PAIN ASSESSMENT WITH ELECTRODERMAL ACTIVITY SIGNALS BY MACHINE LEARNING APPLICATIONS

by

Busra Tugce Susam

B.S in Electrical and Electronics Engineering, University of Istanbul, 2013

Submitted to the Graduate Faculty of
the Swanson School of Engineering in partial fulfillment
of the requirements for the degree of
Master of Science

University of Pittsburgh

2017

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This thesis was presented

by

Busra Tugce Susam

It was defended on

November 27th, 2017

and approved by

Murat Akcakaya, Ph.D., Assistant Professor, Department of Electrical and Computer Engineering

Ervin Sejdic, Ph.D., Associate Professor, Department of Electrical and Computer Engineering

Zhi-Hong Mao, Ph.D., Associate Professor, Department of Electrical and Computer Engineering

Thesis Advisor: Murat Akcakaya, Ph.D., Assistant Professor, Department of Electrical and Computer Engineering

PAIN ASSESSMENT WITH ELECTRODERMAL ACTIVITY SIGNALS BY MACHINE LEARNING APPLICATIONS

Busra Tugce Susam, M.S.

University of Pittsburgh, 2017

Objective and automated pain detection has been one of the key concerns of clinical researches for many years. The modern techniques for automated pain detection are based on machine learning methods that distinguish pain condition from no-pain condition mainly relying on the information extracted from heart rate, electromyography, blood volume pressure and recorded video. Moreover, recent studies based on the statistical analyses of the electrodermal activity (EDA) showed that EDA carry salient information about different pain phases in patients. This is mainly due to the fact that EDA signals contain rich information about autonomic nerves system that can identify different stress levels in patients induced due to varying pain levels. Due to noninvasiveness and portability of EDA measurement systems, a pain detection method based on EDA is highly valuable for clinical settings to quickly, accurately and objectively identify the pain levels of patients. In this thesis, we present a novel and highly accurate pain level detection algorithm based on EDA. Specifically, with the aim of distinguishing pain from pain-free conditions with EDA signals, we develop a comprehensive machine learning framework. In this framework, we employ timescale decomposition (TSD) to extract salient statistical features from EDA signal associated with pain. TSD uniquely disintegrates a signal in time domain to capture short and long-term changes over the signal. We characterize these changes in the EDA signals through the computation of mean, variance and entropy over all the decomposed segments of the EDA. We use these statistical and information theoretic identities to form feature vectors from the EDA. The dimensions of the feature vectors are then reduced through principle

component analysis. In order to develop a pain detection method, these reduced features are then used in three different classifiers, linear discriminant analysis (LDA) and support vector machines (SVMs) with linear and radial basis function (RBF) based kernels. We test our algorithm using a dataset which is obtained through our collaboration with University of California San Diego Medical School. The EDA data in this dataset are measured from children who had undergone a laparoscopic appendectomy. Particularly, we aim to identify two phases of pain including acute and ongoing pain. Ongoing pain is as the name suggests in pain which is continuous in a period of time after the surgery while acute pain is defined as the pain induced as a result of a press in the surgical area by the research staff. Our results show that, using linear SVM, acute pain is differentiable with an accuracy of 83.33% (sensitivity=86.67%, specificity=80%) while ongoing pain is distinguishable with an accuracy of 87.5% (sensitivity=87.5%, specificity=87.5%) using LDA. Requiring less setup complexity than the existing methods, these results are comparable to the state-of-the-art pain detection methods.

Keywords: machine learning, electrodermal activity signals, timescale decomposition, ongoing pain assessment, acute pain assessment.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

I would first like to thank my thesis advisor Dr. Murat Akcakaya. The door to Dr. Akcakaya was always open whenever I ran into trouble spot or had a question about my research or writing. He consistently allowed this paper to be my own work, but steered me in the right direction whenever he thought I needed it. I would also like to thank the experts who were involved in this research project: Dr. Richard Vincent Palumbo of the Bouve College of Health Sciences at Northeastern University, Dr. Matthew S. Goodwin of the Bouve College of Health Sciences at Northeastern University, Alex A. Ahmed , and Oliver Wilder-Smith at Northeastern University, Dr. Jeannie S. Huang at University of California San Diego School of Medicine, Dr. Ken Craig at the University of British Columbia. Without their passionate participation and input, my research could not have been successfully conducted. Finally, I must express my very profound gratitude to my parents and to my colleagues for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.

1.0 INTRODUCTION

Electrodermal activity(EDA) is the general term to define autonomic changes in the electrical properties of skin. EDA is a measurement of the activity in eccrine sweat glands, which are innervated by sudomotor nerves [1] [2]. As the sudomotor nerves prompt the production of sweat, the conductivity measured on the skin surface changes [3]. The EDA has two components: tonic and phasic. Tonic EDA is represented by the skin conductance level(SCL) which indicates slow changes on skin conductance. Phasic EDA reflects skin conductance changes(SCR) in skin conductance. The SCR arises within a pre-defined response window (15s) after stimulus onset and satisfying a minimum amplitude criterion (0.05 μS) [4]. In particular, EDA signals are widely used in neuroscience research, including information processing, emotional and cognitive processes, and clinical researches because the EDA signals provide valuable information about the autonomic nerves system response to a broad range of external stimuli [5] [6] [7] [8]. Another reason of its popularity is the ease of obtaining a distinct electrodermal response [3].There are three standard methodologies to record the EDA. The first one is called endosomatic measurement which consists in measuring the potential difference between two skin sites. Although the endosomatic measurement does not require specific amplifiers and circuitry, it is rarely used due to change in external affects, such as the change of the ionic potential in the sweat ducts [3]. The other two methods consists on an exosomatic approach in which a small external current is injected into the skin. In exomatic approaches, two different methods exist; the direct current method(DC) and the alternating current method(AC). DC procedures are the most used method in exosomatic approaches, although AC methods allow measuring capacity changes in the electrodermal responses [9]. In spite of the variety of EDA recording devices which use different measurement techniques(e.g., endosomatic and exosomatic measurement approches), the EDA has a

great potential in ecological scenario if monitoring is done sequentially. As a matter of fact, wearable sensors have become more prominent due to their wireless communication ability, comfort and portability. Although there are several proposed body locations to measure EDA , van Dooren et al. [10] investigated 16 different body locations to measure the EDA response respect to emotional film frames. Findings illustrates that fingers, foot and shoulders are most responsive areas whereas arm, back , armpit and thighbone are less responsive areas. Despite the variety of EDA analysis methods, the conventional and model-based approaches will be described in following sections.

1.0.1 Conventional Approaches:

The goal of conventional approaches is to extract features from EDA signal. A conventional data analysis algorithm can be defined in a sequential way. Data filtering is performed to reduce the observation noise. EDA signals are segmented to capture the peaks that stimuli-evoke. Then, defining the threshold to determine which EDA peaks cannot be considered significant. After the identification of peaks, feature extraction can be performed. These features are usually qualitative and semi-quantitative models [2]. Generally, time-domain features of EDA used to define the overall ANS activity which (i.e.,the sudomotor activity)causes SCRs. According to the knowledge of SCRs arrise within 1-5s after the stimulus, features are usually calculated into time window of 5s. [2] [4] [11].In spite of extracted feature variabilities within the time response windows of 5s, most common features are the number of significant SCRs , maximum value of the tonic curve, mean value or standard deviation of the tonic or phasic components. In frequency domain, Power Spectral Density(PSD) of the EDA signal has been proposed to assess ANS by Posada-Quintero [12]. The aim of this work is to analyze if a similar relationship of low frequency component response(LF, 0.045-0.15 Hz) of heart rate variability (HVR) to sympathetic function exist with EDA signal. In cardiovascular research, the prevalent way to asses the ANS is to examine the PSD of the HRV [13]. The HRV spectrum consists of three components: low-frequency bandwidth (LF, 0.045-0.15 Hz) that is effected by sympathetic and parasympathetic nervous system activity , high-frequency bandwidth(HF, 0.15- 0.4 Hz) which is influenced by parasympathetic ner-

vous system, and very low frequency bandwidth(VRF, 0.0033-0.04 Hz). To assess the ANS balance, the ratio between LF and HF was generally used. However, LF band contains also parasympathetic dynamic, that leads a question of how accurate to use the ratio between LF and HF. the PSD of the EDA was used to investigate the assumption that if EDA can indicate the cardiac and sympathetic nervous system activity in LF , since EDA is directly controlled by sympathetic system. In their study, a main interest was tonic stress responses; therefore ,the two components of EDA signal was considered; the tonic(SCL) and non-specific SCRs (NsSCRs) which is considered a tonic measure. These time-domain features SCL and the NsSCRs are explicitly related to sympathetic nervous system, but highly variable between subjects [2]. Therefore, the PSD of EDA was calculated. After three experimental paradigms, they concluded the sympathetic nervous system activities in EDA signal can be found in the low frequency band by the range of 0.045 - 0.25 Hz. Moreover, they stated the frequency domain analysis of EDA can lead less inter-subject variability as compared to the time domain features. Overall, the PSD analyses of EDA signal revealed that EDA provides a quantitative functional measurement of ANS as compared with the LF power of HVR.

1.0.2 Model-Based Approaches:

The model-based approaches describes the relationship between observable variable(e.g., skin conductance) and unobservable processes (e.g., sympathetic arousal) using mathematical equations that formulate psycho-physiological assumptions [14].Therefore, the model indicates the SC time series, and the independent variable consists of unobservable psychological status (e.g., sudomotor nerve activity).In the analysis of experimental data , the observed SC data (i.e. EDA data) is known but not central state, and we try to estimate the time series of the central state that generated these SC data. In the literature, several model-based approaches have been proposed to explain how ANS activity (in particular the sudomotor activity) causes SCRs using decomposition the phasic signal into individuals SCRs associated with each stimuli. Alexandar et al. [15] proposed the first LTI model for EDA analysis. In their model, SC is the result of a convolution between discrete bursting incidents of the sudomotor nerve activity (SMNA) and a biexponential impulse response

function (IRF) assumed to be known a priori and time invariant In general, the relationship can be formulated by

$$SC = SMNA * IRF \tag{1.1}$$

where SC is skin conductance, SMNA is sudomotor nerve activity,* is the convolution operator , and IRF is a skin conductance impulse response function [2]. Through Alexandar's assumption,the SC time series in only influenced by SMNA not by other confounding factors(i.g. noise).To improve the realibity of the model, Benedek and Kaernbach developed two new models to modify the LTI assumption taking into account the variability in SCR shape. These models are known as the non-negative deconvolution model [4] and the continuous deconvolution model [16]. Both models divide the SMNA into two parts; describing the actual phasic activity and representing EDA variations of different origins (e.g., noise). They adopted a biexponential IRF, called the Bateman function. They estimate a noisy SMNA and then recover a filtered phasic component using a low-pass filter and a subsequent heuristic and prefixed peak-detection scheme. The Bateman function is described by

$$b(t) = e^{\frac{-t}{\tau_2}} + e^{\frac{-t}{\tau_1}} \tag{1.2}$$

where τ_1 and τ_2 are parameters that define the shape of function. The Bateman function is characterized by a steep onset followed by a slow recovery period, controlled by τ_1 and τ_2 [4]. There have been several softwares that use the Bateman equations due to its minimal computational cost , and these softwares are capable of determining the location of SCRs. For instance, Bach et al. [17] displayed the SCRalyze toolbox.Broadly, SCRalyze algorithms use optimization methods to estimate the model input (SMNA) or paramaters that best explain the observed SC data [2]. Another well-known EDA analysis software is Ledalab which is very similar to SCRalyze. Both models use the Bateman equations as an impulse response to locate the SCRs. Whereas SCralzye model was derived by curve-fitting to a large number of data sets, the Ledalab model uses two approaches to invert the peripheral model; Discrete Deconvolution Analysis and Continuous Deconvolution Analysis. [4] [16]. Morever, the SCRalyze requires event labels but the Ledalab does not require event labels , and the Ledalab is computationally expensive due to its optimization process [18]. Recently, Groce et al. [19] proposed method called CvxEDA to estimate SMNA using a convex optimization

approach. The model consists on Bayesian statistics and the observed SC as the sum of components; a slow tonic component;the output of the convolution between an IRF and a compact non-negative SMNA phasic driver; and an additive noise term. This procedure aims to distinguish peaks by convex optimization, maximum a posterior probability and sparsity. This model also uses the Bateman functions and the parameter set of IRF (i.e, τ_1 and τ_2) was optimized though the minimization of cost function given by the sum of the number of points of the phasic components of EDA signal [2] [19].

Overall, EDA signals is used in a wide range of experimental setups because it is easy to measure and providing rich information about autonomic nerves system(ANS) corresponding to a broad range of external stimuli. Specifically, EDA signals are commonly used to determine the levels of arousal related to emotional and cognitive processes [2] [20] [21] [22]. Moreover, there has been growing interest in EDA signals for pain detection. Many studies have been proposed to the use of EDA signal for pain detection non invasiveness and portability of EDA measurement systems [23] [24]. In literature, various face-recognition approaches have been proposed to identify the pain and pain-free phase [25] [26] [27]. Although the high accuracy of pain detection was achieved by those approaches, several researches have used pixel based methods for representation of pain features which can get influenced by scaling, rotation and illumination [28]. To improve the detection of pain and reduce the high computational complexity, our contributions are listed below;

1. Since most of the face-recognition approaches are sensitive to external factors(e.g, heads movement, illumination), and have high computational complexity, EDA signals are used to distinguish pain from pain-free condition to reduce the high computational complexity through pattern recognition approaches.
2. We proposed timescale decomposition (TSD) technique to extract the statistical features from EDA signals. This approach uniquely decomposes a signal in time domain to capture short and long term changes over the signal. Additionally, the graphical representation of TSD allows users to see the possible changes over the signal. That can also reduce the computational time of analysis by defining the interest area over the signal.

2.0 FEATURE EXTRACTION

The fundamental concepts explained in this section include a novel technique for feature extraction called timescale decomposition(TSD), and the framework of TSD, feature extraction on TSD, and the implementation of TSD on EDA signals.

2.1 TIMESCALE DECOMPOSITION

The primary goal of feature extraction methods is to provide comprehensive information about signals. These methods can be defined either as time, frequency or time-frequency domain algorithms. The aim of time-frequency approaches is to explore the energy concentration along the frequency axis at given time instant. Ideally, time-frequency domain approaches would reveal concealed information about the frequency components occurring in at any given time [29]. Most current approaches in time-frequency domain are wavelets and Short-time Fourier transform (STFT). As wavelet analysis uses basis functions called mother wavelet by windows of variable length to assess the power spectrum in shorter and longer window, STFT uses a constant sliding window to identify frequency components at each time instant. STFT is very simple to implement and wavelet analysis provides better time-frequency resolution. However, such techniques have limitations to analyze the intensive longitudinal data (i.e., time series). An analysis of data generally depends on the timescale in which data assessed. To illustrate, if trying to examine an exercise intervention increases heart rate, statistical analysis of 10 milliseconds (i.e, shorter than 1 heart beat) is unlikely to provide accurate answer because the timescale is shorter than the process of changes. On the contrary, one week of heart rate data might be more likely uninformative neither, because

the long-term timescale can occult short-term changes. The plausible timescale to assess the process of interest is somewhere in-between. Thus, flexibility of window length is crucial to capture short and long-term changes on the intensive longitudinal data. STFT uses constant window length and that causes lower time-frequency resolution, Although wavelet analyses might provide better time-frequency resolution using windows of variable length, that depends a basis function [30]. Therefore, STFT with constant window length and wavelet analyses with a challenging basis functions may not be desirable techniques to examine time series. The main objective of timescale decomposition is to capture characteristics of the temporal behavior of time series across all possible time scales and starting positions of intervals based on a user-chosen statistical measure. TSD can also be used for multivariate and univariate analyses, such that mean, standard deviation and correlation.

2.1.1 Timescale Decomposition Work Scheme

TSD is a method designed to simultaneously measure short and long-term changes in time series data. The procedure is a simple extension of the sliding window method, which segments time series data into consecutive, overlapping, fixed length windows, and calculates a given metric (e.g., mean, standard deviation) on each segment. The sliding window procedures result in a continuous measure of change over time based on the given metric and window length. Whereas a sliding windowed approach results in a single vector from a single window length, TSD iterates this procedure, calculating the given metric in all possible window lengths at all possible starting points, and systematically organizes results into a single matrix (see Figure 2.1). The resulting triangular matrix is organized such that consecutive rows differ by a window length of 1, with progressively increasing window lengths. Thus, each row constitutes a continuous sliding windowed analysis of the data, with the top row using the shortest window length, and the bottom row using a single window length equal to the length of the data (i.e., this cell is an aggregate of the complete data set). Consecutive columns differ in their starting position by 1, progressively shifting the starting point of the windows to the next time point. Thus, all windows in a given column have the same starting point, but progressively later ending points. The first cell in a column is an aggregate of the

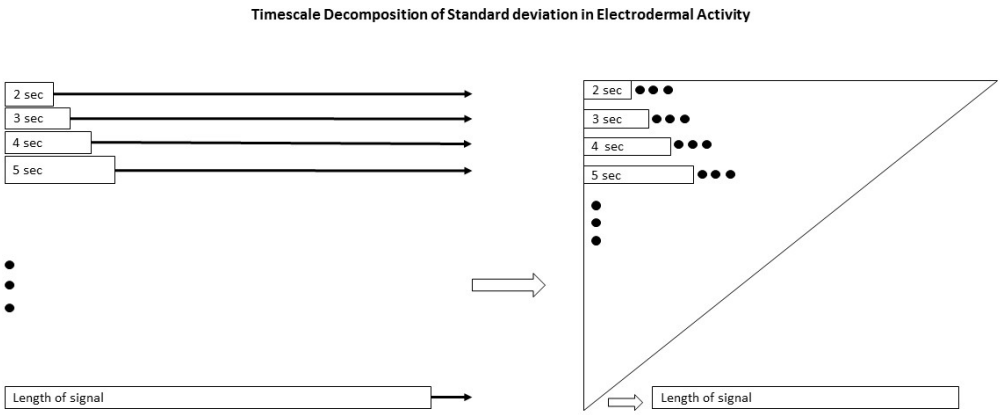

Figure 2.1: The framework of Timescale Decomposition

shortest window length from the given starting point, whereas the last cell in each column is an aggregate of the data from the given starting point to the end of the data. The resulting matrix is therefore a complete decomposition of the data, organized by starting time (left to right) and timescale (top to bottom). To illustrate, we performed TSD on the EDA data using standard deviation (SD) as the metric. First, we took a normalized EDA signal (see Figure 2.2 top), and computed a TSD , then used a heat map to plot the resulting matrix (see Figure 2.2 bottom). Note that the TSD plot shows how the SD changes depending on the timescale (window length) and starting point. Heat maps are graphical techniques that hierarchically represent values in a matrix with colors, allowing simple visual inspection of large amounts of data [31]. Thus, graphical representations of TSD support users in the discovery of a wide range of potentially hidden patterns by taking advantage of the human ability to visually detect such patterns

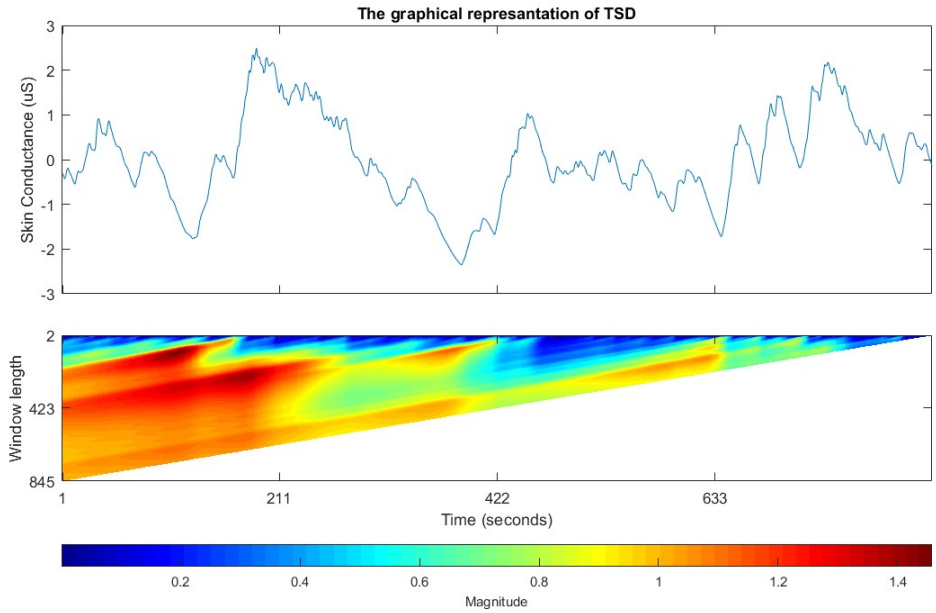

Figure 2.2: The heat-map of Timescale decomposition.

2.1.2 Feature Extraction on Timescale Decomposition

The necessity of having a reasonable amount of informative features requires a feature extraction algorithm on TSD output matrices for machine learning implementations. The feature extraction on TSD output matrix can be performed in two ways; column and row extraction. As a column of TSD output matrix illustrates the calculation of desired metric (e.g., mean, standard deviation and correlation) in all timescales from the given location forward, a row of TSD output matrix indicates the desired statistical measurement of the signal at given timescale. First we extract the desired event times on TSD output matrix. To define statistical characteristic of each row or column of TSD, the basic statistical measurements are used; mean, standard deviation and entropy [32]. For row extraction on TSD output matrix, we calculate the mean , standard deviation(std) and entropy of each row,

and then each mean, std and entropy column matrices is transposed into final feature matrix (See Figure 2.3). Same algorithm is performed on each column of TSD output matrix (See Figure 2.4).

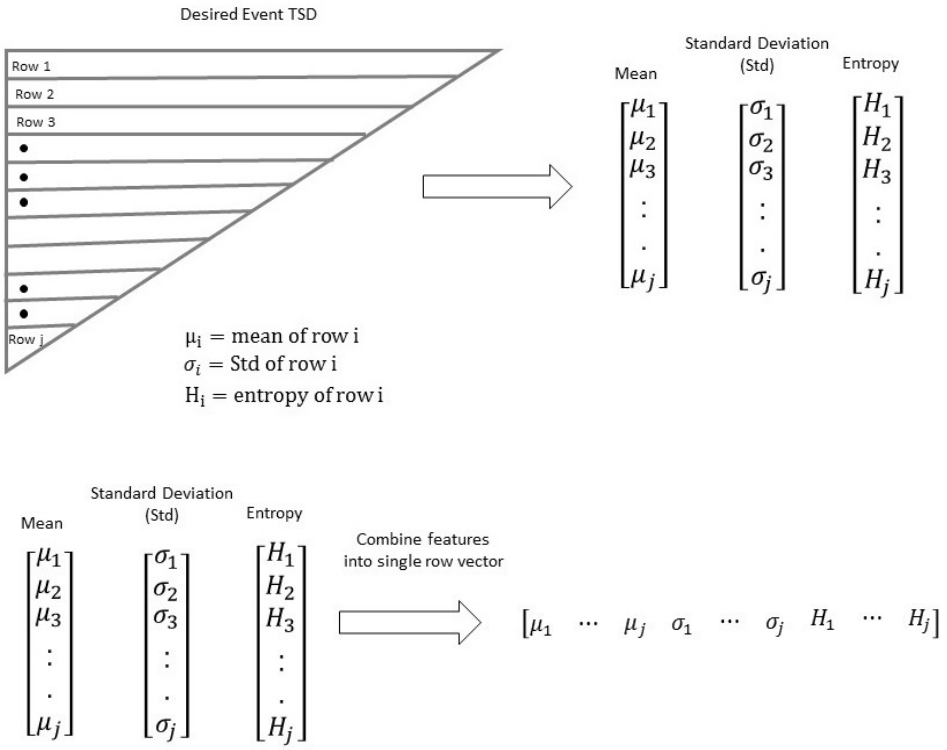

Figure 2.3: The row feature extraction on TSD

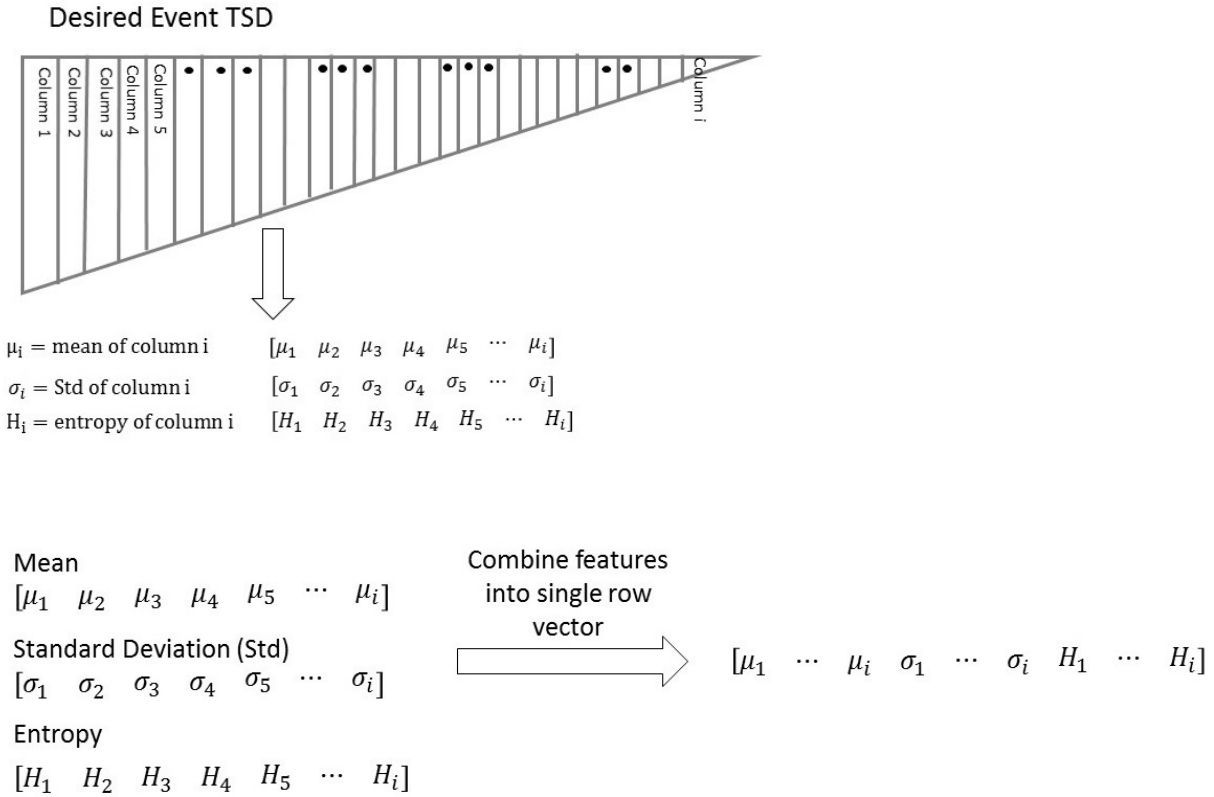

Figure 2.4: The column feature extraction on TSD

2.1.3 Timescale Decomposition implementation on EDA signals

This subsection is dedicated to fully understand the advantages of TSD. To check the feasibility of TSD for pain detection, it was applied to EDA signals obtained to study the role of visual connection in dyads' synchrony of heterosexual couples. PCA and RBF SVM were used for feature reduction and classification respectively.

2.1.3.1 Data set and Experimental setup: A total number of 18 heterosexual couples between the ages of 18 and 22 years (males mean = 19.53, females mean = 18.87) completed the study. Due to a technical issue, data from two trials was excluded, and leaving data from 16 couples available for analysis. Couples reported knowing each other for an average of 3.89 years (SD = 3.62). For the trials, each couple was brought into a quiet room and had physiological sensors, a J&J Engineering I-330-C2+, 12-channel biofeedback unit at a sampling rate of 10 hertz (Hz), to simultaneously measure each partners EDA. Couples were asked to remain quiet for the trials. Participants were seated back-to-back (BB) in separate chairs for the first 17-minutes. The first two minutes was considered an acclimation phase, followed by the BB phase. Following this phase, a tone sounded to alert participants to turn their chairs to face each other. Participants then sat face-to face(FF) in silence for another 15 minutes. A set of 16 random dyads were then generated by pairing the data from a male participant with the data from a random female participant from a different dyad, while controlling for duplicate pairings. The following hypotheses were tested:

1. Correlations between romantic couples' EDA during the BB phase are classifiably different than the FF phase.
2. Correlations between romantic couples' EDA during the BB phase are classifiably different than random dyads.
3. Correlations between romantic couples' EDA during the FF phase are classifiably different than random dyads.

2.1.3.2 Methodology: Each EDA time series was smoothed using a 0.35 Hz low pass filter. This filter setting was designed to remove high frequency components associated with artifacts such as movement, while maintaining all other signal components up to the fastest (∼3 seconds) components of an electrodermal response [21]. After filtering, data was down sampled from 10 Hz to 1 Hz to reduce analysis time and simplify interpretability of the TSD. Finally, EDA data was normalized to have a mean of 0 using z-score normalization. Normalization is particularly important with EDA data, as level is impacted by individual attributes such as skin thickness and density of sweat glands at the recording site [21]. A complete TSD of the correlation between male and female EDA was then calculated for each couple. For each TSD, the starting point,shortest window length, was set to 2 and correlation was calculated in each increasing window until the length of the EDA signal (N = 1,920). Each TSD was then plotted using a heat map, such that positive correlations show

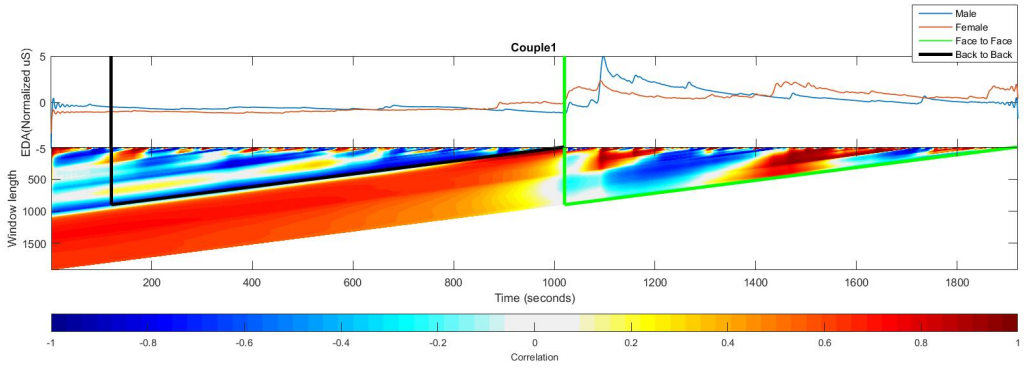

Figure 2.5: The heat map of TSD using correlation in each window

as dark red and negative correlations show as dark blue. As the correlation approaches zero, colors fade to light gray (see Figure 2.5). EDA data was plotted above each heat map and time aligned to the first row of the TSD. To segment areas of the plots exclusive to each condition, we added exclusion onset and offset lines for each condition. All data within the exclusion borders are specific to the given condition, BB phase and FF phase respectively. After, we completed TSD using correlation on each good quality, normalized, filtered, down sampled EDA signal from each participants' EDA signal. We then extracted features from each TSD matrix for use in machine learning classification. First, we separated the segment

areas from the BB and FF phases of each couple's TSD of correlation. This lead to two complete TSDs of correlation for each of the 16 couples, one specific to the BB phase, and one specific to the FF phase, for a total of 32 TSDs. We then calculated the mean, SD, and entropy of each row and each column of each TSD. There are three feature matrices were generated from these data, one specific to row features, one specific to column features, and one combining all features from rows and columns see section 2.1.2. For instance, for couple-1, we first separated the TSD of the complete trial into the data exclusive to the BB and FF phases. For the BB-TSD, we calculated the mean correlation of each row. This results in a single column in which each row equals the mean correlation for the given row of the TSD. This column was then entered into the row features matrix, and the complete features matrix. We then calculated the mean correlation in each column of the BB-TSD for couple-1, resulting in a single row in which each column equals the mean correlation for the given column of the TSD. This row was then transposed and entered into the column features matrix and the complete features matrix. We repeated these steps on the BB-TSD calculating SD, then entropy of all rows and columns, and the same procedures were then performed on the FF-TSD. All steps were repeated for each 16 couple's TSD. Note that the mean of a row indicates the average correlation between a couple's EDA when correlations are calculated at the given timescale, whereas the mean of a column indicates the average correlation across all timescales from the given location forward. The SD of the rows is an estimate of variability in correlations at each timescale, whereas the SD within a column reflects the variability in correlations as timescale increases from the given starting point forward. Entropy is a measure of data structure, such that higher entropy indicates greater randomness, whereas lower entropy indicates greater structure. Therefore, the entropy of rows can be interpreted as a measure of how structured the correlations are over time within a given timescale, whereas the entropy of columns indicates how structured the correlations are across timescales from the given point forward.

A two-class classification scheme was then used to optimize differentiation between features from the BB and FF conditions. This procedure begins with feature reduction, which diminishes the computational complexity of the classification, and leads to a more robust result with lower classification error. Feature reduction was performed using principal compo-

nent analysis (PCA), a well-known and extensively used data reduction technique employing eigenvalues and eigenvectors to transform data into a lower dimensional space [33]. The values of each eigenvector represent a direction, while the corresponding eigenvalue represents the variance of the data along that direction. To keep of most of the variation in the data is crucial and the decision on how many principal components is challenging. Therefore, a threshold was set so that between 90 % and 99% of the original variance remained in the feature matrix [34]. This procedure was repeated with each feature matrix. Finally, classification scheme was completed using Support Vector Machines(SVMs). SVMs, a popular machine learning technique, was used to determine if the two experimental phases, i.e. BB vs FF, could be successfully classified. To determine the class, in our case either BB or FF, SVM attempts to find a hyperplane that can be used to separate the data specific to each class [35]. SVMs offer additional appeals to find the hyperplane; such that, basis function(RBF), polynomial kernels and sigmoid kernels. In this work, RBF SVMs was employed. To avoid overfitting with RBF SVM, leave one out cross validation(LOO) which is a machine learning technique to leave one samples for testing phase, and train the classifier with rest of samples until all samples are used in training set [36]. After the features where extracted and then reduced with PCA LOO removes a single trail which will be used later for testing. The rest of the 31 trails then go into the training phase. During the training phase, a range of box constraints and sigma values were tested to determine the pair that gives the best accuracy. Both parameters were ranged between .001 and 100 initially and then the interval was reduced in secondary runs to optimize the classification. To determine the optimal parameter pair, each combination of parameters was tested using a LOO scheme. To complete the analysis of one parameter pair first a single trail was removed for validation and then training was completed on the remaining 30 trails to determine the decision boundary. The validation point was then tested and the trails class determined. This training and validation was then repeated, keeping the same parameter pair, so that each of the 31 test points were removed and used for validation. Finally, the accuracy for those 31 tests was calculated for the given parameter pair. This procedure was then run for each parameter pair combination resulting in a testing accuracy specific to each pair. The parameter pair that resulted in the best accuracy was then used in the testing phase to train

and find a decision boundary and determine the class of the initial test point. See Figure 2.6 for a visualization of this process. This process, training and test phase, was then repeated 31 times so each point was used as the training trail.

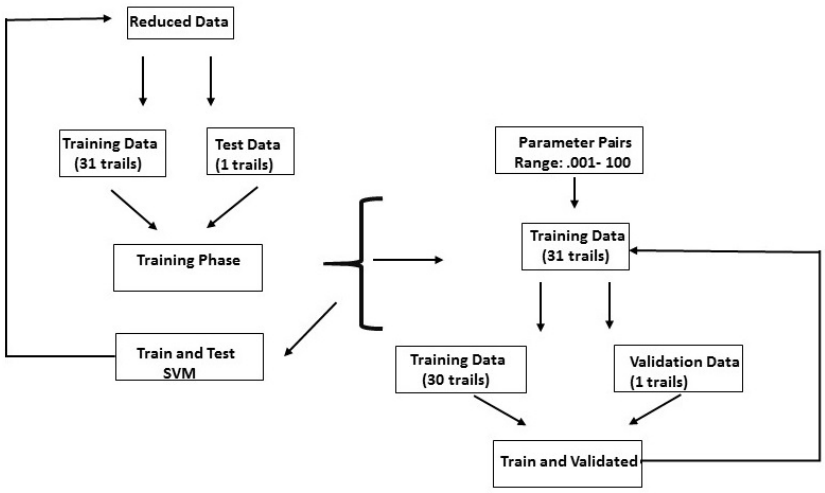

Figure 2.6: The classification scheme of couples' EDA data

These steps were repeated in order to determine whether true couples were classifiably different from random dyads. We first tested the accuracy of differentiating true couples and random dyads using data from the BB phase. Next, we tested the accuracy of differentiating true couples and random dyads using data from the FF phase. Finally, we tested the accuracy of classifying the random dyads' BB and FF phases. To assess whether temporally decomposed data added information beyond a single timescale, we repeated the classification procedures described above using features from individual timescales. As prior research in this area has used windowed correlations in 15-second windows [37], as well as correlations from complete trials [38], we used machine learning approaches to classify the BB and FF periods using features from a 15-second windowed correlation, and again using the overall correlation from each condition.

2.1.3.3 Results: In order to identify couples' EDA would differ by condition, dyad-level features within (rows) and across(column) timescales were extracted to classify the data as originating from BB or FF phase. These features are categorized into three feature matrices, such that row extraction, column extraction and combination of row and column extraction of TSD output matrices in BB and FF phase. After PCA implementation, followed by SVMs classification. Results by row extraction of TSDs indicated an accuracy of 75%, with a sensitivity of 81% and specificity of 50%. Results by column extraction of TSDs illustrated an accuracy of 66%, with a sensitivity of 50% and a specificity of 81%. Finally, the combination of row extracted features and column extracted features' classification demonstrated an accuracy of 66% , with a sensitivity of 56% and a specificity of 75%. Overall, these findings indicate that within timescale (row extraction of TSDs) structures perform better in differentiation of BB and FF phase in true couples. To fully understand the advantages of a proposed method TSD, we ran these procedure by using a single timescale. First, we completed using window lengths of 15 seconds and, then using window length of 15 minutes, both of which was proposed by [37] and [38]. As a result, using 15 seconds window lengths in same classification procedure show poor accuracy (50%) with a sensitivity of 75% and a specificity of 25%. Similarly, the 15 windows length had a low accuracy of 50%, with a sensitivity of 63% and specificity of 38%.

Further analyses were conducted to identify whether the correlation structure were unique to defined condition or the given couples, we ran the same classification on differentiation between random dyads and true dyads in the given conditions. First, we tested the performance of differentiation between random couples and true couples in BB phase. The row extractions of TSD output matrices indicated an accuracy of 63%, with a sensitivity of 44% and a specificity of 75%. Results by the column extraction of TSD matrices demonstrated an accuracy of 91%, with a sensitivity of 88% and a specificity of 94%. Finally, the combination of row and column extraction of TSD matrices illustrated an accuracy of 63%, with a sensitivity of 50% and a specificity of 75%. For FF phase, the row extraction on TSD output matrices led to an accuracy of 53% (sensitivity=63%; specificity=44%), the column extraction resulted in 94% accuracy (sensitivity=88%; specificity=100%) and the combination of row and column features led to an accuracy of 72% (sensitivity=69%;

specificity=75%). Remarkably, these findings demonstrate that cross-timescale(the column extraction on TSDs) was highly accurate in differentiating between true couples and random dyads in both phases. Finally, we attempt to differentiate the defined conditions (BB and FF phase) in random dyads. This led to an accuracy of 63% (sensitivity=75%; specificity=56%) when the combination of rows and columns on TSD was used, an accuracy of 69% (sensitivity=63%; specificity=75%) when only features were considered, and an accuracy of 66% (sensitivity=75%; specificity=56%) when only row feature extraction was used.

2.1.3.4 Discussion and Conclusion: Timescale decomposition was used to decompose time series data in time domain to provide underlying comprehensive information about data set by assessing the all possible time intervals. As timescale can play an important role in the outcome of many analyses and aggregating the data at any one timescale could be misleading, TSD uses all possible time intervals to assess the data by implementing any statistical metric. For example accuracy was poor (50%) in two attempts to classify true couples condition (BB or FF phase) using a correlation from a single timescale (15second and 15 minutes window). On the other hand, the accuracy was increased to 75% when row feature extraction of TSD was used in differentiation of the given conditions (BB and FF phase). Similarly, for both conditions(BB and FF phase), true dyads were differentiated from random couples by using the column extraction of TSD was used with the accuracy of 91% in BB phase and the accuracy of 94% in FF phase. In both conditions, the correlation structures in true couples EDA were different from the random dyads, indicating that correlations between individuals EDA are sufficient to differentiate between people who were together, and people who were under the same conditions with someone else at a different time. Overall, these results suggest that interpersonal physiological relationships develop between partners quietly sitting together, and that synchrony patterns can be used to differentiate between times when they are facing toward and away from each other. In addition, as the current approach generated relatively equal differentiation between true couples and random dyads in both phases whereas prior results did not [39], findings suggests that important information is contained in the temporal structure of the data. Despite the numerous advantages of TSD, several limitations exist in current approach. Due to the number of iterations inherit in TSD

framework, the computational cost of TSD is high. One built-in solution is the adaptation of optimization techniques to reduce the computational cost. In addition, if the data is longer than 30 minutes, the visualization of TSD can be improved to provide more feasible changes over the signal. To solve that problem, changing the scale or color scheme of a TSD heat-map can lead to substantial differences in the appearance of the plots.

Overall, Timescale decomposition was designed to expose temporal structures in univariate and multivariate time series data. The procedure can be used to generate interpretable visualizations, and resulting decomposed data made available for statistical analysis. The simulated and motivating examples presented herein indicate that temporal information can be exposed through TSD. Moreover, TSD greatly improved classification results, suggesting useful information can be obtained by decomposing the time domain with extending sliding window approach. Though additional work is needed to further develop these procedures, their flexibility, simplicity, and interpretability potentiate their future utility.

3.0 PAIN DETECTION WITH EDA SIGNALS THROUGH MACHINE LEARNING ALGORITHMS

This section will provide information about the use of EDA signal in pain assessment through pattern recognition algorithms, beginning with an overview for pain assessment.

3.1 EDA FOR PAIN ANALYSES

3.1.1 Introduction

The meaning of pain has been a widely discussed and debated issue in many research areas. In 1900, Charles Sherington was one of the first to bring attention to the definition of pain. In his critique of pain, Sherrington stated pain as "the psychical adjunct of a protective reflex which arbitrates the reflexive retraction of a body part that saves it from injury [40]." However, Sherington's definition of pain has an enormous caveat in existence of the pain of disease because the psychical adjunct of a protective reflex only describes the good pain. Several experts have expressed doubt about Sherington's definition of pain. Contrary to Sherrighton's pain definition, John Locke has claimed pain is "a behavioral drive triggered by excessive stimulation of any sense of organ [41]." In addition, MacKenzie in 1909 has claimed that pain is a subjective personal experience and can be assessed through communication [42]. MacKenzie's statement also called into question how to measure pain. Therefore, this pain enigma has divided researches into two concepts of pain; the psychical adjunct of a protective reflex and a behavioral drive triggered by excessive stimulation of any sense organ. Afterward, John Bonica , the founder of modern pain medicine, explained the diffe-

rence between protective pain and the dreadful pathological pain [41]. According to Bonica, "acute symptomatic pain serves as the useful purpose of warning, whereas chronic pain is a malefic force which causes severe physical, emotional and economic stress on the patient." Another immense contribution of John Bonica's was the establishment of the International Association for the Study of Pain (IASP) to advise professionals from many disciplines to communicate effectively in terms of the analyses of pain. Until the 1960s, pain was defined by IASP as "an unpleasant sensory and emotional experience associated with actual or potential tissue damage that is always stated to be always subjective and to be best assessed by self-report [43]." In recent years, great advances have occurred in the study of pain. The role of external factors on the patient's body have been investigated such as the effect of past experiences, anxiety, genetic differences and expectations. Hence, pain can be described by four broad categories; nociception, perception of pain, suffering and pain behaviors [44]. Nociception pain is "the detection of tissue damage that leads to pain " while perception of pain can be described by " a noxious stimulus." Suffering is a negative response induced by pain, fear, and other psychological states, whereas pain behaviors can be "ascribed to the presence of tissue damage." Moreover, the types of pain can be divided into three categories; transient pain, acute pain and chronic pain. Transient pain is "elicited by the activation of nociceptive transducers on tissues in the absence of any tissue damage." Acute pain is "evoked by substantial injury of body tissue and activation of nociceptive transducers at the site of local tissue damage" and chronic pain is "the type of pain when the injury exceeds the body's capability for healing [44]." Since the best way to assess pain is self-report, there have been many attempts to accurately measure whether the person is in pain and how much pain he or she feels. The results have been disappointing because most of the attempts revealed that pain threshold and tolerance could be measured rather than absolute pain sensation [41]. Preliminary work on the pain scaling system was conducted by Hardy and Wolff. They generated the pain scaling system based on thresholds of volunteers using radiant heat stimuli generated by a dolorimeter. The ten-point pain scale was measured in units called Dols with 0 meaning no pain at all and 10 meaning the maximum pain imaginable [45]. However, the main weakness in the study is that the proposed pain scaling system offers no explanation for measuring clinically relevant pain. Nevertheless, several attempts

at quantifying human pain lead to a common pain scaling method, the visual analogue scale (VAS) or the numerical rating scale (NRS) [46]. The VAS is a coherent outgrowth of the Dol scale proposed by Hardy and Wollf. It is also based on a ten-point scale to rate pain, but it does not require sophisticated equipment. The VAS has been widely used in many research areas due to its ease of use and reliability [47], however, this pain quantification method requires certain basic cognitive and motor skills which specific individuals can be lack of these skills. Therefore, several researchers have worked on accurate pain measurement in these populations; preverbal children, elderly patients with motor difficulties, and individuals with dementia [48] [49] [50]. An increasing number of pain-related studies have found that pain stimuli increases the activation of sympathetic nervous system, which leads to increased heart rate and skin conductance [46] [51] [52] [53]. As mentioned in Section 1, the EDA signals can be used in the pain assessment analyses due to its connection to the sympathetic nervous system.

Several groups have worked on pain detection through statistical analyses. A brief summary of statistical methods proposed for pain detection is presented in Table 3.1. It is clear from table that there is a positive correlation between sudomotor activity and noxious stimuli. Similarly, noxious stimulation elicited higher heart rate. For instance, Schectatsy et al. [23] has studied the correlation between heat pain perception and skin conductance. In the study, various thermal stimuli protocols (i.e., three different levels of heat stimulus; light warm, high warm, and maximum warm) were conducted in 22 healthy volunteers and sudomotor activity was recorded. Sympathetic skin responses and the mean level of EDA were measured. EDA was analyzed using the VAS scaling pain quantifications. After three different levels of heat stimulus, there was a positive correlation between changes in sudomotor activity and temperature perception. Moreover, the mean EDA was drastically higher during the pain phase when comparing pre-perception warmth to post-perception warmth. Most of the researchers also have worked on the correlation between pain intensity measurements and biomedical signals. As mentioned before, the validity of pain intensity measurements depends on the population. Inspired by the association between noxious stimuli and sudomotor activity, EDA signals can be used for pain detection

Recently, machine learning applications were used for pain detection. As it can be seen from Table 3.2, high accuracy achieved by face recognition algorithms. However, most of researchers have used pixel methods which may distorted by external factors; rotation, scaling and illumination. Additionally, the fusion of biomedical signals were used for pain detection through machine learning applications. For instance, Gruss et al. [54] has worked on the detection of pain induced by heat stimulation. In the study, the fusion of EDA, EMG and EEG signals was used. Next, forward selection algorithm was used for feature selection, and RBF SVM classifier was employed in classification procedure. The accuracy of 79.29% was obtained through the classification of pain threshold versus baseline. Due to noninvasiveness and portability of EDA measurement systems, EDA signals are used to classify the pain by using the proposed feature extraction method (i.e., Timescale decomposition) followed by the advantage of machine learning applications in our study. The assessment of pain is divided into two types of pain, acute pain which is generated by evoked-pressure stimuli and ongoing pain which could be defined as post-operative pain without evoked-stimuli. The objective of this study is to identify pain by using the EDA signals. Acute and ongoing pain is analyzed with pattern recognition algorithms, as well as the role of pain score.

Table 3.1: Brief summary of statistical methods proposed for pain detection

Author	**Patients demographics**	**Modalities**	**Features**	**Measures of pain intensity**	**Stimuli types**	**Statistical analyses**	**Statistical findings**
Eriksson et al. [24]	-Twenty-two healty infants	-EDA signals	-Conductance baseline level -Number of waves per second -Mean amplitude of the waves	-PIPP scoring	-Heel-lancing for blood sampling -Tactile stimulation	-The Mann-Whitney U test	-Amplitude; (p=0.006) -PIIP score changes (p=0.004)
Choo et al. [55]	-Ninenty children undergoing surgery	- EDA signals -Heart rate -Systolic blood pressure -Respiratory rate	-Median NFSC during 60 second intervals	-NRS pain scoring -NRS anxiety scoring -FPS-R scoring	X	-Sperman rank correlation coefficients	-NFSC vs NRS anxiety score (P=0.15, P<0.002) -NFSC vs FPS-R (P=0.89, P<0.03) -AUC for NFSC (≥4) = 62.3% -AUC for HR(≥4) = 50.6% -AUC for NFSC(≥7)= 69.1% -AUC for HR(≥7) =50.3%

Table 3.1 (continued).

Author	**Patients demographics**	**Modalities**	**Features**	**Measures of pain intensity**	**Stimuli types**	**Statistical analyses**	**Statistical findings**
Loggia et al. [46]	-Thirty-nine healthy male	-EDA signals -Heart rate	-Mean -Coefficient of variation	-VAS scoring	-Twelve 6-s heat stimuli of different intensities	-The parametric Pearson correlation coefficient -The non-parametric Spearman's rank correlation coefficient	-SC is highly correlated with pain between subject analyses - HR is highly correlated with pain across subjects
Schestatsky et al. [23]	-Twenty-two healty adults	- EDA signals	-SSR latency -SSR amplitude - Mean level of electrodermal activity	-VAS scoring	-Three different levels of heat stimuli i.e., warm, pain, MaxVas-onset	-Bonferroni's test -Pearson's coefficient -Chi-square test	-Warm onset vs warm onset SSR latency (r=0.61) -Pain onset vs pain onset SSR latency (r=0.65) -Higher percantage of SSR amplitude(78%) in pain phases than in warm phases(58%) -MeanEDA is significantly higher in during the pain phase

HR: Heart rate, NFSC:Number of fluctuations of skin conductance, SSR:Sympathetic skin responses, MeanEDA:mean value of EDA signals, PIPP:Premature infant pain profile, PSPI:Prkachin and Solomon pain intensity, NRS:Numeric rating scale, VAS:visual analog scale, FPS-R: Face Pain Scale-Revised, MaxVas-onset:Maximum sensation onset, AUC= Area under the curve

Table 3.2: Brief summary of the recent methods proposed for pain detection with machine learning framework

Author	**Patients demographics**	**Modalities**	**Feature descriptors**	**Measures of pain intensity**	**Classifier**	**Performance measure**	**Accuracy**
Rathee et al. [28]	-Twenty-five subjects with a shoulder pain	-Recorded video	-DML -TPS	-PSPI scoring	-RBF SVM	-Leave one out CV - 10-fold CV	96%
Lucey et al. [26]	-Twenty-five subjects with a shoulder pain	-Recorded video	-PTS -APP	-PSPI scoring	-Linear SVM	-Leave one out CV - 10-fold CV	78%
Ashraf et al. [27]	-Twenty-five subjects with a shoulder pain	-Recorded video	-S-PTS -C-APP	-OPI scoring	-Linear SVM	-Leave one out CV	82%
Gruss et al. [54]	-Eighty-five healthy subjects	- The combination of EDA ,EMG and EEG signals	-159 statistical features (e.g.,amplitude, frequency etc.)	- VAS scoring to heat stimuli	-Forward selection -RBF SVM	-Training on 75% of the data, testing on 25% of the data -3-fold CV in parameter selection	- Baseline vs pain tolerance thereshold 90.94% - Baseline vs pain threshold 79.29%

DML:Distance metric learning, TPS:Thin plane spline, PSPI:Prkachin and Solomon pain intensity, RBF:Radial basis function, CV:Cross validation, PTS:Normalized shape, APP:Appearance, S-PTS:Similarity normalize shape, C-APP:Canonical appearance, OPI:Observer pain intensity, VAS:visual analog scale, EDA:Electrodermal activity, EMG: Electromyography, EEG:Electroencephalography

3.1.2 Data set and Experimental setup

Data set for our experiment was obtained from University of California San Diego Medical School. Sixty-one neurotypical youths, aged 5-17 years, who had undergone a laparoscopic appendectomy at a pediatric tertiary care center provided assent, along with one parent who provided informed consent to participate in a study examining automated assessment of childrens post-operative pain using video and body sensors. Subjects underwent three study visits during which they were assessed for pain across the recovery period following laparoscopic appendectomy from within 24 hours following surgery (inpatient) to up to 42 days later (outpatient). For the acceptability part of this study, only participating youth and their parents attending the final study visit (N=54) were given the opportunity to provide data. Prior to the last study visit, two youth discontinued participation in the study. Demographics of youth who dropped-out did not statistically significantly differ from the main cohort ($p>0.20$ for age, sex, and racial comparisons). Children and their parents were approached for participation after undergoing surgery. Children and their parents provided written informed assent and consent, respectively, for participation in this study, as approved by the local institutional review board (IRB). The consent process reviewed the fact that research participants would not personally or directly benefit from participation in the research. One hundred dollars was given to the parent for the child and parents participation; the level of compensation was based on prior work and with approval of the local IRB in context of other studies compensation rates. For the main study, electrodermal activity via wrist/hand sensors were collected to assess physiological and behavioral reactions arising from spontaneous ongoing pain when at rest and instigated pain when pressure was experimentally applied to the surgical incision site. Manual pressure was applied on the abdomen adjacent to the surgical incision site for two, 10-second intervals. Then, youths were asked to score their pain from 0 to 10 for ongoing pain assessment and instigated pain assessment.

3.1.3 Methodology

EDA data quality was visually assessed by a doctoral level researcher with expertise in EDA who was blind to condition. As data were collected from two sensors during each trial, comparisons of both signals (e.g., left hand sensor and right-hand sensor data plotted together) were used to confirm the quality of each signal. Data were removed if the signal did not match with standard expectations for electrodermal activity based on recommendations from Dawson [21]. If data from multiple sensors was deemed good quality, one signal was randomly chosen to be used in the analyses. Selected good quality data was then smoothed using a 0.35 Hz low pass filter. This filter setting was designed to remove high frequency components associated with artifacts such as movement, while maintaining all other signal components up to the fastest ($\sim$3 seconds) components of an electrodermal response [21]. After filtering, data was down sampled to 1 Hz to reduce analysis time and simplify interpretation of results. Finally, to account for baseline differences in EDA levels, the data was normalized to have a mean of 0 using z-score normalization. This is an important step, as EDA levels can be inconsistent over time within and between individuals, so change in EDA is more interpretable than raw level. Before feature extraction step, EDA signals were trimmed to a fixed length of thirty seconds in acute pain assessment and a fixed length of 5 minutes in ongoing pain assessment. Then, feature extraction began with timescale decomposition of each EDA signals as mentioned in Section 2.1.1. To capture characteristics of the temporal behavior of EDA, we performed TSD on the EDA data using standard deviation (SD) as the metric. We then extracted features from each TSD matrix for use in machine learning classification. To generate these features, we first computed the mean, standard deviation, and entropy of each row of each TSD and entered them into a single feature matrix as mentioned in Section 2.1.2. This feature matrix was then reduced using principal component analyses (PCA). PCA is a widely used dimensionality reduction technique designed to help achieve optimum parsimony [35]. Features were normalized before PCA implementation. PCA was set to select data which maintained over the 90 % of the variance in the data, and remove all other data. To avoid overfitting, the PCA was applied to training set, then testing feature vector was projected on principal components.These results were used to generate a reduced

feature matrix. Next, we used three widely used classifiers on the reduced feature matrix; radial basis support vector machines (RBF SVM), linear support vector machines (Linear SVM) and linear discriminant analysis(LDA). Finally leave one participant out cross validation (LOPO) was used in training phase, followed by the testing phase to avoid overfitting. Support Vector Machines(SVM) are extensively used in data classification. SVM is technique which transforms the data set to a hyperplane in which data set is separable [56]. A classification task commonly involves separating data into training and testing test. Every single instance in training set includes one target value, class label. The objective of SVMs is to produce a model which is based on the training data set to predict the target values of the test data given only the test data attributes [57] [35].

$$s(v) = sign\left[\sum_{i=1}^{L} \alpha_i y_i K(v_i, v) + b\right] \tag{3.1}$$

where $v_i \in R^n$, $y_i \in -1, 1$ α_i is the Lagrange multiplier of the dual optimization problem , b is bias and $K(v_i, v_j)$ is a kernel which provides variability to a classifier to differentiate the data in different forms. In practice several kernel are used and we adopted the most popular kernels; linear and radial basis kernel (RBF). RBF kernel is given by

$$K_{RBF}(v_i, v_j) = \exp\left(\frac{||v_i - v_j||^2}{2\sigma^2}\right) \tag{3.2}$$

where is σ is the scaling factor and RBF SVM also requires parameter optimization. However, Linear SVMs does not require parameter optimization and the kernel for linear classification is given by

$$K_{Linear}(v_i, v_j) = v_i^T v_j \tag{3.3}$$

where v_i is the training vector. Linear Discriminant Analysis (LDA) is another commonly used machine learning algorithm that consists in finding the projection a hyperplane that minimizes the interclass variance and maximizes the distance between the projected means of the classes [58]. Let $x_1, \ldots x_n \in R^m$ be a set of n data samples belonging to two different class sets. For each class, A and B, sample means can be defined :

$$\tilde{x}_A = \frac{1}{N_A} \sum_{x \in A} x, \tilde{x}_B = \frac{1}{N_B} \sum_{x \in B} x \tag{3.4}$$

where N_A, N_B are the number of samples in each class, respectively. Then, the scatter matrices, positive semidefinite, for each class can be described by equations:

$$S_A = \sum_{x \in A} (x - \tilde{x}_A)(x - \tilde{x}_A)^T, S_B = \sum_{x \in B} (x - \tilde{x}_B)(x - \tilde{x}_B)^T \tag{3.5}$$

where S_A, S_B is the sample variability in each class. Exquisitely, we want to find a hyperplane, defined by the vector ϕ, in which data samples' variances would be minimal. That can be defined as:

$$\min_{\phi}(\phi^T S_A \phi + \phi^T S_B \phi) = \min_{\phi} \phi^T (S_A + S_B)\phi = \min_{\phi} \phi^T S \phi \tag{3.6}$$

where $S = S_A + S_B$, and the scatter matrix between the classes is given by

$$S_{AB} = (\tilde{x}_A - \tilde{x}_B)(\tilde{x}_A - \tilde{x}_B)^T \tag{3.7}$$

Finally, we maximize the distance between the projected means of classes and minimize the interclass variance to find a hyperplane in which classes are separable. That can be described by maximization of Fisher's criterion:

$$\max_{\phi} \mathscr{F}(\phi) = \max_{\phi} \frac{\phi^T S_{AB} \phi}{\phi^T S \phi} \tag{3.8}$$

Detailed classification procedures for each analyses will be described in following sections

3.1.3.1 Acute Pain Analyses Classification Scheme: As mentioned before, filtered, downsampled and normalized signals were entered into TSD. Next, row extraction on TSD matrices was conducted and leave one participant out cross validation was employed. Then, the PCA was applied to training set, and test feature vector was projected on principal components to avoid overfitting. In radial basis support vector machines classifier, leave one out cross validation was used in training phase, followed by the testing phase used to determine the optimal sigma and box constrain. Both parameters were ranged between 0.001 and 100 for the training phase, see Figure 3.1. In linear support vector machines and linear discriminant analyses, TSD was implemented on processed signals, followed by the row extraction on TSD matrices. LOPO was used to avoid overfitting and then PCA was implemented to reduce features (see Figure 3.2). To differentiate the pain from pain-free condition, three scenarios were employed. First, same participants in initial visit and final visit were used to assess the role of pain score the regardless of pain score (N=21). Second, those participants were grouped based on their pain scores during initial manual press to identify which pain score threshold is significant to distinguish pain from pain-free condition. To verify whether pain or environment was classified, additional three scenarios were employed. Preliminary, 22 subjects whose pain score is equal and greater than four during the initial manual press was chosen from initial visit, then randomly 22 participants out of 37 whose pain score is less than four during the initial manual press were selected from final visit session. These data then was used to train the classifier. To create a testing set, 9 participants whose pain score is equal and greater than four during the initial manual press was used from middle visit, then randomly 9 participants whose pain score is less than four during the initial manual press was chosen from middle visit. In this scenario, the classifier was trained by the data set from different environments, and was tested on middle visit. Secondly, all available data from initial visit (N=34) and from final visit (N=37) was used to train the classifier. Then, the classifier was tested on all available data set (N=27) from middle visit. Finally, 10 fold- cross validation was employed on all data set.

10 fold cross validation is a technique to enhance the accuracy and overcome the over-fitting problem. Basically, randomly chosen 10 fold was left out, the classifier was training on the rest of the data, and tested on those 10 folds. Hence, 10 fold-cross validation uses randomly chosen 10 fold, Monte Carlo simulation was employed with 100 runs, and results were averaged.

3.1.3.2 Ongoing Pain Analyses Classification Scheme: Proceed EDA signals were entered into TSD , followed by the row extraction on TSD matrices as aforementioned above. Once, LOPO was conducted, the PCA was implemented to reduce the features. Similar to Acute pain classification, RBF support vector machines was performed on the reduced feature set (see Figure 3.1). For LDA and Linear SVM, please see the classification scheme shown in Figure 3.2. In order to distinguish pain from pain-free condition, three schemes are applied. Initially, same subjects were chosen from initial and final visit to analyze the role of pain score without consideration of their pain scores (N=20). Secondly, those subjects were sorted by their pain score during the resting time. Finally, further analyses were conducted to validate whether pain or environment was classified. First strategy was employed to verify whether pain or environment is classifiable. To create a training set, 13 subjects whose pain score is equal and greater than four during the resting time was selected from initial visit session, then randomly 13 participants out of 36 whose pain score is less than four during the resting time were chosen from final visit session. For testing set, 10 participants whose pain score is equal and greater than four during the resting time was used from middle visit, then randomly 10 participants out of 21 whose pain score is less than four during the resting time was chosen from middle visit. In this scenario, the classifier was trained by the data set from different environments, and was tested on middle visit. Second, all available data from initial visit (N=35) and from final visit (N=36) was used to train the classifier. Then, the classifier was tested on all available data set (N=31) from middle visit. Finally, 10 fold-cross validation was employed on all data set. Similarly, randomly chosen 10 fold was left out, the classifier was training on the rest of the data, and tested on those 10 folds. After that, Monte Carlo simulation was employed with 100 runs, and results were averaged.

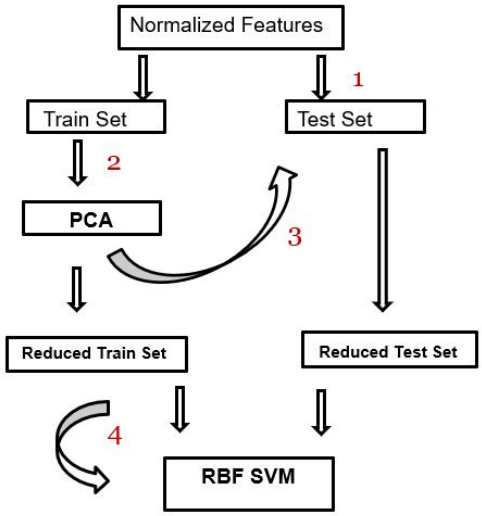

Figure 3.1: The implementation of RBF support vector machines

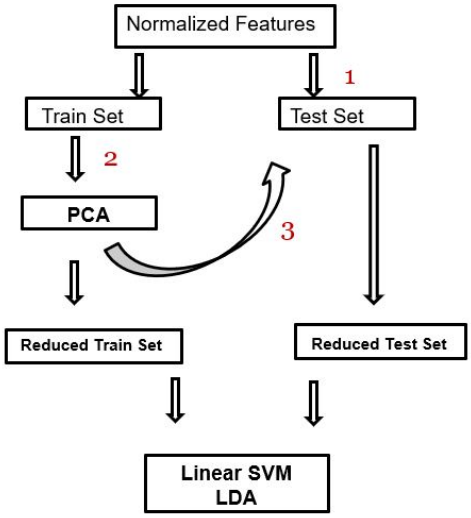

Figure 3.2: The implementation of Linear Discriminant Analysis

3.1.4 Results

In order to identify whether pain would be differ than pain-free condition though EDA signals, proposed feature extraction method TSD was applied to capture the sort and long terms on EDA signals. After that, machine learning algorithms were conducted (see Section 3.1.3.1 and 3.1.3.2). In our procedure, acute pain and ongoing pain was analyzed through proposed feature extraction method and pattern recognition approaches. We also analyzed the role of pain score. Despite of the variety of applied pattern recognition methods , the best results will be illustrated , as follows.

Acute Pain: The first set of analysis investigated whether pain condition would be differ than pain-free condition without the consideration of pain score. Same participants' EDA signals during initial manual press time were assessed from initial visit and finial visit. In spite of the variety of machine learning implementations, linear SVM has performed better by the accuracy of 66.67% with a sensitivity of 61.90% and a specificity of 71.43%. To determine the role of pain score, those participants were grouped by their pain score report. For instance, if pain score threshold is four, the participants whose pain score is equal and greater than four in initial visit, and less than four in final visit were chosen. Figure 3.3 indicates the performance of linear SVM by pain score thresholding criteria as mentioned above. In this analysis, all possible pain score thresholds were considered. The best pain score threshold is four by the accuracy of 83.33%, with a sensitivity 86.67% and a specificity of 80%. Significantly, these findings indicates that when the pain score threshold was considered, the accuracy was enhanced compared to first analyses on the same data set. Therefore pain score threshold was settled to be four in further analyses. Remarkably, these findings also demonstrate that the data set can be generalized since LOOP was applied. To validate whether pain or environmental changes were distinguishable, three strategies were employed. Namely, a classifier was trained and test by the data set from different environments (i.e., initial visit at home, and finial visit at hospital) to check how well the classifier performs in different environment (i.e, middle visit at hospital). First , we look at the affect of randomization on data set to assess the pain from pain-free condition by the consideration of possible environmental changes (e.g., home or hospital), a classifier was

trained and tested by taking a consideration of environmental changes. Results by LDA indicated an accuracy of 72.22% with a sensitivity of 88.89% and a specificity of 55.56%. After that, a classifier was trained by all data set from initial and final visit, and then tested on all the data set from middle visit session. As the data set is unbalanced, it is crucial to beat the chance level which is 66.66%. LDA led to an accuracy of 70.37% (sensitivity=11% and specificity=100%). Although performance is not ideal, we still beat the chance level. Finally, 10-fold cross validation with Monte Carlo simulation was employed on the all available data set. Average results by LDA classifier demonstrated an accuracy of 67.40% with a sensitivity 3.03% and a specificity of 97.18%. Overall, the question regarding to environmental changes are eliminated by the use of randomization affects on the data set. Although, other findings depicted a significant accuracy with a inadequate sensitivity, there is certainly room for improvement.

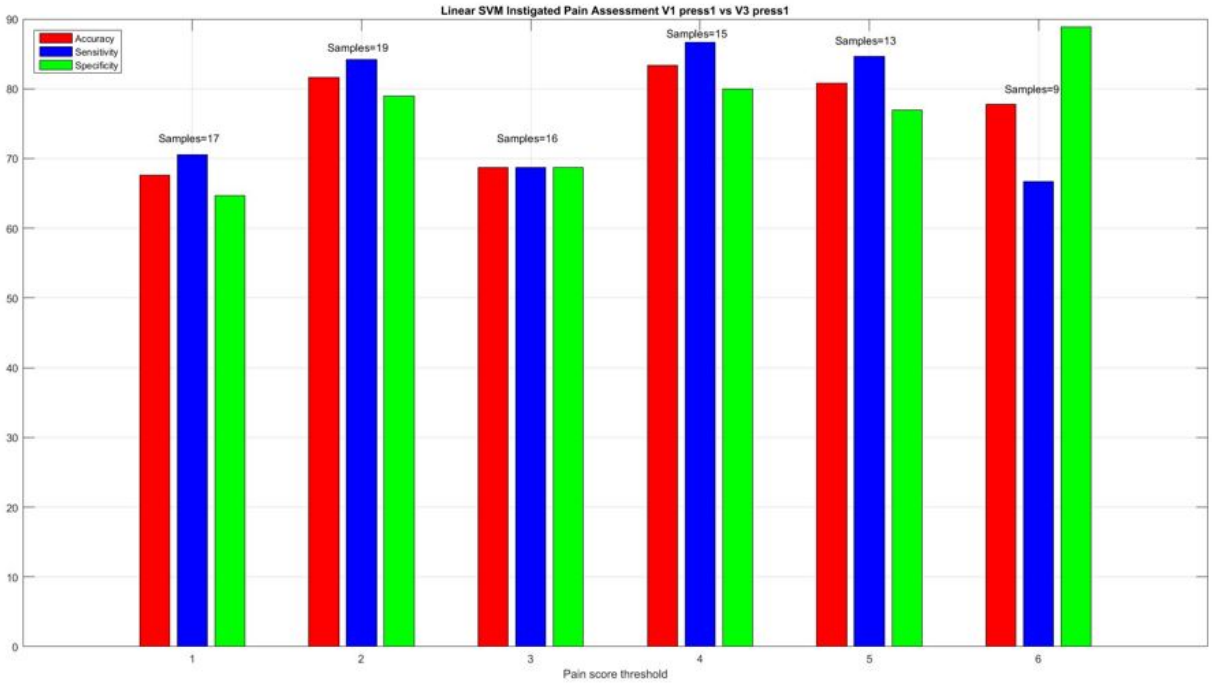

Figure 3.3: Linear SVM performance on acute pain assessment though pain score thresholding

Ongoing Pain: In order to distinguish pain from pain-free condition, same participants during the five minutes resting time were used from initial visit and final visit. Firstly, pain was analyzed without the consideration of pain score report as mention in Section 3.1.3.2. Results by LDA classifier illustrated an accuracy of 72.5% with a sensitivity of 75% and a specificity of 70%. To understand the role of pain score report, same participants were sorted by their pain scores as mentioned above. Then, data set was aligned based on pain score thresholds. This analysis also led to the best threshold being four by implementation of LDA classifier (see Figure 3.4). By pain score threshold four, LDA classifier led an accuracy of 87.5% (sensitivity=87.5% and specificity=87.5%). Therefore, this findings indicated that pain score thresholding approach enhanced the correctly identified pain rate. Additionally, this finding also proved that pain classification can be generalized with the advantage of LOOP cross validation. However, three strategies was applied to a possible issue of environmental changes on pain classification. First of all , data set was balanced using randomly chosen samples. EDA signals during the five minutes resting time was analyzed from initial visit(i.g, signals were recored at hospital) and final visit (i.e., signals were recorded at home). These signals were used in training set, and then EDA signals from middle visit was used to test the classifier. LDA classifier resulted an accuracy of 70% (sensitivity=70% and specificity=70%). After that , all data samples from initial and final visit was used to train a classifier, followed by the use of all data set from middle visit to test the classifier. Hence, data set are unbalanced with a chance level of 67.74%, the Linear SVM provided a best result with an accuracy of 70.97% (sensitivity=30% and specificity=90.48%). Finally, 10-fold cross validation was employed to all data set with Monte Carlo simulation, and results were averaged. Result by Linear SVM depicted an accuracy of 73.52% with a sensitivity of 7.35% and a specificity of 92.78%. Although Linear SVM classifier surpassed the chance level, it did not perform well to classify true positives

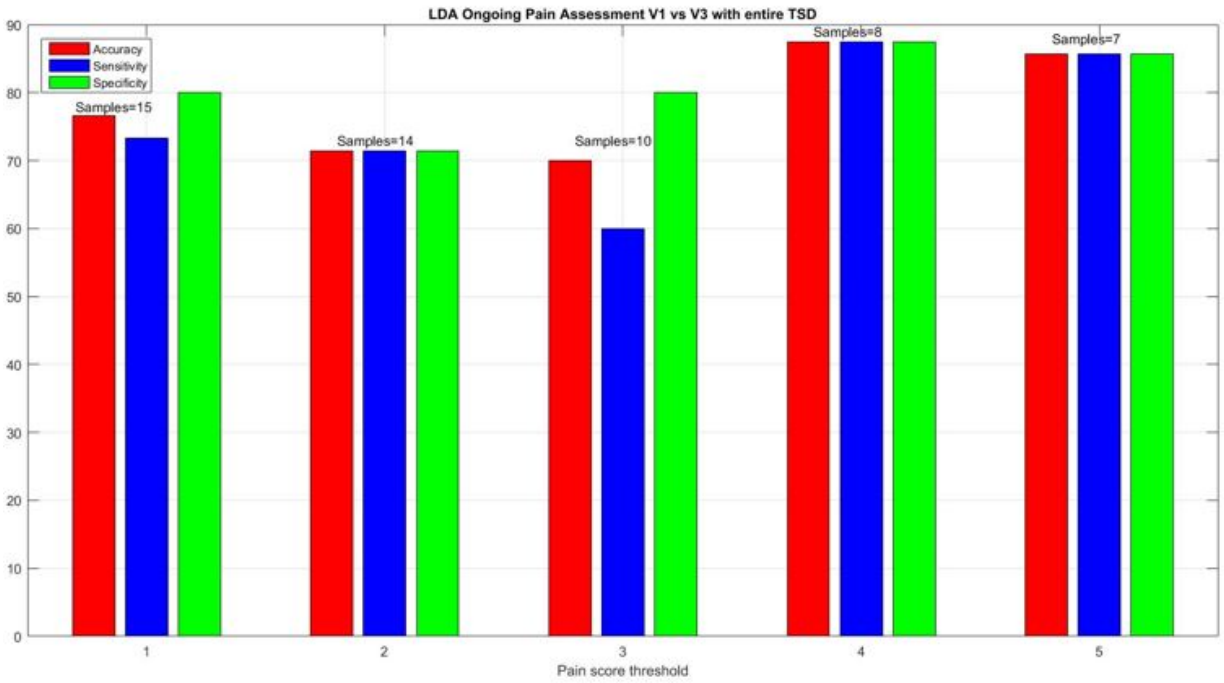

Figure 3.4: LDA performance on ongoing pain assessment though pain score thresholding

3.1.5 Discussion and Conclusion

In this study, pain assessment has been conducted though EDA signals because EDA signals are easy to obtain, and are connected to autonomic nerves system. Despite of the various conventional approaches for EDA analyses, we used our previous feature extraction method Timescale decomposition which exposes EDA signal in time domain to provide comprehensive information about data set by assessing the signal with all possible time intervals. Hence TSD provides more flexibility to seize long and short term changes over the signal with an option of any statistical metric to be assessed, the characteristic of EDA signals in pain phase can be captured adequately. Moreover, the graphical representation of TSD allows users to detect changes over the signal conveniently, therefore, the time of analysis can be reduced. To our knowledge, TSD is the first feature extraction algorithm to investigate the characteristic of EDA signals in given condition with all possible time intervals. As pain detection is

the key concern in many research areas, there are many proposed algorithms to understand the characteristic of pain , intensity of pain and the role of pain scoring systems (e.g, VAS, PIPP). Therefore, we first investigated whether acute pain was distinguishable from pain-free condition by the consideration of pain score report. The most remarkable result to emerge from acute pain investigation is that pain can be differentiable from pain-free condition by an 83.33 % (sensitivity=86.67%, specificity=80%) when pain score was considered. To confirm that hypothesis, all data set was used without the consideration of pain score. As a result, the accuracy of pain detection is lower(66.67%) than the consideration of pain score report scenario(83.33%). This findings illustrated that the accuracy of pain detection was improved by the consideration of pain score report. Similarly, ongoing pain assessment was conducted to investigate whether ongoing pain is differentiable from pain-free condition. The accuracy by the consideration of pain score report is higher (87.5%) than regardless of pain score report scenario (72.5%). This findings confirm the usefulness of pain score report. As the focus of the study was pain detection, there is a possibility that environmental changes have impact on classification procedure. To solve that problem, further analyses were conducted in ongoing and acute pain assessment. When all possible environmental changes were considered, the randomization effect on the data set in acute pain assessment indicated that acute pain still can be distinguishable from pain-free condition in different environments by an accuracy of 72.22%, with a sensitivity of 88.89% and a specificity of 55.56%. Furthermore, the randomization effect on the data set in ongoing pain assessment led to an accuracy of 70%, with a sensitivity of 70% and a specificity 70%. Despite the fact that there was environmental change effects on the data set, pain still can be distinguishable from pain-free condition.

4.0 CONCLUSIONS

In this thesis, EDA signals were used to identify the characteristic of pain due to fact the EDA signals provides salient information about autonomic nerves system, and are easy to obtain. Despite the variety of EDA analysis methods, we have proposed timescale decomposition. This method uniquely decomposes a signal in time domain to capture short and long-term changes over the signal with all possible time intervals. To check the feasibility of TSD for pain detection, it was applied to EDA signals obtained to study the role of visual connection in dyads' synchrony of heterosexual couples. Then, PCA and RBF SVMs were used for feature reduction and classification respectively. Consequently, it was found that TSD improved the classification accuracy when using incremental window compared to moving window by 25% accuracy. Inspired by such significant accuracy improvement, we perform TSD to extract informative statistical features form EDA signal associated with pain. We then characterize these changes in EDA signals through the computation of mean, standard deviation and entropy. These statistical and information identities are used to form feature vectors from the EDA , followed by implementation of PCA to reduce the dimension of these features. To develop a pain detection method, these reduced features are then used in three different classifiers: LDA, RBF SVMs, and linear SVMs. Particularly, the pain classification is divided into two main categories: ongoing pain assessment and acute pain assessment. In the analysis of acute pain, the role of pain score report was analyzed . As a result, acute pain was differentiable by an accuracy of 83.33% (sensitivity=86.67%, specificity=80%) using linear SVM when data set was sorted by pain score threshold. To validate the importance of pain score report, available data set was used without the consideration of pain score. Result by linear SVM classifier indicated acute pain was differentiable by an accuracy of 66.67% (sensitivity=61.90%, specificity=71.43%), this result shows that the consideration of

pain score improved the accuracy of acute pain identification. Despite the fact that there is a limitation due to environmental changes on pain classification, we proposed a technique in which a classifier was trained by the data set from first visit and final visit, and tested on the data set from middle visit to overcome the possibility of classifying environmental changes besides of the pain classification. Randomization affect on the data set led to an accuracy of 72.22% (sensitivity=88.89%, specificity=55.56%). This result also demonstrates that acute pain is still differentiable than pain-free condition in spite of the environmental differences. Similarly, ongoing pain assessment was conducted by the effect of pain score report. When pain score report was considered , ongoing pain was differentiable from pain-free condition with an accuracy of 87.50% (sensitivity=87.50%, specificity=87.50%) using LDA. In order to check the importance of pain score report, all data set was used regardless of pain score report. Results by LDA classifier indicated an accuracy of 72.50% , with a sensitivity of 75.00% and a specificity of 70%. This finding shows that the consideration of pain enhanced the identification of ongoing pain. In order to overcome the issue of environmental differences in pain classification, same machine learning framework was applied to ongoing pain assessment. Results by the randomization effect led an accuracy of 70% , with a sensitivity of 70% and a specificity of 70% using LDA classifier. Therefore, ongoing pain is differentiated from pain-free condition. Additionally, we can state that pain can be differentiable, since LOPO was used in machine learning framework. Namely, a classifier was trained by different participants , and tested on another participant in the implementation of LOPO.

In conclusion, different pain phases can be identified through electrodermal activity signals by the advantage of novel feature extraction method (i.e., timescale decomposition) and pattern recognition algorithms. Although, ongoing and acute pain are classified from pain-free condition, there is a room for improvement of environmental differences issue. This study is first step towards enhancing our understanding of the role of EDA signals in automated pain assessment. Future work will concentrate on whether frequency analyses of EDA signals improve the classification of pain. Additionally, the identification of different pain levels will be considered. To further our research, we plan to use the fusion of EDA signals with video signal as a input of classification procedures to reinforce the accuracy of pain identification.

BIBLIOGRAPHY

[1] H. Critchley and Y. Nagai, *Electrodermal Activity (EDA)*. New York, NY: Springer New York, 2013, pp. 666–669. [Online]. Available: https://doi.org/10.1007/978-1-4419-1005-9_13

[2] A. Greco, G. Valenza, and E. P. Scilingo, *Advances in Electrodermal Activity Processing with Applications for Mental Health: From Heuristic Methods to Convex Optimization*. Springer, 2016.

[3] W. Boucsein, *Electrodermal activity*. Springer Science & Business Media, 2012.

[4] M. Benedek and C. Kaernbach, "Decomposition of skin conductance data by means of nonnegative deconvolution," *Psychophysiology*, vol. 47, no. 4, pp. 647–658, 2010.

[5] H. Sequeira, P. Hot, L. Silvert, and S. Delplanque, "Electrical autonomic correlates of emotion," *International journal of psychophysiology*, vol. 71, no. 1, pp. 50–56, 2009.

[6] A.-A. Dube, M. Duquette, M. Roy, F. Lepore, G. Duncan, and P. Rainville, "Brain activity associated with the electrodermal reactivity to acute heat pain," *Neuroimage*, vol. 45, no. 1, pp. 169–180, 2009.

[7] C. L. Stephens, I. C. Christie, and B. H. Friedman, "Autonomic specificity of basic emotions: Evidence from pattern classification and cluster analysis," *Biological psychology*, vol. 84, no. 3, pp. 463–473, 2010.

[8] A. Greco, A. Lanata, G. Valenza, F. Di Francesco, and E. P. Scilingo, "Gender-specific automatic valence recognition of affective olfactory stimulation through the analysis of the electrodermal activity," in *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*. IEEE, 2016, pp. 399–402.

[9] Ø. Martinsen, S. Grimnes, and O. Sveen, "Dielectric properties of some keratinised tissues. part 1: Stratum corneum and nailin situ," *Medical and Biological Engineering and Computing*, vol. 35, no. 3, pp. 172–176, 1997.

[10] M. van Dooren, J. H. Janssen *et al.*, "Emotional sweating across the body: Comparing 16 different skin conductance measurement locations," *Physiology & behavior*, vol. 106, no. 2, pp. 298–304, 2012.

[11] D. F. Levinson and R. Edelberg, "Scoring criteria for response latency and habituation in electrodermal research: A critique," *Psychophysiology*, vol. 22, no. 4, pp. 417–426, 1985.

[12] H. F. Posada-Quintero, J. P. Florian, A. D. Orjuela-Cañón, T. Aljama-Corrales, S. Charleston-Villalobos, and K. H. Chon, "Power spectral density analysis of electrodermal activity for sympathetic function assessment," *Annals of biomedical engineering*, vol. 44, no. 10, pp. 3124–3135, 2016.

[13] A. Bauer, M. Malik, G. Schmidt, P. Barthel, H. Bonnemeier, I. Cygankiewicz, P. Guzik, F. Lombardi, A. Müller, A. Oto *et al.*, "Heart rate turbulence: standards of measurement, physiological interpretation, and clinical use: International society for holter and noninvasive electrophysiology consensus," *Journal of the American College of Cardiology*, vol. 52, no. 17, pp. 1353–1365, 2008.

[14] D. R. Bach and K. J. Friston, "Model-based analysis of skin conductance responses: Towards causal models in psychophysiology," *Psychophysiology*, vol. 50, no. 1, pp. 15–22, 2013.

[15] D. M. Alexander, C. Trengove, P. Johnston, T. Cooper, J. August, and E. Gordon, "Separating individual skin conductance responses in a short interstimulus-interval paradigm," *Journal of neuroscience methods*, vol. 146, no. 1, pp. 116–123, 2005.

[16] M. Benedek and C. Kaernbach, "A continuous measure of phasic electrodermal activity," *Journal of neuroscience methods*, vol. 190, no. 1, pp. 80–91, 2010.

[17] D. R. Bach, "A head-to-head comparison of scralyze and ledalab, two model-based methods for skin conductance analysis," *Biological psychology*, vol. 103, pp. 63–68, 2014.

[18] M. Kelsey, M. Akcakaya, I. R. Kleckner, R. V. Palumbo, L. F. Barrett, K. S. Quigley, and M. S. Goodwin, "Applications of sparse recovery and dictionary learning to enhance analysis of ambulatory electrodermal activity data," *Biomedical Signal Processing and Control*, vol. 40, pp. 58–70, 2018.

[19] A. Greco, G. Valenza, A. Lanata, E. P. Scilingo, and L. Citi, "cvxeda: A convex optimization approach to electrodermal activity processing," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 4, pp. 797–804, 2016.

[20] J. Wagner, J. Kim, and E. André, "From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification," in *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on.* IEEE, 2005, pp. 940–943.

[21] M. E. Dawson, A. M. Schell, and D. L. Filion, "The electrodermal system," *Handbook of psychophysiology*, vol. 2, pp. 200–223, 2007.

[22] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," *IEEE transactions on pattern analysis and machine intelligence*, vol. 23, no. 10, pp. 1175–1191, 2001.

[23] P. Schestatsky, J. Valls-Solé, J. Costa, L. León, M. Veciana, and M. L. Chaves, "Skin autonomic reactivity to thermoalgesic stimuli," *Clinical Autonomic Research*, vol. 17, no. 6, pp. 349–355, 2007.

[24] M. Eriksson, H. Storm, A. Fremming, and J. Schollin, "Skin conductance compared to a combined behavioural and physiological pain measure in newborn infants," *Acta paediatrica*, vol. 97, no. 1, pp. 27–30, 2008.

[25] P. Lucey, J. F. Cohn, I. Matthews, S. Lucey, S. Sridharan, J. Howlett, and K. M. Prkachin, "Automatically detecting pain in video through facial action units," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 41, no. 3, pp. 664–674, 2011.

[26] P. Lucey, J. Cohn, S. Lucey, S. Sridharan, and K. M. Prkachin, "Automatically detecting action units from faces of pain: Comparing shape and appearance features," in *Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on.* IEEE, 2009, pp. 12–18.

[27] A. B. Ashraf, S. Lucey, J. F. Cohn, T. Chen, Z. Ambadar, K. M. Prkachin, and P. E. Solomon, "The painful face–pain expression recognition using active appearance models," *Image and vision computing*, vol. 27, no. 12, pp. 1788–1796, 2009.

[28] N. Rathee and D. Ganotra, "A novel approach for pain intensity detection based on facial feature deformations," *Journal of Visual Communication and Image Representation*, vol. 33, pp. 247–254, 2015.

[29] L. Cohen, *Time-frequency analysis.* Prentice Hall PTR Englewood Cliffs, NJ:, 1995, vol. 778.

[30] E. Sejdić, I. Djurović, and J. Jiang, "Time–frequency feature representation using energy concentration: An overview of recent advances," *Digital Signal Processing*, vol. 19, no. 1, pp. 153–183, 2009.

[31] W. Akers, "Visual resource monitoring for complex multi–project environments," *International Journal of System of Systems Engineering*, vol. 6, no. 1-2, pp. 112–126, 2015.

[32] D. C. Montgomery, *Introduction to statistical quality control.* John Wiley & Sons (New York), 2009.

[33] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley interdisciplinary reviews: computational statistics*, vol. 2, no. 4, pp. 433–459, 2010.

[34] I. T. Jolliffe, "Principal component analysis and factor analysis," *Principal component analysis*, pp. 150–166, 2002.

[35] S. Theodoridis, K. Koutroumbas *et al.*, "Pattern recognition," *IEEE Transactions on Neural Networks*, vol. 19, no. 2, p. 376, 2008.

[36] C. M. Bishop, *Pattern recognition and machine learning*. springer, 2006.

[37] C. D. Marci, J. Ham, E. Moran, and S. P. Orr, "Physiologic correlates of perceived therapist empathy and social-emotional process during psychotherapy," *The Journal of nervous and mental disease*, vol. 195, no. 2, pp. 103–111, 2007.

[38] J. K. Baker, R. M. Fenning, M. A. Howland, B. R. Baucom, J. Moffitt, and S. A. Erath, "Brief report: A pilot study of parent–child biobehavioral synchrony in autism spectrum disorder," *Journal of autism and developmental disorders*, vol. 45, no. 12, pp. 4140–4146, 2015.

[39] S. Liu, Y. Zhou, R. Palumbo, and J.-L. Wang, "Dynamical correlation: A new method for quantifying synchrony with multivariate intensive longitudinal data." *Psychological methods*, vol. 21, no. 3, p. 291, 2016.

[40] C. Sherrington, *The integrative action of the nervous system*. CUP Archive, 1910.

[41] F. Cervero, *Understanding pain: exploring the perception of pain*. Mit Press, 2012.

[42] D. Sinclair, G. Weddell, and W. Feindel, "Referred pain and associated phenomena," *Brain*, vol. 71, no. 2, pp. 184–211, 1948.

[43] H. Merskey, "Classification of chronic pain," *Description of chronic pain syndromes and definitions of pain terms*, pp. 1–213, 1994.

[44] J. Loeser, "Perspectives on pain," in *Clinical Pharmacology & Therapeutics*. Springer, 1980, pp. 313–316.

[45] J. D. Hardy, H. G. Wolff, and H. Goodell, "Studies on pain: discrimination of differences in intensity of a pain stimulus as a basis of a scale of pain intensity," *Journal of Clinical Investigation*, vol. 26, no. 6, p. 1152, 1947.

[46] M. L. Loggia, M. Juneau, and M. C. Bushnell, "Autonomic responses to heat pain: Heart rate, skin conductance, and their relation to verbal ratings and stimulus intensity," *PAIN*®, vol. 152, no. 3, pp. 592–598, 2011.

[47] R. L. Quiton and J. D. Greenspan, "Across-and within-session variability of ratings of painful contact heat stimuli," *PAIN*®, vol. 137, no. 2, pp. 245–256, 2008.

[48] L. M. Breau, P. J. McGrath, C. Camfield, C. Rosmus, and G. A. Finley, "Preliminary validation of an observational checklist for persons with cognitive impairments and in-

ability to communicate verbally," *Developmental medicine and child neurology*, vol. 42, no. 9, pp. 609–616, 2000.

[49] D. L. LaChapelle, T. Hadjistavropoulos, and K. D. Craig, "Pain measurement in persons with intellectual disabilities," *The Clinical journal of pain*, vol. 15, no. 1, pp. 13–23, 1999.

[50] P. J. McGrath, C. Rosmus, C. Canfield, M. A. Campbell, and A. Hennigar, "Behaviours caregivers use to determine pain in non-verbal, cognitively impaired individuals." *Developmental medicine and child neurology*, vol. 40, no. 5, pp. 340–343, 1998.

[51] G. Hampf, "Influence of cold pain in the hand on skin impedance, heart rate and skin temperature," *Physiology & behavior*, vol. 47, no. 1, pp. 217–218, 1990.

[52] K. C. Kregel, D. R. Seals, and R. Callister, "Sympathetic nervous system activity during skin cooling in humans: relationship to stimulus intensity and pain sensation." *The Journal of Physiology*, vol. 454, no. 1, pp. 359–371, 1992.

[53] A. Möltner, R. Hölzl, and F. Strian, "Heart rate changes as an autonomic component of the pain response," *Pain*, vol. 43, no. 1, pp. 81–89, 1990.

[54] S. Gruss, R. Treister, P. Werner, H. C. Traue, S. Crawcour, A. Andrade, and S. Walter, "Pain intensity recognition rates via biopotential feature patterns with support vector machines," *PloS one*, vol. 10, no. 10, p. e0140330, 2015.

[55] E. K. Choo, W. Magruder, C. J. Montgomery, J. Lim, R. Brant, and J. M. Ansermino, "Skin conductance fluctuations correlate poorly with postoperative self-report pain measures in school-aged children," *The Journal of the American Society of Anesthesiologists*, vol. 113, no. 1, pp. 175–182, 2010.

[56] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[57] C.-W. Hsu, C.-C. Chang, C.-J. Lin *et al.*, "A practical guide to support vector classification," 2003.

[58] P. Xanthopoulos, P. M. Pardalos, and T. B. Trafalis, *Robust data mining*. Springer Science & Business Media, 2012.