like alcohol, drowsiness, fatigue, aging and distraction. More than 30% of the crashes are a result of alcohol impairment (NHTSA 2013). Drowsy driving was responsible for 2.3-2.5% of fatalities on road as per report published by NHTSA. Distracted driving compared to the above impairments, is one of the important form as it contributes majority of the crashes and fatalities on road. NHTSA reports 3477 people were killed in 2015 alone and 391,000 were injured in motor vehicle crashes due to distracted driving. Distracted driving has become a major concern in the modern era of innovative technology and has drawn attention from human behavior researchers and policy makers in the field of transportation safety (Box 2009). Distraction while driving is primarily caused by performing secondary tasks such as having engaging conversation with the passengers, eating, drinking, using cell phones for making and receiving calls or text messages, or using one of the several in-vehicle information systems (IVISs) namely navigation system and radio (Regan et al. 2008). The increasing use of several IVISs and technologies like phone is becoming critical as they induce distraction in drivers. Assisting the drivers to make use of the IVISs without compromising on the safety is an important prevailing challenge.

Distraction can be categorized into two main groups visual distraction and cognitive distraction. Visual distraction happens when the driver has his/her "eyes-off-road" and cognitive

distraction corresponds to “mind-off” the primary task of driving (Victor 2005). Both these distraction/inattention state result is affecting the driving performance of the driver by reducing the response time during hazards, large variation in lane, less effective visual scanning of the road. Moreover, these distractions can occur in combination with each other causing a higher order of distraction which creates major concerns to detect them in time.

Among several secondary tasks that diverts drivers’ attention form the primary task of driving, it is observed phone conversation, or using the phone to send and receive text messages induces higher order of distraction. Much importance is given to aforesaid scenarios since they include both visual and cognitive distraction that diverts the attention of the driver from task of driving completely (Liang et al. 2007). Developing a warning system that detects and mitigates the inattention state of the driver in early stages of distraction is one way to solve this problem (Donmez et al. 2003). This created a need for detecting the distraction accurately and warning the driver early in real-time.

1.1 Need for physiological signal monitoring

Developing an early inattention state warning mitigation system is a critical challenge. There have been several researches conducted to detect distraction using modalities recorded based on tracking driver’s eye gaze (Liu et al. 2015), eye lid movement (Metz et al. 2011), lane tracking (Young et al. 2011) and video cameras that captures behavior of the driver to detect distraction (Wege et al. 2013). Although these techniques demonstrated successful performance they suffer from major drawbacks such as invading privacy and mainly identifying and responding of distraction rather late only after it is visually noticeable. The above methods mainly focus on identifying visual distraction alone. These modalities lack identifying or detecting cognitive distraction as the above methods cannot record the cognitive behavior of the driver. Thus, to overcome these limitations there is a need for quick and reliable distraction detection modalities, which is achieved by continuous monitoring of physiological signals.

1.2 Introduction to physiological Signals

Measurements generated due to physiological process of human being such as heart-beat rate (electrocardiogram or ECG), respiratory rate, muscle current (electromyography or EMG), brain activity (electroencephalogram or EEG) or skin conductance (electrodermal activity (EDA) or galvanic skin responses (GSR)) are generally classified as physiological signals. In the recent years

physiological signals have become conventional measures in driver distraction experiments. Physiological signals have been used in several studies to identify level of mental load, mental stress and drowsiness during driving (Begum 2013).

The authors (Alizadeh & Dehzangi 2016) used EEG dynamics captured during a naturalistic driving experiment to developed a systematic framework that extracts effective descriptors and measures the impact of in-vehicle secondary task on the cognitive state of the driver during driving. They used several classification methods to detect distraction. This study demonstrated high detection accuracy for distraction. The studies conducted by authors (Lin et al. 2011; Kim et al. 2013) also demonstrated state-of-the-art results that are comprehensive and reliable. However, complexity to setup and analyzing data is the major limitation of the above-mentioned studies making those systems expensive and intrusive to implement.

In the study (Deshmukh & Dehzangi 2017) ECG modality was used to predict drivers distraction. Wavelet packet transform was used to localize the impact of distracting element and perform frequency sub-band analysis. Linear discriminant analysis (LDA) was used to reduce the generated WPT feature space. Kernel transformations were also employed to address the non-linear nature of the input feature space. Kernel transformation demonstrated best identification accuracy. Although ECG is comparatively easy to record than EEG and a quick reliable indicator of distractions, it faces similar issues of intrusive implementation as EEG.

To overcome the limitations of EEG and ECG modality, in this thesis we investigate skin conductance (SC) also known as galvanic skin responses (GSR) an electrical conductance of skin, is low cost, robust physiological signal which is a minimally intrusive modality that can be recorded on wrist and fingers easily (Ciabattoni et al. n.d.; Rajendra & Dehzangi 2017; Dehzangi et al. 2018). The change in electrical property of skin caused by interaction between environmental events and physiological state of a person is known as galvanic skin response. GSR is one of the most sensitive physiological indicator of cognitive and emotional state. Emotional stimulation triggers skin conductance mainly in the hand and foot region. As the skin of a human is good conductor of electricity, when small amount of weak electric current is passed through skin, changes in skin conduction can be calculated. Higher value of skin conductance corresponds to higher arousal. Measuring skin conductivity can reveal changes in sympathetic nervous system which drives human behavior, cognitive and emotional state on a subconscious level (Nourbakhsh et al. 2012; Nourbakhsh et al. 2013).

1.3 Literature Review

To inspect the impact of cognitive state change, synchronously recorded GSR signals were investigated in several experiments. In study (Nourbakhsh et al. 2012), the authors used GSR as an indicator of mental and emotional stress to evaluate users' stress due to workload while performing reading and arithmetic task. Eight arithmetic tasks with four difficulty level were performed by the experimental subjects, similarly the reading tasks included silent reading, where each task comprised of four slides of text and each slide was displayed for 30 seconds with three levels of difficulty. Time domain and frequency domain features were explored and concluded that frequency domain features showed to be promising in measuring the cognitive workload compared to the time domain features.

In paper (Kurniawan et al. 2013) authors investigate how classification techniques can be used to automatically determine periods of acute stress relying on information contained in GSR and/or speech of a person. The experimental setup was designed to record information while the subjects were performing three levels of task where the difficultly level increased incrementally. The subjects were made to perform Stroop-Word congruent color test and few basic arithmetic tasks corresponding to the level of difficulty. Features such as mean, minimum, maximum and standard deviation of skin conductance and peak height; and the total number and the cumulative amplitude, rising time and energy of startle responses in segment were extracted from GSR signal. Four classifiers: K-means, GMM, SVM and decision tree were employed for performing identification task. SVM demonstrated 70% accuracy for detecting stress. Finally, the study concluded that using other measures and not relying only on GSR for stress detection might improve the performance of their classification model.

In the previous work (Lew et al. 2008), a novel method for analyzing skin conductance (SC) and pupillometry time-series data using Short Time Fourier Transform (STFT) was employed to extract estimation of mental work load with high enough temporal bandwidth to be useful for augmented cognition application. Measures such as SC, pupil dilation, blink rate, and visual scanning pattern were recorded while the participants were engaged in performing process control task based on DURESS simulation. Participants were to diagnose errors introduced in the form of fault events and perform control adjustments to maintain the operating simulator in the target range. The authors explored STFT of the extracted measures to check whether the STFTs produced visible effects demonstrating change or increase in the mental workload and stress associated.

Graphical data analysis of the STFT showed notable increase in the power spectrum across a range of frequencies directly following fault events.

(Ayata et al. 2017) used time domain, wavelet and empirical mode decomposition (EMD) based features extracted from galvanic skin responses (GSR) to perform emotion recognition. They categorized valence and arousal and used several machine learning algorithms such as random forest, decision tree, k-nearest neighbor and support vector machine to study relationship between physiological signals and arousal and valence. This study demonstrated that there is a relationship existing between GSR signal, arousal and valence. Prediction accuracy of 71.53% and 71.04% were achieved for predicting arousal and valence respectively using time domain features only. Employing EMD yielded increase in classification performance compared to time-domain only features. It demonstrated prediction accuracy of 85.07% and 82.81% for arousal and valence respectively.

The previous work (Liu et al. 2016) proposed a novel human emotion recognition technique that selects GSR features automatically. Wavelet function was employed to de-noise the data and data was normalized to eliminate individual difference. They also employed covariance-based feature selection on the 30 generated features to improve the recognition performance. The covariance method selected 15 features that contains the most discriminative information to accurately represent the GSR signal. Support vector machine (SVM) was used to classify human emotion and the radial basis function (RBF) kernel SVM demonstrated accuracy of 66.67%.

The above investigations provide evident proof to use skin conductance signal to identify human emotion. However, very few investigations are conducted to evaluate cognitive workload or distraction during naturalistic driving experiment. (Min et al. 2013) investigated the effect of distraction task, such as using navigation system and sending text message with cellphone, on the driving performance of the driver using galvanic skin responses as the modality. The participants performed three driving tasks namely: driving only, driving + sending text message and driving + using navigation system, where each task was 2 min long. Then, relative change in skin conductance level was calculated and comparative analysis was performed for three tasks. The study resulted in the tasks driving + sending text message and driving + using navigation system demonstrating significant increase in relative skin conductance. However, this work did not include any significant quick and reliable distraction detection technique and the experiment was conducted using a driving simulator which is another major limitation of this study.

The authors in (Chen et al. 2017) used several physiological signals like electrocardiogram (ECG), galvanic skin response (GSR) and respiration to identify distraction while driving on the road. Spectral, time and wavelet multi-domain features were extracted for 10 second interval of data. Optimal feature sets were selected by combining Sparse Bayesian Learning (SBL) and Principal Component Analysis (PCA). Kernel based classifiers were employed to identify stress while driving. Experimental results of this study reveal that employing physiological measures to in-vehicle intelligent system to assist drivers on road for identifying and alerting stress early. In this study raw GSR signal was used with combination of other physiological measures and only time domain features were explored.

1.4 Thesis and Contribution

In this thesis, driver monitoring and intervention system on the edge using a non-invasive wrist band wearable galvanic skin response is proposed. To meet the requirement of quick and reliable distraction detection and mitigation system short-term segments were employed to detect the impact of secondary tasks of calling and texting on drivers during naturalistic driving. This thesis consists of three main published research works.

The research work (Rajendra & Dehzangi 2017) was a preliminary analysis of driver distraction during naturalistic driving experiment using raw galvanic skin response signal. This work solely focused on two scenarios (i) normal driving (non-distracted state) and (ii) driving while having an engaging phone conversation (distracted state). Few standard statistical measures were generated and binary support vector machine (SVM) was used for distraction detection. The main aim was to investigate the discriminative power in the raw GSR space between normal and distracted driver state. An average detection accuracy of 91% using binary SVM was achieved.

The research work (Dehzangi Omid, Rajendra 2017) investigated Galvanic Skin Responses (GSR) using a wrist band wearable to conduct an empirical characterization of driver GSR signals during a naturalistic driving experiment. Three scenarios were investigated: 1) **Normal** driving focusing attention on the primary task of driving 2) **Phone Distracted** driving while having an engaging phone conversation and 3) **Text distracted** driving while writing and sending texts while driving. Here phone distraction is investigated as cognitive distraction element in comparison to text distraction, which represent cognitive and visual distraction at the same time. Time and frequency domain were explored to extract relevant features to capture the changes/patterns at the physiological level. Due to the fact that feature extraction is a manual process and to normalize

the feature space toward the identification task, the feature space was transformed using linear discriminant dimensionality reduction to discover discriminative bases of the GSR multivariate feature space that identify distraction. That would eliminate both the computational complexity and the redundancies in the manually generated feature space. Due to multi-class nature of the identification task, there might be biases between the distracted and non-distracted categories that can bias the estimation of between- and within-class scatter matrices. Therefore, a class dependent weight was incorporated to calculate the scatter matrices. The proposed weight aims to increase the flexibility of the discriminative bases vectors to capture the factors that focus on eliminating the overlap between distracted vs. non-distracted in the generalization phase. The experimental results demonstrated high cross validation accuracies of distraction identification using raw GSR signals (i.e. 85.19%). Conducting dimensionality reduction using LDA resulted in slight improvement in accuracy (i.e. 85.94%) using only two discriminant bases. The generalization accuracy was further improved by applying our proposed weighting mechanism (i.e. 88.92%).

In this research (Dehzangi et al. 2018) the raw GSR signal was decomposed into its phasic and tonic components using Continuous Decomposition Analysis (CDA), and then the continuous phasic component containing relevant characteristics of the skin conductance signals were investigated for further analysis. A high resolution spectro-temporal transformation of the GSR signals was generated for non-distracted and distracted (calling and texting) scenarios to visualize the associated behavior of the decomposed phasic GSR signal in correlation with distracted scenarios. According to the spectrogram observations, relevant spectral and temporal features were extracted that capture the patterns associated with the distracted scenarios at the physiological level. Then feature selection was performed using support vector machine recursive feature elimination (SVM-RFE) in order to: (1) generate a rank of the distinguishing features among the subject population, and (2) create a reduced feature subset toward more efficient distraction identification on the edge at the generalization phase. Support vector machine (SVM) was employed to generate the 10-fold cross validation (10-CV) identification performance measures. The results demonstrated cross-validation accuracy of 94.81% using all the features and the accuracy of 93.01% using reduced feature space. The SVM-RFE selected set of features generated a marginal decrease in accuracy while reducing the redundancy in the input feature space toward shorter response time necessary for early notification of distracted state of the driver.

Chapter 2 Experimental Setup

This chapter presents the experimental design and data acquisition procedure using wrist band like wireless galvanic skin response wearable device.

A custom designed wearable data acquisition platform was developed comprising a synchronized multi-modal solution to acquire the physiological signals using a comprehensive wearable sensor network, which is used to collect data (Dehzangi & Williams 2015). This platform is capable of collecting large amounts of heterogeneous drivers' physiological including Galvanic Skin Response (GSR) during naturalistic driving. Experiments conducted in this thesis were approved by the Institutional Review Board (IRB) of the University of Michigan with the Submission ID: HUM00102869. The driver participants were given a consent form to sign before the experiments, in which the nature of the data collection was described. The experiments were conducted under constant supervision of two research investigators for a reliable data acquisition process. Wherein one of the research investigator sat beside the driver in the front to continuously monitored the driving subject to avoid any drastic situation. While the second research investigator sat behind to monitor and record the driving data from the participant continuously to acquire reliable recording for further investigations. The participants between the age group of 20–40 years that participated in our experiment were legally permitted to drive. Healthy male subjects were considered for this thesis to rule out the inconsistency. The subjects were asked to stay away from any alcoholic beverages or pharmaceuticals that would trade off their sharpness amid the investigation. Figure 2.1 depicts a driver subject while conducting our naturalistic driving experiment. Figure 2.2 illustrates the wireless galvanic skin response wrist band like wearable device used for data acquisition. Three scenarios of driving were considered in this thesis, each of which were performed by the subjects for ≈2 min and their corresponding GSR signals were recorded. The three driving scenarios were (i) driving under normal condition, (ii) driving while engaging in a phone conversation and (iii) driving while using the phone for texting. Normal driving (non-distracted) is represented by Scenario (i) while distracted driving is represented by scenarios (ii) and (iii).

Figure 2.1: A driver subject during our naturalistic driving experiment.

Figure 2.2: Wireless galvanic skin response (GSR) wearable device on the right-hand wrist

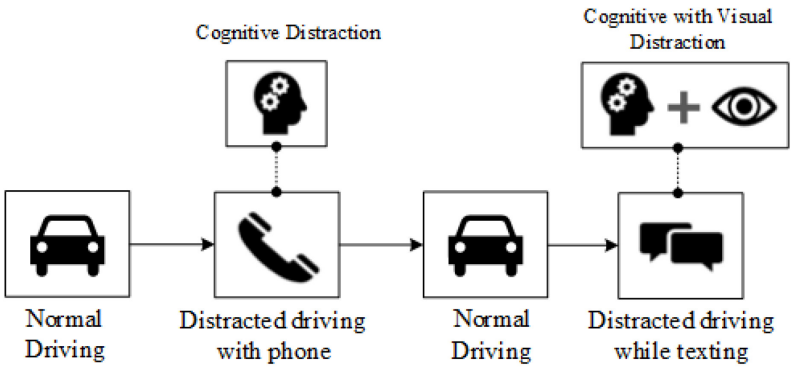


Figure 2.3: Illustrates data collection order during the naturalistic driving experiment

Figure 2.3 illustrates data collection order during our driving experiment. We employed a 10th order Butterworth low-pass filter <20 Hz on the recorded raw GSR to cope with and remove artifacts such as high frequency-noise, motion artifacts and electromyography (EMG) artifacts that might interfere due to the movement of hand and finger during phone and texting experiment sessions.

Chapter 3 Preliminary investigation of raw galvanic skin response for driver distraction detection

This chapter presents the preliminary analysis employed on the raw galvanic skin responses recorded during naturalistic driving experiment to detect driver distraction. Six driver subjects participated in this realistic driving experiments. In this investigation two tasks including driving under normal condition and driving while having an engaging conversation on the phone were considered. The former would represent normal and the latter would represent distracted driving. Standard statistical features were extracted from the recorded raw GSR signal. The experimental results demonstrated high accuracies of detection under subject dependent scenarios. The possibility of subject independent distraction detection employing non-linear space transformation based on kernel analysis and support vector machines (SVM) were also investigated.

3.1 Analysis method

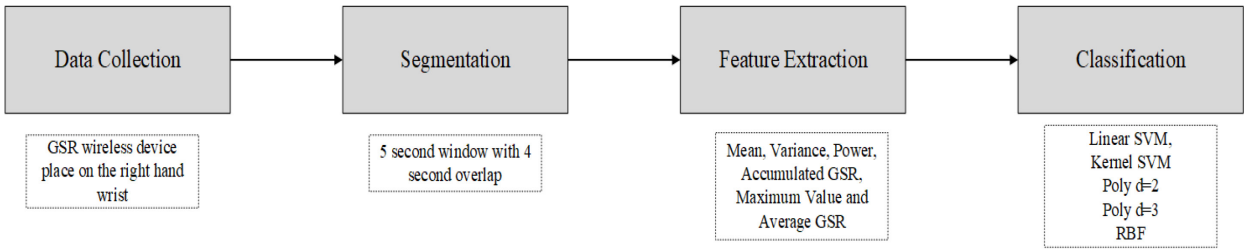


Figure 3.1: Flowchart for the proposed distraction detection system using raw galvanic skin conductance signal

Figure 3.1 illustrates the general flow of the proposed driver distraction detection system. Raw GSR signal were recorded using a wrist-band wearable wireless GSR device. The GSR signals were recorded for two scenarios 1) **Normal** driving focusing attention on the primary task of driving and 2) **Distracted** driving via having an engaging phone conversation while driving. Then segmentation of 5 second window with 80% overlap was employed. Statistical features were extracted from each window. Subsequently, state-of-the-art support vector machine (SVM) and its

kernel were employed for detection driver distraction. In this section the implementation of each step is discussed in detail.

3.1.1 Segmentation

After obtaining the clean raw GSR signal, segmentation was applied to obtain short windows and features were extracted from these windows. Each window had a size of 5 seconds with an overlap of 4 seconds. This was performed in order to meet the short-response time requirement of driver monitoring or alerting system.

3.1.2 Feature Extraction

After applying the segmentation, the following features for each window were extracted: mean, variance, accumulated GSR, average GSR, maximum Value, and power.

- Mean: It is the average of all the data points present in each window i.e. sum of all the data points divided by total number of data points in that window.
- Variance: It is the average of squared distance from the mean.
- Accumulated and Average GSR: Accumulated GSR is summation of GSR values of that window over task time (Nourbakhsh et al. 2013). Average GSR is normalized (Nourbakhsh et al. 2012) GSR value of a window over task time. Average GSR and accumulated GSR after applying normalization function to entire subjects

$$Normalise_{GSR(i,k,t)} = \frac{GSR(i,k,t)}{\frac{1}{m}\sum_{j=1}^{m}\sum_{t=1}^{Tij} GSR(i,j,t)} \quad 3.1$$

GSR(i,k,t) is values of each data point at time *t* of task *k* of subject *I* and *m* is the number of tasks. They also calculated accumulated and average GSR for task *k* of subject *I* as below:

$$accGSR(i,k) = \sum_{t} Normalised_{GSR(i,k,t)} \quad 3.2$$

$$avgGSR(i,k) = \frac{\sum_t Normalised_{GSR(i,k,t)}}{T} \quad 3.3$$

- Maximum Value: Maximum value of GSR in a window was the resulting value of this feature.

- Power: The sum of the absolute squares of a signals time-domain samples divided by the signal length, or, equivalently, the square of its root mean square (RMS) level.

3.1.3 Identification Task

3.1.3.1 Support vector machine (SVM)

Classification task was conducted using Linear Support Vector Machine (SVM). The SVM approach offer an effective classification strategy to separate input vector in a 2 class problem and is investigated in many different applications (Vapnik 1995; Guyon et al. 2002). SVM projects input vector x into a scalar value $f(x)$ as the output score.

$$f(x) = \sum_{i=1}^{N} \alpha_i y_i K(x_i, x) + b \qquad 3.4$$

where, the *vectors* $\{ x_i | i = 1, \ldots, N\}$ are the support vectors, *N* is the number of support vectors, $\alpha_i > 0$ are adjustable weights, $y_i = \{-1, +1\}, b$ is the bias term, and the function $K(x_i, x)$ is the kernel function. For the 2-class classification, the class decision is made based on the sign of $f(x)$. As it can be seen, the classifier is constructed from sums of the kernel function expressed as,

$$K(x_i, x) = \varphi(x_i)^t \varphi(x) \qquad 3.5$$

where $\varphi(x)$ is a mapping from the input space to a possibly infinite dimensional space.

3.1.3.2 SVM Kernel

- **Polynomial Kernel (Poly d=2)**: It is a kernel function that represents similarity of vectors in a feature space over polynomials of the original variables, making it possible for learning nonlinear models. The polynomial kernel determines the similarity of the input samples not only by looking at the features of the input but also combinations of the features. The polynomial kernel for degree-d is defined as:

  $$K(x, y) = (x^t\, y + c)^d \qquad 3.6$$

  where *x, y* are vectors in input space. When $c = 0$ the kernel is called homogenous and $c \geq 0$ is a free parameter trading off the influence of higher-order versus lower-order terms in the polynomial.
- **Radial Basis Function (RBF) Kernel:** The Radial Basis Function kernel, also called as RBF kernel, or Gaussian kernel, is in the form of radial basis function more specifically a Gaussian function.

$$K_{RBF}(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right) \quad 3.7$$

where x and x' are two samples, represented as feature vectors in some input space. $\|x - x'\|^2$ is the squared Euclidian distance between the two feature vector. σ is a free parameter.

3.1.4 10-Fold cross validation

Linear and kernel-based SVM were employed on the data corresponding to each subject separately for training and then evaluation of the predictive model was conducted using 10-fold cross validation (10-CV). In 10-CV, the original dataset is partitioned into 10 equal size subsets. Of the 10 subsets, a single subset is retained as the validation data for testing the model, and the remaining nine subsamples are used as training data. The cross-validation process is then repeated 10 times (the folds), with each of the 10 subsets used exactly once as the validation data. The 10 results from the folds can then be averaged to produce a single estimation. The advantage of this method is that all observations are used for validation in exactly one out of the 10 iterations without being participated in the training process whatsoever for that iteration.

3.2 Experimental Results

In order to evaluate the recorded raw GSR for detection of driver distraction, two types of predictive analysis were performed: 1) subject dependent analysis, and 2) subject independent analysis.

3.2.1 Subject dependent analysis

The detection accuracy on a subject by subject basis were evaluated. In this approach the data corresponding to each subject separately was considered for training and then evaluation of the predictive model was performed using 10-fold cross validation. The scatter plot of subject 1 (fairly similar across all the subjects) shown in Figure 3.2 depicts the classes "normal" (blue) vs. "distracted" (red), which shows a good separation between the two classes with few overlapping segments. The results for subject dependent analysis is shown in Table 3.1. Linear SVM produced an average accuracy of 89.45% averaged over all the subjects. A maximum accuracy of 97.4% was achieved with subject 5 and 95.8% for subject 1. The lowest accuracy of 80% was achieved with subject 4; several factors could affect the accuracy such as noise and motion artifacts while

collecting data. It can be observed from Table 3.1 that using kernels with SVM can only improve the detection results with RBF kernel (18% relative improvement).

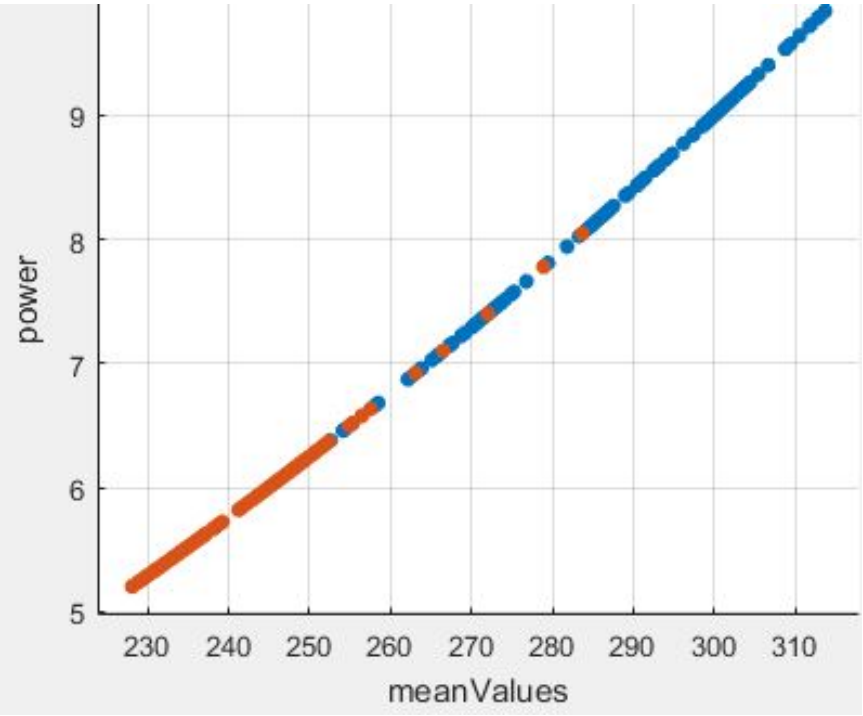


Figure 3.2: Scatter plot of the extracted segments of GSR for subject 1 using mean value and power features

Table 3.1:Classification accuracy for subject dependent analysis

Subjects	Accuracy in %			
	Support Vector Machine			
	Linear	Poly d = 2	Poly d = 3	RBF
Subject1	95.8%	96.7%	96.3%	96.3%
Subject2	82.1%	81.2%	83%	86.6%
Subject3	87.5%	86%	87.1%	89.2%
Subject4	80.6%	83.6%	83.8%	84.6%
Subject5	97.4%	97.4%	95%	97.1%
Subject6	93.3%	93.3%	94.2%	94.2%
Average	**89.45%**	**89.70%**	**89.90%**	**91.33%**

3.2.2 Subject independent analysis

In this case, all the subjects were combined as follows. The “normal” data corresponding to all subjects were pooled into one class and all subject “distracted” data were pooled into another.

Figure 3.3 shows the scatter plot of meanValues and power features after pooling the subjects and high levels of overlap in the input feature space. Table 3.2 reports detection accuracies for subject independent analysis using different kernels with SVM.

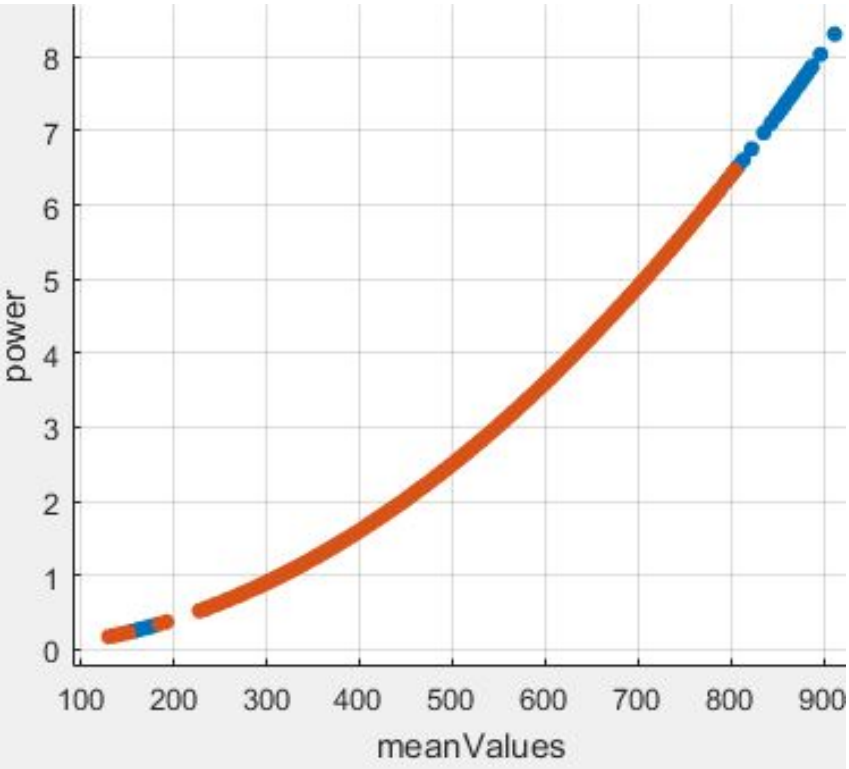


Figure 3.3: Scatter plot of subject independent analysis when all the subjects were pooled together

The linear SVM detection rate degraded dramatically to 70.1% which was predictable based on Figure 3.3 scatter plot. RBF kernel effectively improved the detection accuracy to 76.7% (22% relative improvement) by inducing non-linearity into the feature space using a Gaussian kernel. The result for subject independent analysis is shown in Table 3.2. The kernel SVM namely polynomial and RBF demonstrated a promising result with an accuracy of 76.7%. The Gaussian RBF kernel could classify the non-linear data of subject independent task effectively. The generated results demonstrated high accuracies of detection under both subject dependent and subject independent task. Linear SVM on subject dependent task showed high average accuracy of 89.45% for all the six subjects. SVM kernels were employed on subject independent task. Combining all the subjects made the input feature space more overlapped. Kernels demonstrated acceptable results on the subject independent analysis. RBF showed a promising result of 76.7%.

Table 3.2: Accuracy results for subject independent analysis using SVM with different kernels

Subjects	Accuracy in %			
	Support Vector Machine			
	Linear	Poly d = 2	Poly d = 3	RBF
Concatenated all subjects	70.10%	70.40%	71.70%	76.70%

3.3 Conclusion

GSR was investigated to analyze whether it can be used to detect distraction under naturalistic driving conditions using a wrist band wearable. Experimental results demonstrated high accuracy of detection under both subject dependent task as well as subject independent task. Predictive accuracy for subject dependent resulted in high average accuracy of 91.33% which is a promising result using minimally intrusive GSR technology. Next all the subjects were combined to investigate about subject independent analysis. Non-linear kernels employed demonstrated better performance for this scenario due to high levels of overlap in the input feature space which resulted in the accuracy of 76.7% which is about 90% of the subject dependent accuracy. This study provided evident results to employ unobtrusive GSR technology in the vehicle in order to detect distraction under naturalistic driving.

Chapter 4 Frequency, Time and Statistical Domain Feature Analysis

Analysis conducted in the previous chapter provided evident and satisfactory results to use GSR signal as good indicator of distraction while driving. In this chapter further analysis is performed on the raw GSR signal by extracting relevant spectral and temporal features that contain most discriminative information to characterize driver distraction. In this chapter, three scenarios were investigated 1) **Normal** driving focusing attention on the primary task of driving 2) **Phone Distracted** driving while having an engaging phone conversation and 3) **Text distracted** driving while writing and sending texts while driving. Phone distraction is considered as cognitive distraction element in comparison to text distraction, which represent cognitive and visual distraction at the same time. A custom designed frequency domain, time domain and statistical feature space was generated for the recorded raw GSR signal that represents most discriminative information to characterize driver distraction. Due to the fact that feature extraction is a manual process, the redundancy and computational complexity of the feature space might limit the distraction identification performance. Therefore, linear discriminant dimensionality reduction was employed to discover discriminative bases of the GSR multivariate feature space to more efficiently identify distraction. A class dependent weight was incorporated to the calculation of scatter matrices, aiming to alleviate limitations of LDA with multiple distracting categories and to eliminate the overlap between distracted vs. non-distracted scenarios in the generalization phase.

4.1 Analysis Method

Figure *4.1* depicts a flowchart of the proposed methodology. The recorded raw GSR signal was segmented to meet the real-time requirement of a driver monitoring system.

Segmentation method of 5 second windows with 4 second overlap was employed to meet the requirement of short-response time for a driver monitoring or alerting system, bringing a real-time approach to our model. From every window custom designed relevant frequency domain, time domain and statistical features were extracted. Further dimensionality reduction technique

using linear discriminant analysis and weighted linear discriminant analysis were employed. Linear and

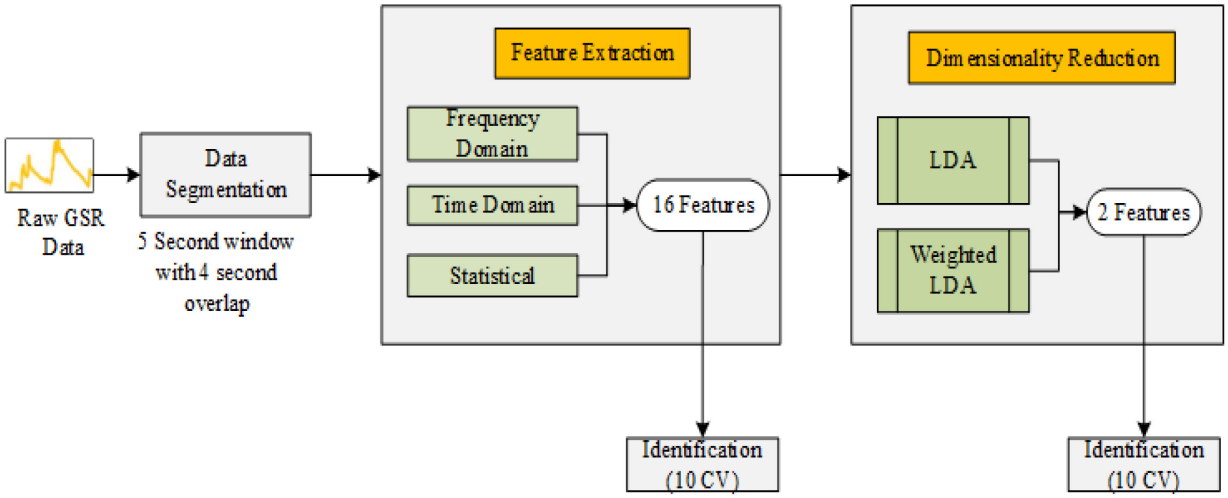


Figure 4.1: Flowchart of the proposed methodology

kernel-based support vector machine (SVM) with 10-fold cross validation were employed to generate identification results. Implementation of each step is discussed in detail in this section.

4.1.1 Feature Extraction

In this chapter custom designed frequency and time domain features the contained most discriminative information were generated form from every segmented window for the raw GSR signal. The above features along with the statistical features extracted in chapter 3 was used to create a high dimensional (16-D) feature space. The output of these features was labeled appropriately and considered as the sample data point of the feature space. Table 4.1 shows the list of extracted features.

- Short Time Fourier Transform (STFT): STFT is a technique that is quite often used to analyze physiological signal. It can be used to conduct analysis of signals both in frequency and time domain. After dividing the original signal into equal segments, Fourier Transform is applied independently to each segment. In (Lew et al. 2008), STFT analysis on SC data showed effective results in detecting work load. We extracted four STFT coefficients each representing a sub-band with 12.5 Hz bandwidth between 0 and 50 Hz.

Table 4.1: List of extracted features

Feature Domain	Feature Names
Statistical	Mean, Variance, Accumulated GSR, Average GSR and Maximum Value
Frequency	Short Term Fourier Transforms (4 features)
Time	Fractal Dimensions and Auto Regressive Coefficients

- Fractal Dimension: Fractal dimension is used to determine the chaotic or fractal nature of a physiological signal. It is an impressive mathematical tool to model various complex physiological signals and an index to quantify the complexity of a fractal pattern. One of the common techniques to characterize the time-series data, Katz method was used to extract one fractal dimension feature. Katz's approach calculates the fractal dimension as follows:

$$D = \frac{\log\left(\frac{L}{a}\right)}{\log\left(\frac{d}{a}\right)} = \frac{log(n)}{\log(n) + log\left(\frac{d}{L}\right)'} \quad 4.1$$

  where total length of the curve is given by *L*, that is sum of the distance between successive points. Euclidian distance is used to calculate distance between two points of the signal. The maximum distance between the initial point and any other point of the sample is given by *d*, *a* is the average distance and *n* is *L* divided *a* (Esteller et al. 2001).

- Auto-Regressive (AR) coefficients: In an AR model of order *p*, the current output is estimated based on the linear combination of past *p* output values plus some noise term. The mean squared prediction error of an auto-regression model is minimized by calculating the weights on the previous *p* outputs. The model with current output value $y(n)$ and zero mean white space noise input $x(n)$ is:

$$y(n) = \sum_{k=1}^{p} a(k)y(n-k) = x(n) \quad 4.2$$

  The parameters of order *p* were selected as the features for each window. The order of *p* was set to 5 which generated five AR features.

4.1.2 Discriminative Feature Space Analysis and Transformation

An overall of 16 features (includes statistical, frequency and time domain features) were generated for each subject. To normalize the feature space towards the identification task, transformation of the feature space was employed using linear discriminant dimensionality reduction to discover discriminative bases of the GSR multivariate feature space purposed toward distraction identification. Further a class dependent weight was proposed to incorporate to the calculation of scatter matrices, aiming to alleviate limitations of LDA with multiple distracting categories and to eliminate the biases between distracted vs. non-distracted scenarios for unseen data at the generalization phase.

4.1.2.1 Linear Discriminant Analysis (LDA):

LDA is a linear transformation technique that performs dimensionality reduction of the feature space while preserving the class discriminatory information as much as possible (Yan et al. 2007) . It assumes the distribution of each class is Gaussian and all the classes have same covariance matrix. LDA tends to find the directions also known as 'linear discriminants' that represents the axes along which classes are best separated (Farooq & Dehzangi 2017). A linear projection matrix $y = U^t x$ is computed by maximizing the ratio of between class scatter S_b and within class scatter S_w and expressed as follows:

Within-class scatter matrix:

$$S_w = \sum_{i=1}^{c} \sum_{j=1}^{n_i} (Y_j - M_i)\,(Y_j - M_i)^T \tag{4.3}$$

Between-class scatter matrix:

$$S_b = \sum_{i=1}^{c} (M_i - M)\,(M_i - M)^T \tag{4.4}$$

A transformation space is computed by LDA that maximizes the between class scatter and minimizes the within class scatter, expressed as follows:

$$max \frac{\left|\tilde{S}_b\right|}{\left|\tilde{S}_w\right|} = max \frac{|U^T S_b U|}{|U^T S_w U|} \tag{4.5}$$

4.1.2.2 Linear Discriminant Analysis with Class-Dependent Weight

In distraction identification task, there are multiple states to distinguish namely, phone and text. Since phone conversation only affects the driver's cognitively it is expected to be less distracting. However, texting has further effects on the driver because the driver is visually distracted as well. Therefore, it was hypothesized that impact on the recorded signals would be different as well. A weight was incorporated to calculate within class scatter matrix for each class. The distance between the centroid of each class and the global mean is given by weight ω_i.

$$S_w = \sum_{i=1}^{c} \frac{1}{\omega_i} . \sum_{j=1}^{n_i} (Y_j - M_i)\,(Y_j - M_i)^T \qquad 4.6$$

4.2 Experimental Results

In order to characterize and identify driver distraction with short response time first the recorded GSR signal is segmented into small windows and then 16 different features were generated which includes statistical, frequency and time domain features from each window. Linear and kernel-based support vector machine was employed on the data corresponding to each subject separately for training and then evaluation of the predictive model was performed using 10-fold cross validation. Initially all the 16 features were used for predicting accuracy. Table *4.2* shows the classification models accuracy generated using 16 features.

The scatter plot of subject 1 (fairly similar across all the subjects) generated using t-sne (Van Der Maaten & Hinton 2008) shown in Figure 4.2 depicts the classes "normal" (red), "phone" (green) and "text" (blue). The t-sne is a state of the art 2-D non-parametric representation of the manually generated 16-D feature space. Figure 4.2 shows a distinguished trend of change in the GSR-based feature space from normal driving toward phone distraction and further toward text distraction. This observation suggests that text distraction introduce higher amount of changes to the GSR-based features space, which might be due to the impact of visual distraction on top of cognitive distraction. The classification models resulted with an average accuracy of 85.19% averaged over all the subjects. A maximum accuracy of 92.5% was achieved with subject 4. The results reported in Table *4.2* demonstrated that GSR-based time and frequency feature extraction is useful to detect the state of distraction for each subject effectively. Examining Figure 4.2, there is still a great deal of overlap exists between the normal vs. phone scenarios.

Table 4.2: Classification accuracy of all subject using 16 features

Subjects	Accuracy in %			
	Support Vector Machine			
	Linear	Poly d=2	Poly d=3	RBF
Subject 1	84.60	84.30	63.40	88.60
Subject 2	71.00	76.10	76.30	79.60
Subject 3	84.70	88.20	87.30	87.30
Subject 4	86.60	89.30	92.50	91.30
Subject 5	75.30	79.10	78.40	81.40
Subject 6	80.60	91.20	88.20	92.10
Subject 7	71.90	75.30	78.80	76.00
Average	**79.24**	**83.36**	**80.70**	**85.19**

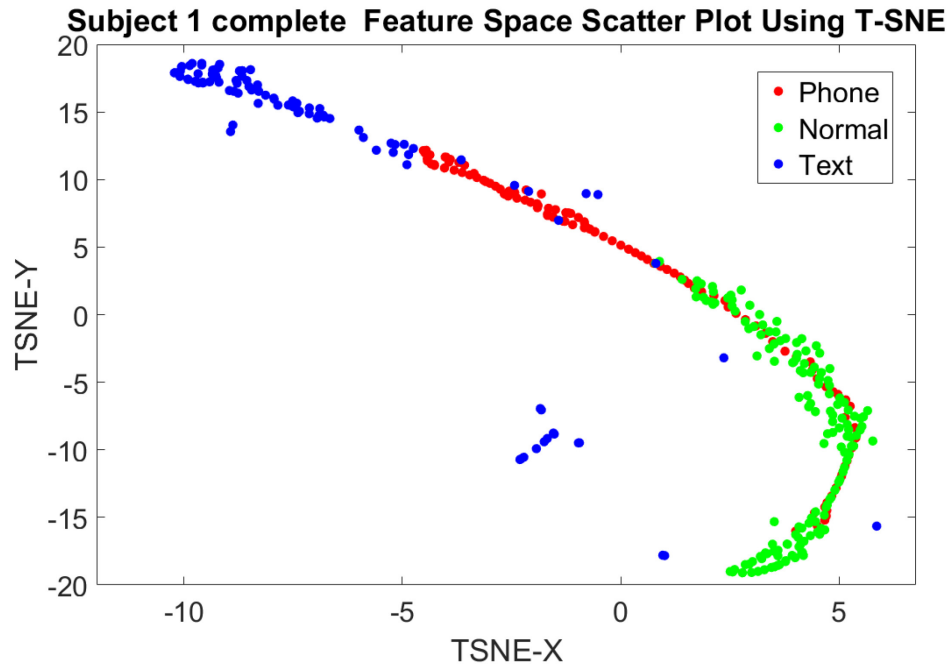


Figure 4.2: Subject 1 scatter plot of complete feature space using t-sne representation

Figure *4.2* also reports dramatic improvement from linear SVM to the best kernel-based SVM 10-CV accuracy (~29% relative improvement) which is an evidence of high degrees of overlap in the feature space. Although kernel-based SVM can alleviate the overlap in the feature

space using kernel non-linear expansion. However, it can also cause a great deal of complexity in time and space of the detection procedure.

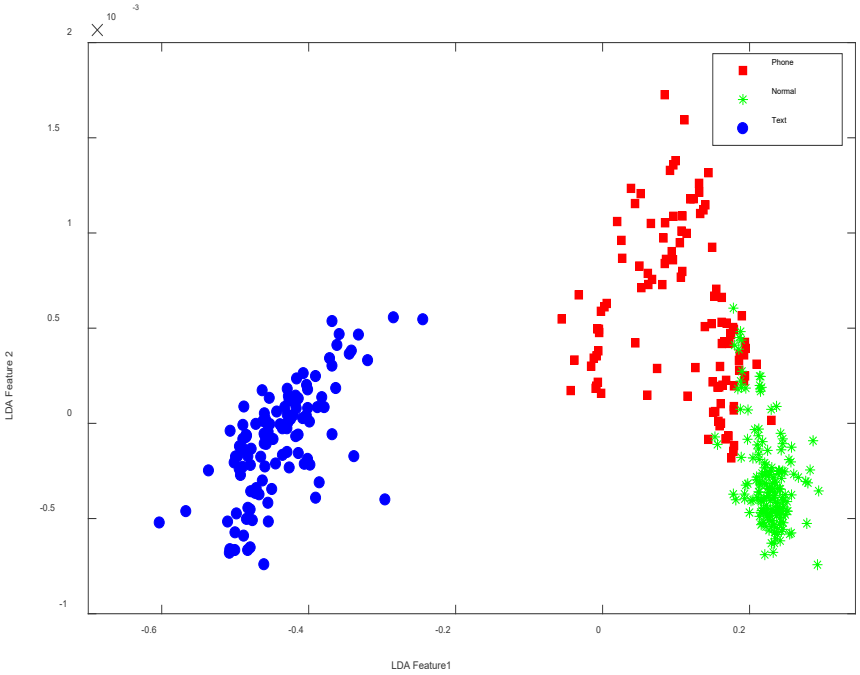


Figure 4.3: Subject 1 scatter plot after applying linear discriminant analysis (LDA) dimensionality reduction technique

Table 4.3: 10-CV identification accuracy of subjects after applying LDA

Subjects	**Accuracy in %**			
	Support Vector Machine			
	Linear	**Poly d=2**	**Poly d=3**	**RBF**
Subject 1	86.40	89.40	88.90	87.50
Subject 2	75.50	76.30	69.30	77.40
Subject 3	85.60	89.40	89.40	89.50
Subject 4	92.30	95.30	95.50	95.00
Subject 5	77.10	80.20	80.30	78.90
Subject 6	88.00	91.20	91.20	92.30
Subject 7	75.80	78.30	74.20	81.10
Average	**82.95**	**85.73**	**84.10**	**85.94**

LDA is an effective linear dimensionality reduction as it transforms the feature space into lower dimension by finding linear mapping that maximizes the class separability along with the direction in which most discriminative features of multiple classes are present. LDA was conducted to improve discrimination and reduce redundancy in the original 16-D space. Figure 4.3 shows a scatter plot of the transformed feature space after applying dimensionality reduction using LDA, which shows better separation between three classes in a 2-D discriminant bases space.

Prediction accuracy after applying dimensionality reduction using LDA is shown in Table 4.3. LDA proved to be successful in transforming the high dimensional feature space into low dimension while preserving class discriminating features by demonstrating an average accuracy of 85.94% averaged over all the subjects.

Table 4.4: 10-CV identification accuracy of subjects after applying weighting

Subjects	**Accuracy in %**			
	Support Vector Machine			
	Linear	**Poly d=2**	**Poly d=3**	**RBF**
Subject 1	93.4	93.2	**93.7**	93.5
Subject 2	82.5	**84.6**	84.3	84.3
Subject 3	89.6	89.4	89.4	**90.4**
Subject 4	**95.5**	95.3	95.5	95.5
Subject 5	82.7	83.2	**83.3**	83.1
Subject 6	91.3	**92.2**	92.2	92
Subject 7	80.5	82.3	82.9	**83.6**
Average	**87.93**	**88.60**	**88.75**	**88.91**

A best case with maximum accuracy of 95.5% was demonstrated by Subject 4. This cost reduction in time, computational complexity and reduced feature space will able our current model to move towards real-time driver monitoring system. Table 4.3 also reports less improvement from linear SVM to the best kernel-based SVM 10-CV accuracy (~17% relative improvement) which is an evidence of lower degrees of overlap in the feature space. However, as shown in Figure *4.3*, phone class has more overlap with normals driving compared to texting. LDA can generate an optimistic estimate of between- and within-class scatter matrices if one or more of the classes are

well-separated even if there are other highly overlapped classes. In order to take that into account, we employed our proposed weighting scheme to reduce the impact of within class scatter matrix for the text class in learning LDA transformation matrix as discussed in section 4.1.2.2. The aim of the presented weighting mechanism is to alleviate the biased estimation of the within class scatter matrices and improve the prediction accuracy. The aim was to focus on reducing the overlap between the normal vs phone classes while lowering the contribution of text class in the optimization, which might not introduce any negative impact.

Table 4.4 also reports minimal improvement from linear SVM to the best kernel-based SVM 10-CV accuracy (~8% relative improvement) which is an evidence of much lower degrees of overlap in the feature space. This observation is specifically significant since linear detection and identification systems are generally more preferable due to lower complexity and higher reliability in practice.

4.3 Conclusion

In this chapter, raw GSR was investigated to check whether it can be used for empirical characterization of driving task under naturalistic driving conditions using a wrist band wearable. Seven driver subjects participated to perform three tasks i.e. driving normally, driving while having phone conversation and driving while using phone for texting.

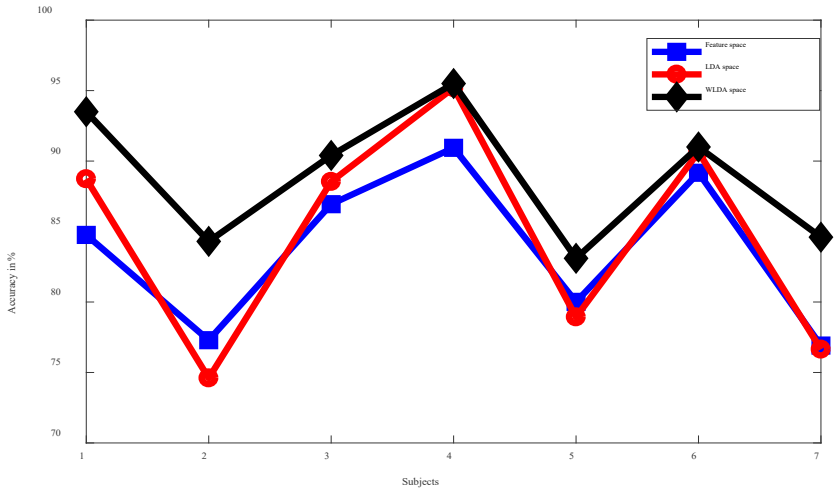


Figure 4.4: Accuracy % chart for all subjects

Most relevant 16 features were extracted that included statistical, frequency domain and time domain features to capture changes/pattern at physiological level. As feature extraction is a manual

process and also to normalize the feature space towards identification task the original 16-dimensional feature space was transformed using linear discriminant dimensionality reduction to discover discriminative bases of the GSR multivariate feature space to more efficiently identify distraction.

A class dependent weight was incorporated to for the calculation of scatter matrices, aiming to alleviate limitations of LDA with multiple distracting categories and to eliminate the overlap between distracted vs. non-distracted scenarios in the generalization phase.

Figure 4.4 shows the accuracy % comparison of the original 16-D feature space vs. reduced feature space with LDA and weighted LDA. This analysis provided evident results that the extracted frequency and time domain features contained the most discriminative information to characterize and identify driver distraction using raw GSR signals. These extracted measures were then made as standard to represent the GSR signal to identify distraction and were further used in the next analysis of this thesis.

Chapter 5 Continuous Decomposition of Galvanic Skin Responses

In this chapter the raw GSR signal is decomposed into its phasic and tonic component using continuous decomposition analysis (CDA) technique (Benedek & Kaernbach 2010). Skin conductance (SC) or galvanic skin response (GSR) is basically made up of two components i) skin conductance level-SCL (tonic component) and ii) skin conductance responses-SCR (phasic component). The phasic component contains the most useful information to characterize the GSR signal. This was further supported by a conducting high resolution Spectro-temporal analysis on the phasic component of GSR signal. Thus, only the phasic component of GSR was used for identification of distraction in this chapter. A population of 10 driver subjects participated in this study during real driving experiments. Three scenarios were investigated: (1) normal driving focusing attention on the primary task of driving; (2) phone distracted driving while having an engaging phone conversation; and (3) text distracted driving while writing and sending texts when driving. Based on the results obtained in the previous chapter, several spectral and temporal measures were extracted that characterize the phasic GSR signal in correlation with distracted scenarios. Additionally, two more measures that represented the phasic component of GSR were extracted based on the observation of the high resolution spectro-temporal analysis. Since input feature space is constructed in a manual process, the redundancy and computational complexity of the space might decrease the accuracy and response time of distraction identification in the generalization phase. Therefore, support vector machine – recursive feature elimination (SVM-RFE) (Huang et al. 2014) was employed to remove the redundancies for more efficient processing on the edge. SVM-RFE was employed in order for the following: (1) generate a rank of the discriminative features for the subject population, and (2) create a reduced feature subset with the highest distraction identification accuracy. The results using SVM-RFE demonstrated marginal decrease in accuracy while reducing the computational complexity and the redundancy in the input space towards early notification of distraction state to the driver.

5.1 Analysis Method

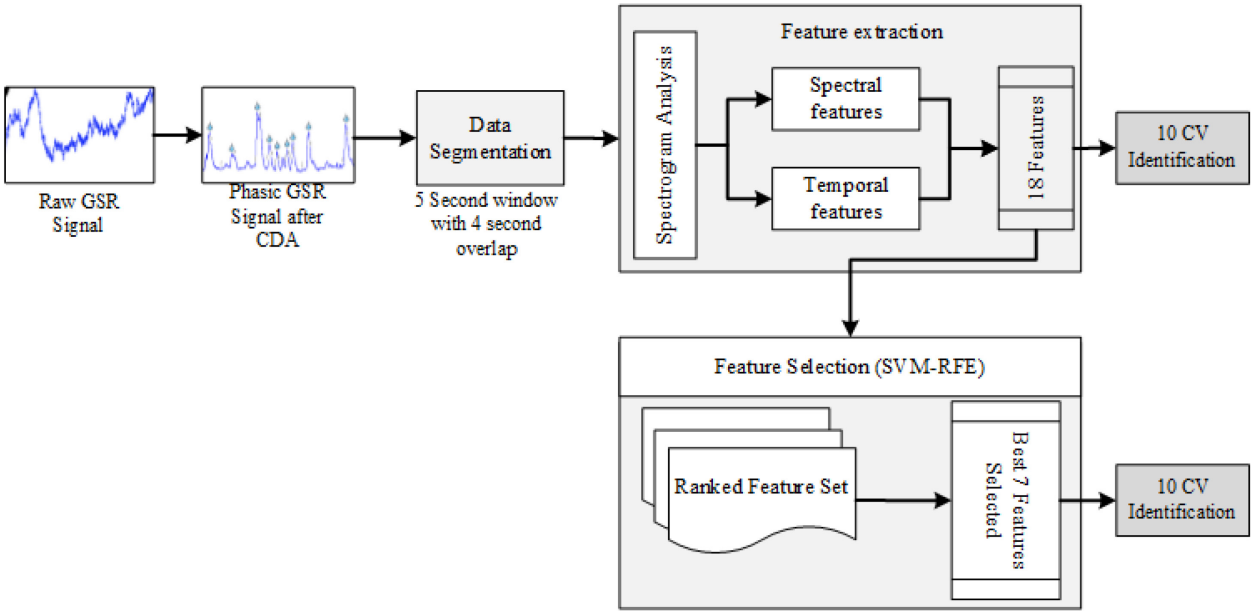


Figure 5.1: Flowchart of the proposed driver monitoring and intervention system on the edge.

Figure 5.1 illustrates the flowchart of the proposed driver monitoring and intervention system on the edge. After preprocessing, the recorded raw GSR signal was decomposed into phasic and tonic component using continuous decomposition analysis (CDA) and based on spectro-temporal analysis and characterization of the phasic GSR, segmentation and feature extraction modules were designed and developed. The extracted features were then used for identification tasks. Furthermore, feature selection was implemented using SVM-RFE to reduce the dimension of the feature space to alleviate the curse of dimensionality and improve the response time. In this section, we discuss the implementation of each steps in detail.

5.1.1 Continuous Decomposition Analysis (CDA)

Skin conductance data is described by the superposition of subsequent skin conductance responses (SCRs). Due to this characteristic of the SCRs, the process of calculating the actual responses to a sympathetic activity in response to an external stimulus becomes tedious. This limitation is overcome by the deconvolution technique that separates the skin conductance (SC) data into phasic and tonic continuous activities. The tonic activity might include noise and shows subject dependencies.

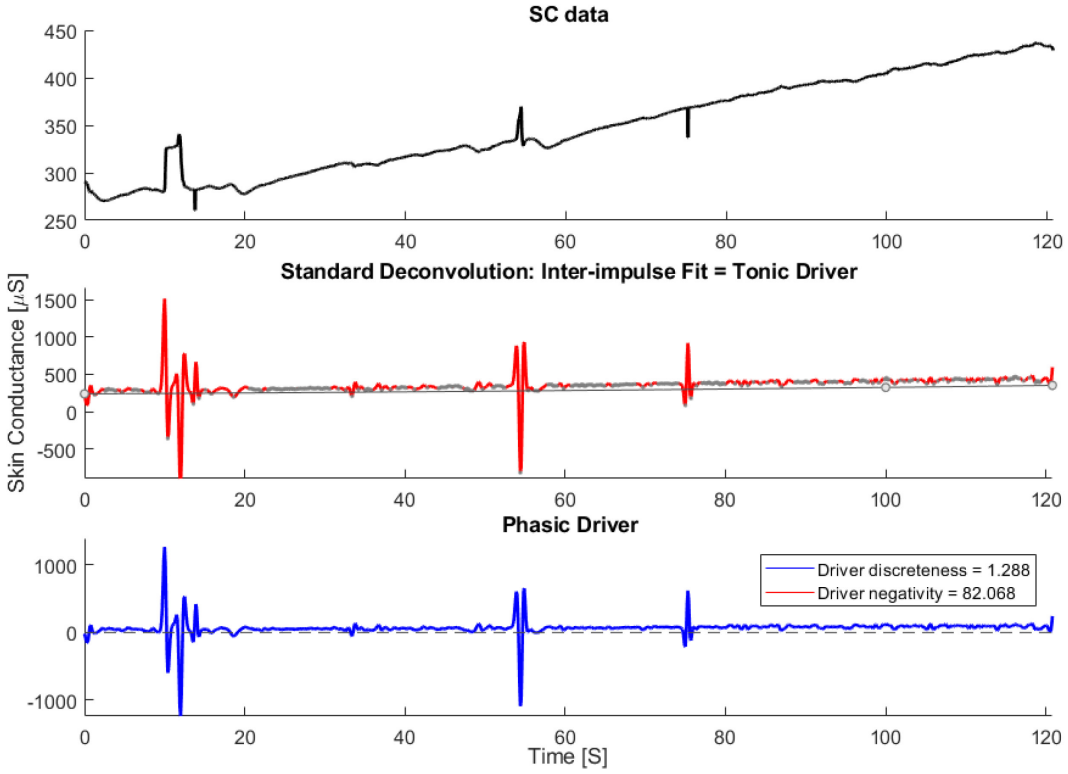


Figure 5.2: Continuous decomposition analysis (CDA) for normal scenario.

The tonic activity can be observed as a trend in the original SC signal in the top subplot of Figure 5.2. On the other hand, the phasic activity of the SC signal is considered for further investigation as this component of SC signal contains the actual response to any event-related sympathetic activity predominantly in the form of distinct burst of peaks with a zero baseline. The phasic component of SC also demonstrated trends in spectro-temporal space in correlation with the distracted scenarios. Extracting the phasic component is done in three steps: deconvolution of galvanic skin response (GSR) data, computation of tonic activity and computation of phasic activity. A particular change in skin conductivity is triggered by secretion of sweat due to activity of the sudomotor nerve. In mathematical terms, the sudomotor nerve activity can be treated as a driver, containing distinct sequence impulse/bursts, which triggers a particular impulse response (i.e., SCRs). The outcome of this procedure can be described by convolution of the driver with impulse response function (IRF). IRF characterizes the shape of impulse response over time (Benedek & Kaernbach 2010):

$$SC_{phasic} = Driver_{phasic} * IRF \quad (5.1)$$

The phasic activity is believed to bestride a gradually changing tonic activity. Hence, SC activity can be considered to be composed as follows:

$$SC = SC_{tonic} + SC_{phasic} = SC_{tonic} + Driver_{phasic} * IRF \quad (5.2)$$

Tonic activity can also be considered as convolution of a driver function with same IRF. SC data can then be written as:

$$SC = (Driver_{tonic} + Driver_{phasic}) * IRF \quad (5.3)$$

Deconvolution is the reverse process of convolution. Skin conductance data's deconvolution incorporates a phasic and tonic fraction. By estimating one of them, the other can be determined easily:

$$\frac{SC}{IRF} = Driver_{SC} = \left(Driver_{tonic} + Driver_{phasic}\right) \quad (5.4)$$

Tonic electro-dermal activity can be observed in the absence of any phasic activity. However, SCRs (representing phasic SC activity) have a slowly recovering trail that may obscure any tonic SC activity. For the driver, the time constant of phasic responses is markedly reduced and so is their overlap. Time intervals between distinct phasic impulses can then be used to estimate tonic activity. Convolution can be conceived as a smoothing operation. Deconvolution has the reverse effect and amplifies error noise. Therefore, the resulting driver is smoothed by convolution with a Gaussian window. According to Equation 5.4, the phasic driver can now be computed by subtracting the tonic driver from the total driver signal. This subtraction results in a signal, which shows a virtually zero baseline and positive deflections reflecting the time-constrained nature of the phasic activity underlying the original SC data. The above methodology was driven from previous work (Benedek & Kaernbach 2010). Figure 5.2 and Figure 5.3 represents the deconvolution/decomposition of skin conductance signal for two scenarios, normal and distracted. The top most subplot in both the figures represents the original SC signal, the middle subplot represents deconvolution of tonic driver and the bottom subplot represents the phasic driver of the SC signal that contains the most discriminating component to characterize distraction. It can be observed from Figure 5.2 that, for a normal scenario, the phasic driver does not show impulse burst (bursts of consecutive peaks) as the subject is not under heavy workload. However, when we observe Figure 5.3 for the distracted scenario, more impulse bursts are observed as the subject is in a distracted state undergoing cognitive and visual distraction.

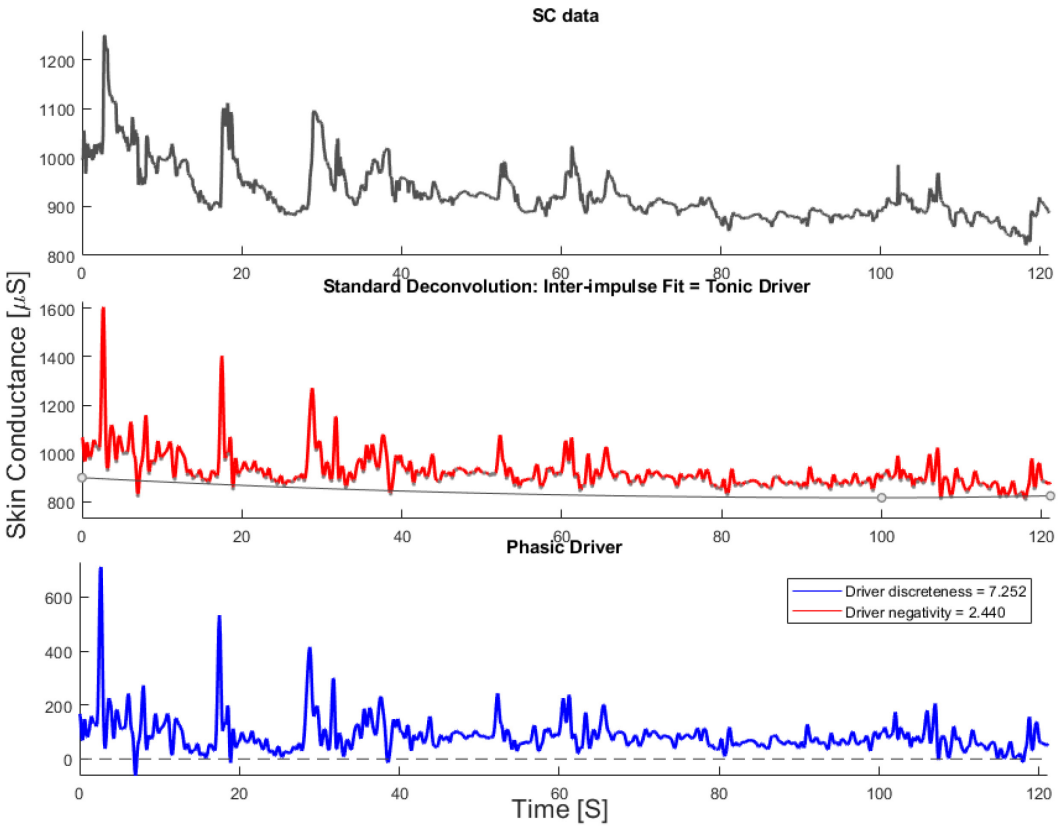
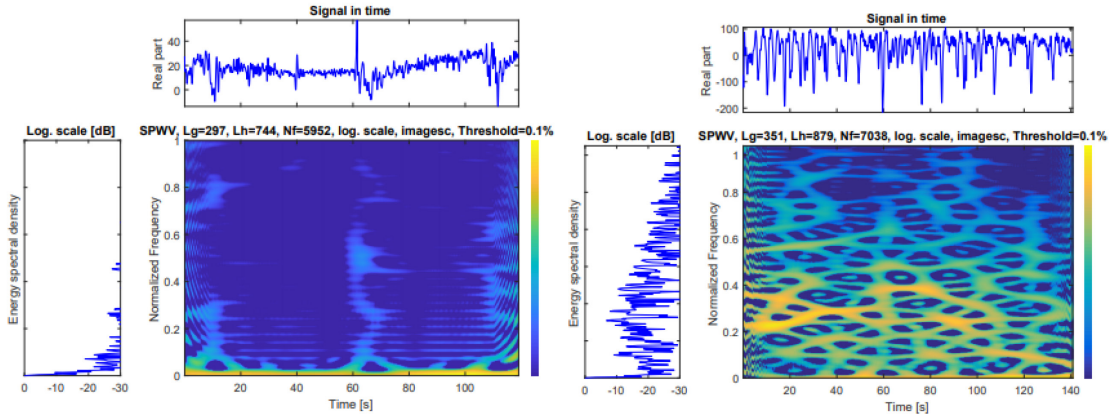


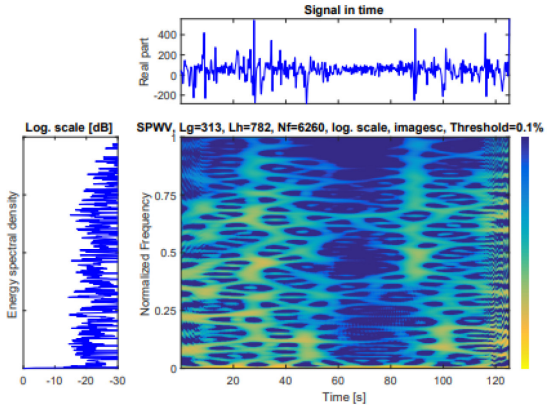
Figure 5.3: Continuous decomposition analysis (CDA) for distracted scenario

5.1.2 Spectral Analysis of Phasic Skin Conductance

Phasic SC as a non-stationary and multi-component signal varies in time and components over the time, and it needs to be analyzed in both temporal and spectral space to have a comprehensive evaluation on its characteristics. In order to investigate the spectro-temporal characteristics of the phasic SC signals related to different states of the driver, Time-Frequency (TF) analysis was conducted and a high-resolution TF representation was designed. TF analysis could be considered as non-stationary signals analysis with frequency content varying with time. TF is a suitable representation for non-stationary and multi-component signals, which is able to describe the energy distribution of the given signal over time and frequency space simultaneously. The TF presents the beginning and end times of the different components of the signal as well as their frequency scope. To achieve the TF representation of the phasic SC signal, we use the Wigner–Ville Distribution (WVD) (Pedersen 1997) approach attempting to attenuate the unwanted cross-terms in the TF space relative to the signal components.



(**a**) TF of phasic GSR for normal scenario (non-distracted). (**b**) TF of phasic GSR for phone scenario (distracted).



(**c**) TF of phasic GSR for text scenario (distracted).

Figure 5.4: High-resolution time-frequency (TF) representation of the phasic component of GSR signal.

Figure 5.4 depicts TF representation of three scenarios, including phone, text, and normal states. Figure 5.4 a-c represents normal, phone, and text scenarios, respectively, where the top panel in all of the sub-figures represents a raw phasic SC signal. The left panel depicts energy spectral density of a subject. The spectrograph represent time on the x-axis and normalized frequency on the y-axis, and the color is used to indicate the power of the TF sample. Based on the Nyquist frequency equation, with the sampling frequency of 50, the frequency range is 0–25 Hz, which was normalized to 0–1 Hz. The spectrograms were generated with normalized frequency observing the spectrograms in Figure 5.4, and we could identify considerably higher frequency components and peaks during distracted scenarios versus normal scenarios.

Based on the observations in the TF space, a signal processing recipe was designed for driver distraction identification on the edge for proactive monitoring and intervention. In the following sections, implementation of segmentation, feature extraction from spectral and temporal domain, feature selection using SVM-RFE are discussed.

5.1.3 Feature Extraction

Along with the spectral and temporal features that were extracted in the previous chapter, additional spatial features were extracted based on the analysis of generated spectrograms. The results of the calculated features were labeled accordingly, and the feature space was generated using these sample data points. Table 5.1 illustrates the list of extracted features.

Table 5.1: List of extracted features

Feature Domain	Feature Names
Temporal	Mean, Variance, Accumulated GSR, Average GSR, Maximum Value, Number of peaks and Sum of Amplitude of peaks, Fractal Dimensions, Auto Regressive
Spectral	Short Term Fourier Transforms (4 features)

- **Number of peaks based on second derivative:** The phasic GSR, *P(i)* signal's first and second derivatives, *Q0(i)* and *Q1(i)* are calculated as in (Kher et al. 2010), as follows:

$$Q_0(i) = |P(i+1) - P(i-1)| \quad 5.5$$

$$Q_1(i) = |P(i+2) - 2P(i) + P(i-2)| \quad 5.6$$

These two arrays are scaled and then summed:

$$Q_2(i) = 1.3 * Q_0(i) + 1.1 * Q_1(i) \quad 5.7$$

This array is scanned until a threshold is met or exceeded:

$$Q_2(j) \geq 1.0 \quad 5.8$$

- **Summation of amplitude of peaks:** The amplitude of number of peaks detected in a window previously is calculated and the amplitude is summed.

5.1.4 Feature Selection Using Support Vector Machine—Recursive Feature Elimination (SVM-RFE)

SVM-RFE feature selection methodology based on support vector machine (SVM) was introduced by (Guyon et al. 2002) and the authors used this methodology to select an important subset of features. It reduces the computational time for classification along with improvement of classification accuracy. The basis of SVM-RFE is backward removal of features iteratively.

Algorithm 1 : SVM-RFE for Ranking features

begin
Given set of features, $F \subset X$
where X is the sample training dataset
Ranked set of features, $R \leftarrow \emptyset$
repeat
Train linear SVM with feature set F
Calculate the weight of each feature ω_i
For each feature $f \in F$ **do**
Compute sorting standard: $c_i = (\omega_i)^2$
EndFor
Find the feature with minimum weight, $f_{min} = arg_{min}\{c\}$
Update $R = R \cup \{f_{min}\}; F = F \setminus \{f_{min}\};$
until all features are ranked
end : output R

Figure 5.5: Algorithm of SVM-RFE for ranking features

The main steps for feature selection is as follows (Huang et al. 2014):

- Input the dataset to be classified,
- Weight of each feature is calculated,
- Removal of features having the smallest weight to obtain ranking of features.

Initially, entire features in the dataset is considered to compute ranking weights for all features. The linear SVM classifier is used to compute the rank weight of each feature. Then, iteratively, features with the lowest rank weight are discarded until only one feature remains in the dataset. Finally, the features will be listed in descending order of generated ranked weights. Figure 5.5 illustrates algorithm for an iteration of SVM-RFE for ranking features in which the weight of each feature ω_i is given by

$$\omega_i = \sum_i (\alpha_i y_i x_i) \quad 5.9$$

where α is a linear SVM classifier used to train the given feature set *F* and is given by

$$\alpha = SVM_{Train}(X, y) \quad 5.10$$

and *X* is the input sample dataset and *y* is the corresponding class label.

5.2 Experimental Results

In this section, the experimental analysis and results of the proposed methodology is presented. In order to meet the short response time to identify distraction, the decomposed phasic GSR signal was segmented into 5 seconds windows with 4 seconds overlap and then extracted 18 different spectral and temporal features from each window based on spectrogram observations. Linear and kernel-based SVM was employed on the data corresponding to each subject separately for training and then evaluation of the predictive model using 10-fold cross validation (10-CV).

Table 5.2: Identification results using linear and kernel-based support vector machine (SVM) including accuracy, prediction speed, and training time with 18D feature space.

Subjects	**Classifier Performance**								
	Support Vector Machine								
	Linear SVM			**Poly d=2**			**Radial basis function**		
	Accuracy	Prediction Speed	Training	Accuracy	Prediction Speed	Training	Accuracy	Prediction Speed	Training
	in %	in ~ obs/sec	Time in sec	in %	in ~ obs/sec	Time in sec	in %	in ~ obs/sec	Time in sec
Subject 1	82.90	6600.00	0.70	87.00	5700.00	0.74	85.40	5100.00	0.78
Subject 2	97.30	7700.00	0.75	97.30	7100.00	0.79	96.70	6800.00	0.81
Subject 3	92.30	6100.00	0.68	94.40	4900.00	0.77	90.60	4700.00	0.79
Subject 4	99.50	6100.00	0.74	99.20	5400.00	0.73	96.80	5300.00	0.79
Subject 5	87.10	6200.00	0.79	96.40	6400.00	1.37	87.40	6200.00	0.76
Subject 6	88.60	8000.00	0.74	93.60	7100.00	0.77	92.80	7300.00	0.89
Subject 7	98.00	6300.00	0.75	97.00	6300.00	0.73	97.50	6200.00	0.78
Subject 8	96.20	10000.00	0.75	97.40	8900.00	0.80	95.90	8500.00	0.90
Subject 9	84.00	8500.00	0.86	90.60	8000.00	1.08	87.40	7600.00	0.86
Subject 10	93.50	6900.00	0.75	95.20	6400.00	0.81	89.20	6300.00	0.81
Average	**91.94**	**7240.00**	**0.75**	**94.81**	**6620.00**	**0.86**	**91.97**	**6400.00**	**0.82**

5.2.1 Identification results using all the features

Initially, the original 18D feature space was evaluated for identification generalization accuracy estimates. Table 5.2 reports the predictive model performance using all 18 features. Table 5.2 demonstrates that using all the 18 features the nonlinear polynomial d = 2 kernel SVM classifier achieved the highest average prediction accuracy of 94.81%, which shows that some critical

discriminative information underlies in the nonlinear feature subspaces. It generated the results with an average prediction speed of 6620 observations per second and average training time of 0.86 s.

Table 5.3: 10-fold cross validation (10-CV) identification results using linear and kernel-based support vector machine (SVM) including accuracy, precision, recall and F-Score generated with 18D feature space.

Subjects	Performance Measures in %											
	Support Vector Machine											
	Linear SVM				Poly d=2				Radial basis function			
	Accuracy	Precision	Recall	F-Score	Accuracy	Precision	Recall	F-Score	Accuracy	Precision	Recall	F-Score
Subject 1	82.90	75.57	80.85	83.44	87.00	84.42	84.49	84.69	85.40	76.47	93.46	84.12
Subject 2	97.30	95.70	99.63	97.62	97.30	96.36	98.88	97.61	96.70	95.99	98.13	97.05
Subject 3	92.30	84.71	100.00	91.72	94.40	88.82	99.31	93.77	90.60	81.82	100.00	90.00
Subject 4	99.50	99.33	99.33	99.33	99.20	98.04	100.00	99.01	96.80	97.26	94.67	95.95
Subject 5	87.10	84.42	88.42	86.38	96.40	95.34	96.84	96.08	87.40	84.16	89.47	86.73
Subject 6	88.60	75.15	90.71	82.20	93.60	85.16	94.29	89.49	92.80	85.23	90.71	87.89
Subject 7	98.00	98.31	97.22	97.77	97.00	96.15	97.22	96.69	97.50	98.85	95.56	97.18
Subject 8	96.20	96.67	98.51	97.58	97.40	96.91	99.79	98.33	95.90	95.33	99.58	97.40
Subject 9	84.00	79.78	96.00	87.14	90.60	86.98	98.00	92.16	87.40	84.78	94.67	89.45
Subject 10	93.50	93.47	95.02	94.24	95.20	96.22	95.02	95.62	89.20	90.50	90.87	90.68
Average	**91.94**	**88.31**	**94.57**	**91.74**	**94.81**	**92.44**	**96.38**	**94.35**	**91.97**	**89.04**	**94.71**	**91.65**

From analyzing Table 5.2 ,the polynomial kernel classifier also demonstrated an average precision of 92.44% implying low false positive rate, recall of 96.38% implying even lower false negative rate, and f-score of 94.35%, the harmonic mean of the precision and recall, which are consistent with the accuracy results. It can be observed that linear SVM with an average accuracy of 91.94% was the fastest classifier achieving an average prediction speed of 7240 observations per second and an average training time of 0.75 s with a marginal decrease in average accuracy of 2.87%. This shows promising evidence that the generated feature space was effective in identifying the inattention state of the driver subjects even in the case of linear prediction.

5.2.2 Identification Results Using SVM-RFE Selected Feature Subset

To overcome the redundancy in the feature space, and computational complexity of the predictive learner due to the high dimensionality of the feature space, we employed SVM-RFE feature selection method. SVM-RFE also ranks the original feature space based on the feature correlation to the class label. It iterates through all the features calculating the correlation between the feature and class label and assigns a weight based on the level of correlation to the class label as described

in Section 5.1.4. Since the feature selection performed by SVM-RFE is a binary class technique, we considered two scenarios, namely (1) normal vs. phone and (2) normal vs. text. Scenario (1) was considered as cognitive distraction and scenario (2) was considered as cognitive and visual distraction. Table 5.4 shows the SVM-RFE ranking for normal vs. phone scenario and Table 5.5 shows the ranking for scenario normal vs. text.

Table 5.4: Support vector machine-recursive feature elimination (SVM-RFE) feature ranking for normal vs. phone.

Subjects	**Normal vs Phone**																	
	Rank																	
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
Subject 1	6	13	15	17	14	18	16	9	5	1	12	7	8	3	3	4	2	10
Subject 2	16	15	6	17	14	9	18	13	1	5	12	7	11	8	3	2	4	10
Subject 3	16	15	6	14	17	18	9	13	12	5	1	7	8	11	3	2	4	10
Subject 4	15	13	16	14	6	18	9	17	1	5	7	12	11	8	3	2	4	10
Subject 5	1	15	6	17	16	14	18	9	13	5	12	7	11	8	3	2	10	4
Subject 6	15	13	17	6	14	18	9	16	5	1	12	2	4	8	3	11	7	10
Subject 7	6	15	14	16	9	18	17	13	12	11	7	5	1	4	2	3	8	10
Subject 8	6	15	13	14	9	16	18	17	12	5	1	7	11	8	3	4	2	10
Subject 9	6	17	1	15	14	18	13	16	9	5	12	7	11	8	3	2	10	4
Subject 10	15	6	14	17	18	13	9	16	5	12	1	11	7	2	4	3	8	10
Frequent	**6**	**15**	**6**	**17**	**14**	**18**	**9**	**13**	**5**	**5**	**12**	**7**	**11**	**8**	**3**	**2**	**4**	**10**

Table 5.5:Support vector machine-recursive feature elimination (SVM-RFE) feature ranking for normal vs. text

Subjects	**Normal vs Text**																	
	Rank																	
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
Subject 1	15	6	17	14	18	9	16	13	5	1	12	7	11	3	8	2	10	4
Subject 2	15	6	17	14	18	16	9	13	12	7	11	5	1	3	8	10	2	4
Subject 3	5	6	15	17	18	14	16	9	13	1	12	7	11	3	4	2	8	10
Subject 4	17	15	13	18	14	6	16	9	1	12	5	7	11	3	8	2	10	4
Subject 5	6	17	15	16	14	18	9	13	5	1	12	7	11	3	8	2	4	10
Subject 6	17	16	6	15	18	14	9	13	1	2	4	12	5	8	3	11	7	10
Subject 7	15	17	9	14	18	16	13	12	8	11	10	7	6	5	4	3	2	1
Subject 8	5	15	6	17	14	18	16	9	13	12	7	8	11	10	4	3	2	1
Subject 9	5	6	17	15	18	14	9	16	13	1	12	7	11	8	3	2	4	10
Subject 10	15	17	12	14	18	13	9	6	16	11	5	1	8	3	2	4	7	10
Frequent	**15**	**6**	**17**	**14**	**18**	**14**	**9**	**13**	**13**	**1**	**12**	**7**	**11**	**3**	**8**	**2**	**2**	**10**

We observed a slight variance in the feature ranks between Table 5.4 and Table 5.5. However, considering subsets of features with various sizes, consistently similar set of features

were selected for both distracted scenarios. In addition, by inspecting subject by subject responses in Table 5.4 and Table 5.5, a high level of similarity in the feature ranks across all subjects is observed. These observations helped to pick a unified set of features that are highly relevant for the task of distraction identification and are consistent between different scenarios and subjects.

Table 5.6: Identification results of linear and kernel-based support vector machine (SVM) including accuracy, prediction speed, and training time with the reduced 7D Feature Space.

Subjects	**Classifier Performance**								
	Support Vector Machine								
	Linear SVM			**Poly d=2**			**Radial basis function**		
	Accuracy	Prediction Speed	Training	Accuracy	Prediction Speed	Training	Accuracy	Prediction Speed	Training
	in %	in ~ obs/sec	Time in sec	in %	in ~ obs/sec	Time in sec	in %	in ~ obs/sec	Time in sec
Subject 1	79.70	7000.00	0.68	85.60	6200.00	1.42	84.60	6300.00	0.69
Subject 2	97.10	9200.00	0.68	96.30	8600.00	1.89	97.30	8000.00	0.76
Subject 3	85.80	6600.00	0.75	91.20	5700.00	1.39	88.50	5800.00	0.71
Subject 4	99.70	7300.00	0.66	99.70	5700.00	0.79	98.70	5000.00	0.77
Subject 5	75.00	7800.00	1.27	95.40	7300.00	3.70	88.80	6500.00	0.78
Subject 6	82.90	8800.00	0.75	88.60	8700.00	0.76	86.80	8000.00	0.70
Subject 7	94.00	7900.00	0.65	96.30	7100.00	0.69	95.80	5500.00	0.78
Subject 8	92.50	12000.00	0.78	93.80	10000.00	0.77	94.60	9600.00	0.89
Subject 9	81.40	9800.00	1.49	88.00	8400.00	1.13	86.50	8200.00	0.77
Subject 10	95.40	8200.00	0.66	95.20	7200.00	0.77	94.50	7000.00	0.70
Average	**88.35**	**8460.00**	**0.84**	**93.01**	**7490.00**	**1.33**	**91.61**	**6990.00**	**0.76**

Table 5.7: 10-fold cross validation (10-CV) identification results using linear and kernel-based support vector machine (SVM) including accuracy, precision, recall and F-Score with the reduced 7D feature space.

Subjects	**Performance Measures in %**											
	Support Vector Machine											
	Linear SVM				**Poly d=2**				**Radial basis function**			
	Accuracy	Precision	Recall	F-Score	Accuracy	Precision	Recall	F-Score	Accuracy	Precision	Recall	F-Score
Subject 1	79.70	72.94	81.05	76.78	85.60	80.12	86.93	83.39	84.60	75.26	93.46	83.38
Subject 2	97.10	95.36	99.63	97.45	96.30	95.29	98.13	96.69	97.30	95.70	99.63	97.62
Subject 3	85.80	75.26	99.31	85.63	91.20	83.14	99.31	90.51	88.50	78.69	100.00	88.07
Subject 4	99.70	99.34	100.00	99.67	99.70	99.34	100.00	99.67	98.70	98.01	98.67	98.34
Subject 5	75.00	69.16	82.63	75.30	95.40	93.85	96.32	95.06	88.80	83.33	94.74	88.67
Subject 6	82.90	65.41	86.43	74.46	88.60	75.76	89.29	81.97	86.80	75.00	81.43	78.08
Subject 7	94.00	96.43	90.00	93.10	96.30	96.09	95.56	95.82	95.80	95.53	95.00	95.26
Subject 8	92.50	92.93	97.66	95.24	93.80	93.39	98.94	96.08	94.60	93.98	99.36	96.59
Subject 9	81.40	76.94	95.67	85.29	88.00	83.33	98.33	90.21	86.50	81.49	98.33	89.12
Subject 10	95.40	96.62	95.02	95.82	95.20	95.83	95.44	95.63	94.50	94.29	95.85	95.06
Average	**88.35**	**84.04**	**92.74**	**87.87**	**93.01**	**89.61**	**95.83**	**92.50**	**91.61**	**87.13**	**95.65**	**91.02**

The best subset of features was selected by calculating the most frequently occurring features across all subjects. Then first seven frequently occurring features were selected as they demonstrated higher 10-CV identification accuracies compared to other feature subset sizes. It was observed that, despite the fact that the order of first seven features is different for both scenarios, the set of the best seven features were the same for both distraction scenarios. The selected seven best features are listed below with their corresponding feature name: (i) number of peaks in a window; (ii) STFT's 2nd feature; (iii) Katz fractal dimension; and (iv–vii) Auto-regressive features.

Table 5.6 and

Table *5.7* provide the identification results corresponding to the selected subset of 7D feature space in comparison with the results from 18D space. From Table 5.6 results using the subset of features that SVM-RFE selected, it is observed that it was possible to achieve the average identification accuracy of 93.01% with a prediction speed of 7490 observations per second using polynomial kernel SVM classifier. When compared to the original 18D feature space, the results showed a marginal decrease of 1.8% in identification accuracy while providing a considerable improvement in the prediction speed (increase of 870 observations per second).

From

Table 5.7, we can see that the predictive model generated after reducing the feature space to 7D also demonstrated similar performance as the 18D predictive model for precision, recall and f-score measure of 89.61%, 95.83% and 92.50%, respectively. This minimal performance decrease in the identification accuracy from the original 18D feature space to 7D is a trade-off that results in improved response time for the online identification task in hand. Reducing the dimensions of the feature space can also potentially alleviate the effect of curse of dimensionality and increase the robustness of the predictive model in the generalization phase.

Table 5.6 demonstrates the nonlinear polynomial and RBF kernel SVM classifier generated the highest average accuracy of 93.01% and 91.61% respectively compared to linear SVM classifier with 88.35%. From

Table *5.7*, it can be observed that polynomial kernel SVM demonstrates the highest average precision, recall and f-score of 89.61%, 95.83% and 92.50%, respectively, which is consistent with the results based on the original feature space. These observations indicate that the selected subsets of features did not disrupt the discriminative capabilities of the original feature space by following

similar performance trends to identify the distracted state from the non-distracted state with minimal decrease in the prediction accuracy and a considerable gain in response time.

5.3 Conclusion

In this thesis, a continuous measure of phasic GSR component was investigated that can be used to identify distracted driving state under naturalistic driving conditions using a wrist band wearable. In contrast to other state-of-the-art driver monitoring and alerting systems using intrusive physiological signal measures such as EEG and ECG, GSRs are minimally intrusive. In this thesis GSR was evaluated toward real-time identification of distraction using short-term segmented windows. Since raw GSR is composed of two components (tonic and phasic), decomposition on the raw GSR was conducted to obtain a continuous phasic GSR signal that contained the most discriminative characteristics to identify driver distraction. Then a high-resolution spectrogram of the phasic GSR signal were generated to visualize and better understand its behavior. 18 features were extracted that included spectral and temporal features bases on the analysis performed in the previous chapters, that capture the pattern changes from normal to distracted scenarios at a physiological level. The nature of feature extraction is manual and might include redundancies that can increase the computational complexity and decrease the robustness of the predictor. Therefore, SVM-RFE technique was applied to alleviate this limitation. SVM-RFE then selected seven features based on the rank assigned to the features that best characterized the distracted state from the non-distracted state. Linear and kernel-based SVM with 10-fold cross validation (10-CV) was used to generate identification results on both the original 18D and reduced 7D feature space. Further investigating the results demonstrated marginal reduction in prediction accuracy and considerable increase in the prediction speed. The results across the population of subjects also demonstrated a high level of consistency.

In thesis two approaches were employed to deal with curse of dimensionality i) linear discriminant analysis dimensionality reduction and ii) feature selection using SVM-RFE. Although the dimensionality reduction technique significantly reduced the dimension of the feature space the feature specific information was lost due to the transformation of the space. However, using the feature selection technique preserved the information about the useful feature and demonstrated better accuracies as well. The proposed driver monitoring and identification system on the edge in this thesis provided evident results using GSR as a reliable indicator of driver distraction while meeting the requirement of early notification of distraction state to driver.

Bibliography

Alizadeh, V. & Dehzangi, O., 2016. The impact of secondary tasks on drivers during naturalistic driving: Analysis of EEG dynamics. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, pp. 2493–2499. Available at: http://ieeexplore.ieee.org/document/7795957/ [Accessed June 26, 2017].

Ayata, D., Yaslan, Y. & Kamasak, M., 2017. Emotion recognition via random forest and galvanic skin response: Comparison of time based feature sets, window sizes and wavelet approaches. In *2016 Medical Technologies National Conference, TIPTEKNO 2016*.

Begum, S., 2013. Intelligent driver monitoring systems based on physiological sensor signals: A review. In *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*. pp. 282–289.

Benedek, M. & Kaernbach, C., 2010. A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, 190(1), pp.80–91.

Box, S., 2009. New data from Virginia Tech Transportation Institute provides insight into cell phone use and driving distraction.

Chen, L. lan et al., 2017. Detecting driving stress in physiological signals based on multimodal feature analysis and kernel classifiers. *Expert Systems with Applications*, 85, pp.279–291.

Ciabattoni, L. et al., Real-time mental stress detection based on smartwatch.

Dehzangi, O., Rajendra, V. & Taherisadr, M., 2018. Wearable driver distraction identification on-the-road via continuous decomposition of galvanic skin responses. *Sensors (Switzerland)*, 18(2).

Dehzangi, O. & Williams, C., 2015. Towards multi-modal wearable driver monitoring: Impact of road condition on driver distraction. In *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks, BSN 2015*.

Dehzangi Omid, Rajendra, V., 2017. Wearable Galvanic Skin Response for Characterization and Identification of Distraction During Naturalistic Driving. , pp.1–14.

Deshmukh, S. V. & Dehzangi, O., 2017. ECG-Based Driver Distraction Identification Using Wavelet Packet Transform and Discriminative Kernel-Based Features. In *2017 IEEE International Conference on Smart Computing, SMARTCOMP 2017*.

Donmez, B., Boyle, L. & Lee, J.D., 2003. Taxonomy of Mitigation Strategies for Driver Distraction. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 47, pp.1865–1869.

Esteller, R. et al., 2001. A Comparison of waveform fractal dimension algorithms. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 48(2), pp.177–183.

Farooq, M. & Dehzangi, O., 2017. High accuracy wearable SSVEP detection using feature profiling and dimensionality reduction. In *2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. IEEE, pp. 161–164. Available at: http://ieeexplore.ieee.org/document/7936032/ [Accessed June 26, 2017].

Guyon, I. et al., 2002. Gene selection for cancer classification using support vector machines. *Machine Learning*, 46(1–3), pp.389–422.

Huang, M.L. et al., 2014. SVM-RFE based feature selection and taguchi parameters optimization for multiclass SVM Classifier. *Scientific World Journal*, 2014.

Kher, R. et al., 2010. Implementation of derivative based QRS complex detection methods. In *Proceedings - 2010 3rd International Conference on Biomedical Engineering and Informatics, BMEI 2010*. pp. 927–931.

Kim, J.Y. et al., 2013. Highly reliable driving workload analysis using driver electroencephalogram (EEG) activities during driving. *International Journal of Automotive Technology*, 14(6), pp.965–970.

Kurniawan, H., Maslov, A. V. & Pechenizkiy, M., 2013. Stress detection from speech and galvanic skin response signals. In *Proceedings of CBMS 2013 - 26th IEEE International Symposium on Computer-Based Medical Systems*. pp. 209–214.

Lew, R. et al., 2008. Exploring the Potential of Short-Time Fourier Transforms for Analyzing Skin Conductance and Pupillometry in Real-Time Applications. *Proceedings of the Human Factors and Ergonomics Society 52th Annual Meeting*.

Liang, Y., Reyes, M.L. & Lee, J.D., 2007. Real-time detection of driver cognitive distraction using support vector machines. *IEEE Transactions on Intelligent Transportation Systems*, 8(2), pp.340–350.

Lin, C.T. et al., 2011. Spatial and temporal EEG dynamics of dual-task driving performance. *Journal of NeuroEngineering and Rehabilitation*, 8(1).

Liu, M. et al., 2016. Human Emotion Recognition Based on Galvanic Skin Response Signal Feature Selection and SVM. In *2016 International Conference on Smart City and Systems Engineering (ICSCSE)*. pp. 157–160.

Liu, T. et al., 2015. Detection of Drivers' Distraction Using Semi-Supervised Extreme Learning Machine. In *Proceedings of ELM-2014 Volume 2 SE - 36*. pp. 379–387. Available at: http://link.springer.com/10.1007/978-3-319-14066-7_36%5Cnhttp://dx.doi.org/10.1007/978-3-319-14066-7_36.

Van Der Maaten, L.J.P. & Hinton, G.E., 2008. Visualizing high-dimensional data using t-sne. *Journal of Machine Learning Research*, 9, pp.2579–2605. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=7911431479148734548related:VOiAgwMNy20J.

Metz, B., Schömig, N. & Krüger, H.P., 2011. Attention during visual secondary tasks in driving: Adaptation to the demands of the driving task. *Transportation Research Part F: Traffic Psychology and Behaviour*, 14(5), pp.369–380.

Min, B.-C. et al., 2013. 1G-35 Changes of driving performance and skin conductance level of experienced taxi drivers due to distraction tasks. *The Japanese Journal of Ergonomics*, 49(Supplement), pp.S556–S558.

NHTSA, 2013. National Highway Traffic Safety Administration Preliminary Statement of Policy Concerning Automated Vehicles. *National Highway Traffic Safety Administration*, p.14.

Nourbakhsh, N. et al., 2012. Using galvanic skin response for cognitive load measurement in arithmetic and reading tasks. *Proceedings of the 24th Conference on Australian Computer-Human Interaction OzCHI '12*, pp.420–423. Available at: http://dl.acm.org/citation.cfm?doid=2414536.2414602.

Nourbakhsh, N., Wang, Y. & Chen, F., 2013. GSR and blink features for cognitive load classification. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. pp. 159–166.

Pedersen, F., 1997. Joint time frequency analysis in digital signal processing.

Rajendra, V. & Dehzangi, O., 2017. Detection of distraction under naturalistic driving using Galvanic Skin Responses. In *2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks, BSN 2017*. pp. 157–160.

Regan, M.A. et al., 2008. Sources of driver distraction. *Driver distraction: Theory, effects, and mitigation*, pp.249–279.

Vapnik, V., 1995. Support vector machine. *Machine learning*, 20(3), pp.273–297. Available at: http://mlab.cb.k.u-tokyo.ac.jp/~moris/lecture/cb-mining/4-svm.pdf%5Cnfile:///Users/atb/Documents/Library.papers3/Files/Machine learning 1995 Vapnik.pdf%5Cnpapers3://publication/uuid/5AC02E6D-E5D6-40BC-9188-73C50CB8FCE1.

Victor, T.W., 2005. Keeping Eye and Mind on the Road. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences, 9*, p.83 s. Available at: urn:nbn:se:uu:diva-6241%5Cnhttp://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-6241.

Wege, C., Will, S. & Victor, T., 2013. Eye movement and brake reactions to real world brake-capacity forward collision warnings - A naturalistic driving study. *Accident Analysis and Prevention*, 58, pp.259–270.

World Health Organization, 2015. World report on ageing and health. 2015. *Luxembourg, Luxembourg*, pp.1–260.

Yan, S. et al., 2007. Graph embedding and extensions: A general framework for dimensionality reduction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(1), pp.40–51.

Young, K.L., Lenné, M.G. & Williamson, A.R., 2011. Sensitivity of the lane change test as a measure of in-vehicle system demand. *Applied Ergonomics*, 42(4), pp.611–618.