Chapter 2

Literature Review

This chapter provides an extensive review of research on EDA signal analysis, offering valuable insights to support the thesis work. It encompasses a comprehensive summary of relevant literature, exploring various aspects, including theories of emotion, applications of emotion recognition, frameworks for emotion analysis, techniques for emotion recognition (both non-physiological and physiological approaches), EDA's role in emotion recognition, decomposition techniques for EDA, the influence of segmentation/windowing techniques, feature extraction methods for EDA, time-frequency representation, and the role of machine learning classifiers in emotion recognition.

Theories of Emotion

The study of emotion clusters across different theoretical frameworks reveals a fascinating journey in understanding the complex landscape of human emotions over time. It all began with William James in 1884, whose pioneering work identified foundational emotions such as rage, fear, grief, and love, emphasizing their deep-rooted connection to bodily involvement. This foundational work paved the way for subsequent theorists to expand and refine our understanding of emotions(Gurney 1884). In 1972, Carroll E. Izard proposed a hardwired model encompassing a broader spectrum of emotions: anger, fear, contempt, disgust, distress, guilt, interest, joy, shame, and surprise (Izard 1972). This comprehensive model reflected a deeper exploration into the innate emotional responses that drive human behavior. Similarly, in 1977, Manfred Clynes broadened the scope by identifying emotional clusters such as joy, grief, anger, hate, reverence, love, and sex, focusing on the depicted expressive behavior that underpins emotional experiences (Clynes 1977).

Building upon these earlier frameworks, in 1980, Robert Plutchik introduced a nuanced categorization of emotions based on their relation to adaptive biological processes(Plutchik 1980). Plutchik's model included emotions such as anger, fear, acceptance, anticipation, disgust, joy, surprise, and sadness, reflecting the intricate interplay between emotions and evolutionary adaptation. In 1982, Jaak Panskepp contributed to the field by proposing a theory centered on hardwired emotions like rage, fear, panic, and expectancy, shedding light on the innate biological underpinnings of emotional responses (Panksepp 1986).

In 1982, Paul Ekman marked a significant milestone in the study of emotions by introducing a more comprehensive set of emotions based on universal facial expressions. Their framework encompassed emotions such as anger, fear, sadness, joy, disgust, and surprise, providing a standardized basis for understanding emotional expressions across cultures Ekman 1992. Nico Frijda, in 1986, introduced a novel framework based on forms of action readiness, categorizing emotions into communication, desire, happiness, interest, wonder, sorrow, and surprise. This framework highlighted the functional aspects of emotions and their role in guiding adaptive behavior (Frijda 1986). These theoretical frameworks provide detailed emotional perspectives, ranging from basic physiological responses to comprehensive conceptualizations that integrate biological, psychological, and social dimensions.

Applications of Emotion Recognition Technologies Across Various Sectors

ER is a crucial branch of artificial intelligence and affective computing, dedicated to identifying and interpreting human emotions across various fields such as healthcare, education, marketing, human-computer interaction, and entertainment. In healthcare, ER technologies provide invaluable assistance to clinicians by decoding patients' emotional states through physiological signals and facial expressions, aiding in the early detection of psychological disorders such as anxiety, depression, and post-traumatic stress disorder, and revolutionizing therapy with personalized treatment plans (Picard et al. 1997), (D'mello and Kory 2015). Similarly, in education, ER fosters dynamic learning environments by discerning students' emotional states and adapting instructional approaches accordingly, thereby enhancing engagement and learning outcomes (Ocumpaugh, Baker, and Biswas 2020).

In the automotive industry, ER technologies serve as guardians of driver safety and comfort by monitoring emotional states to preemptively detect signs of drowsiness, stress, or distraction, ensuring safer road travel (Lee et al. 2023). Furthermore, in security and surveillance, ER systems play a vital role in discerning suspicious behaviors and identifying potential threats in various important places, from airports to public spaces, thereby enhancing or strengthening public safety (Den Uyl and Van Kuilenburg 2005)(Crawford and Paglen 2021).

Additionally, ER tools empower Human resources professionals to understand employees' emotional well-being and interpersonal dynamics, fostering supportive work environments and boosting productivity (Atmaca et al. 2020). Moreover, ER technologies aid coaches and athletes in optimizing sports performance and promoting mental resilience while influencing retail strategies by decoding customers' emotional responses to enhance marketing approaches and drive sales (Laborde, Dosseville, and Allen 2016),(J.-J. Wu and Chang 2020). Ultimately, ER technologies shape human interaction and engagement, enriching our understanding and communication in various facets of life.

Frameworks for Understanding Human Emotions

Emotional models offer structured frameworks for understanding and categorizing human emotions. There are two main types of emotional models: discrete emotional models and dimensional emotional models. In discrete emotional models, emotions can be categorized into distinct, discrete categories with clear boundaries, and they can be distinguished from one another based on specific features or characteristics. Examples of discrete emotional models include basic emotion theories, such as Paul Ekman's six basic emotions (anger, disgust, fear, happiness, sadness, and surprise), which posit that these emotions are universal and evolutionarily adaptive.

On the other hand, dimensional emotion models conceptualize emotions as existing along continuous dimensions rather than discrete categories. These models propose that emotions can be represented in terms of underlying dimensions, such as valence and arousal. Valence refers to the degree of pleasantness or unpleasantness of an emotional experience, while arousal reflects the intensity or activation level of the emotion. Dimensional emotion models suggest that emotions can vary along these dimensions, allowing for a deeper understanding of emotional experiences. An example of a dimensional emotion model is the circumplex model of affect (J. Russell 1980). It organizes emotions based on their position within a two-dimensional space defined by valence and arousal axes. Dimensional emotion models allow for a more detailed analysis of emotions by considering multiple dimensions such as valence, arousal, and dominance, providing a richer understanding of emotional experiences. In contrast, discrete emotional models offer simplicity and clarity by categorizing emotions into distinct labels, aiding in straightforward identification and interpretation.

Two-dimensional (2D) models characterize human emotions along two primary dimensions: valence and arousal. Valence denotes the degree of pleasantness or unpleasantness of an emotion, with positive valence representing pleasant emotions and negative valence indicating unpleasant ones. Meanwhile, arousal signifies the level of physiological activation or intensity associated with an emotion, where high arousal corresponds to heightened physiological responses and low arousal denotes calmer states. By plotting emotions within a 2D space defined by valence and arousal axes, these models offer a comprehensive understanding of emotional experiences. Notable examples of 2D models include the circumplex model of affect, which organizes emotions based on their positions within this space (James A. Russell and Barrett 1999).

In contrast, three-dimensional (3D) emotional models enhance the understanding of emotions by incorporating an additional dimension (Sun et al. 2023). Alongside valence and arousal, these models introduce the dimension of dominance. Dominance represents the sense of control or power associated with an emotion, with high dominance indicating feelings of control or mastery and low dominance suggesting submissiveness or helplessness. By integrating valence, arousal, and dominance dimensions, models, such as the PAD (Pleasure-Arousal-Dominance) model, provide a more detailed representation of emotions (Mehrabian 1996).

While 2D emotional models are more widely used for their simplicity and ease of interpretation, 3D models offer a nuanced understanding by incorporating dominance. Dominance measures a person's level of control over their emotional states, contributing to a more comprehensive emotional analysis. However, the complexity of quantifying dominance often leads to a preference for the simplicity of the 2D valence-arousal framework, known as Russell's 2D model.

In addition to these dimensional models, researchers often use tools like the Self-Assessment Manikin (SAM) to assess subjective emotional experiences. The SAM consists of a series of cartoon figures representing different affective states, allowing the individual to select the figure that best corresponds to their current emotional state along dimensions such as valence and arousal. This tool provides a quick and easy way to measure individuals' emotional responses (Morris 1995).

Researchers have devised various types of emotional triggers to delve into mental and cognitive processes. These triggers include standardized collections encompassing words, images, and audio-visual (AV) stimuli such as video clips, significantly facilitating emotional computing research. These stimuli enable researchers to select appropriate triggers and compare outcomes within controlled laboratory settings. Given their resemblance to human emotional experiences, AV stimuli are crucial in eliciting robust emotional responses during experimental studies (Bălan et al. 2019).

Techniques for Emotion Recognition: Non-Physiological Approaches

ER encompasses diverse techniques to decipher individuals' emotional states across various contexts. Traditionally, ER relied heavily on analyzing non-physiological signals, such as facial expressions, speech patterns, and body gestures (Picard et al. 1997), (D'mello and Kory 2015), (Alarcao and Fonseca 2017). Facial expressions serve as a primary channel for understanding emotions, with pioneering work by Ekman and Friesen (1971) laying the groundwork for objective assessment through the Facial Action Coding System (Ekman 1992). Building upon this foundation, subsequent research has refined facial analysis techniques, leveraging geometric and appearance-based features to enhance emotion classification accuracy (Zeng et al. 2007).

In parallel, the study of body language has gained prominence in non-physiological ER. While facial expressions have historically dominated affect recognition research, recent studies highlight the significance of body gestures in conveying emotional cues. Upper-body movements, including hand gestures during the speech, offer rich sources of emotional signals, with kinematic features like velocity and acceleration playing crucial roles in interpreting emotional displays (Glowinski et al. 2011). Speech analysis represents another pivotal domain within non-physiological ER, offering unique insights into emotional states. Researchers have explored the classification of emotions using speech signals, leveraging spectral analysis and harmonic properties to discern a wide range of emotional states accurately (Glowinski et al. 2011), (M. Chen et al. 2018).

Despite the advantages of non-physiological ER methods, challenges persist in ensuring their reliability and consistency. Individuals may consciously or unconsciously control physical signals like facial expressions or speech, potentially masking their true emotional states (Kroupi, Vesin, and Ebrahimi 2013). Therefore, integrating physiological signals becomes imperative for comprehensively assessing emotional responses.

Techniques for Emotion Recognition: Physiological Approaches

Emotions are rooted in the limbic system, particularly the amygdala, which generates emotional impulses leading to various physiological reactions. These reactions encompass a range of bodily responses, including electric activity in facial muscles, sweat secretion, pupil dilation, changes in breath and Heart Rate (HR), blood pressure alterations, and fluctuations in brain electric activity. These tangible indicators of emotional states can be measured using specialized tools, forming the basis for emotion assessment through physiological signals (Koelstra et al. 2011).

Physiological signals utilized for emotion assessment span the Central Nervous System (CNS) and the Peripheral Nervous System (PNS). Within the realm of CNS signals, EEG emerges as a primary tool in affective computing, providing insights into brain activity patterns associated with different emotional states. Conversely, the PNS offers various modalities for emotion assessment, including ECG, EDA, EMG, BVP, and RSP (Shu, Xie, et al. 2018). Research indicates consistent correlations between peripheral signals and emotional events, highlighting the role of the PNS in emotion processing (Cacioppo, Tassinary, and Berntson 2007), (Levenson 2014). For example, heart rate acceleration is associated with emotions like sadness, rage, and fear, while HR deceleration occurs in response to disgust.

ER research has witnessed remarkable advancements in recent years, driven by exploring physiological signals as indicators of human emotions. These signals range from EEG to HR and skin potential. In this literature survey, we delve into the multifaceted landscape of ER utilizing physiological signals, examining seminal studies and recent developments that underscore these techniques' evolving sophistication and applicability.

Recent research has showcased the remarkable versatility of EEG signals in discerning emotional states. Wang et al. 2018 laid the groundwork by employing EEG signals for three-class emotion classification, differentiating between positive, negative, and neutral emotional states.

Building upon this foundation, Tao et al. 2020 advanced the field further by utilizing EEG signals to classify emotions based on valence and arousal, introducing a nuanced perspective on emotional recognition.

Expanding the scope of EEG-based ER, a study ventured into classifying two distinct sets of emotions. In their study, they successfully categorized emotions into positive, negative, and neutral categories while also extending the classification to encompass a broader spectrum, including neutral, happy, sad, and fearful states. This progression highlights the evolving sophistication and applicability of EEG signal analysis in understanding human emotions, from basic classifications to nuanced emotional distinctions, marking a significant stride toward comprehensive emotional intelligence research (Zhong, D. Wang, and Miao 2020).

In 2018, another study used HR signals to discern emotions such as happiness, sadness, and neutrality. Their experimental setup involved audiovisual movie clips and audio music clips conducted in a controlled indoor laboratory environment, particularly during a walking activity (Quiroz, Geangu, and Yong 2018).

Following this, (Shu, Yu, et al. 2020) investigated HR signals to distinguish between emotional states, specifically neutral, happy, and sad, utilizing the Chinese emotional video stimuli dataset within a controlled indoor laboratory setting.

Umair et al. 2021 conducted a study focusing on Heart Rate Variability (HRV) to assess physiological responses to different states, including baseline, stress, and resting periods, induced through stressor tasks such as the Stroop color test and the trier social stress test, along with cycling activities, within a controlled indoor laboratory environment.

Menghini et al. 2019 explored stress detection utilizing HR and HRV signals, specifically focusing on stress induced by the stress reactivity research task conducted within a laboratory setting with limited physical activity.

In 2017,Shu, Yu, et al. 2020 explored ECG signals to discern emotional states, including valence, arousal, joy, sadness, tension, and peacefulness, with stimuli comprising audio clips conducted within a controlled indoor laboratory environment.

Similarly, Shuhao Chen et al. 2021, Shuhao Chen and collaborators investigated skin potential signals to discern emotional states, including happiness, sadness, anger, and fear, employing video stimuli in a controlled indoor laboratory environment.

Studies also combined multiple modalities for ER. Huang et al. 2018 examined using physiological signals such as ECG and HRV to detect mental fatigue states, utilizing a quiz as a stimulus in a controlled laboratory environment indoors.

Schmidt et al. n.d. conducted a comprehensive exploration on utilizing a diverse range of physiological signals, including accelerometer, ECG, EDA, BVP, EMG, RSP, and SKT, to discern baseline, amusement, stress, meditation, and recovery states induced by stimuli such as video exposure, the trier social stress test, and breathing exercises, conducted within a controlled indoor laboratory environment.

Heinisch, Hübener, and David 2018 examined BVP and SKT signals to identify neutral emotional states, conducting experiments using the International affective picture system as a stimulus in an outdoor limited environment.

Following this, Domínguez-Jiménez et al. 2020 investigated the utility of Photoplethysmography (PPG) and EDA signals in discerning emotional states, specifically amusement, sadness, and neutral, utilizing the FilmStim dataset as stimuli in a controlled indoor laboratory setting.

Xiefeng et al. 2019 conducted a study utilizing ECG, HRV, and heart sound signals to distinguish between emotional states, including relaxed, sad, happy, and angry, employing stimuli such as the International affective picture system, International affective digital sounds, and Chinese affective digital sounds, within a laboratory indoor setting.

Similarly, Siirtola n.d. investigated stress detection using SKT, BVP, HR, and HRV signals elicited by the trier social stress test, conducted in a controlled indoor laboratory environment with specific activities.

F. Li et al. 2020 explored using multiple physiological signals, including electrooculogram, EMG, EDA, BVP, SKT, and RSP, to assess emotional valence and arousal levels involving music videos conducted in a laboratory setting.

Zhang et al. 2020 conducted a study focusing on EDA and HR signals to assess emotional valence and arousal levels, utilizing the Continuously Annotated Signals of Emotion (CASE) and Mobile Emotion Recognition with Continuous Annotation (MERCA) datasets, encompassing indoor and outdoor laboratory settings.

Feng et al. n.d. also investigated using physiological signals, HR, EDA, SKT, and walking steps to assess individuals' well-being based on the positive emotion, engagement, relationships, meaning, and accomplishment framework involving outdoor daily activities.

Similarly, Miranda et al. 2021 explored fear detection using a combination of ECG, EDA, and SKT signals, eliciting fear responses through video clips within a controlled indoor laboratory environment.

Modalities such as EEG and EMG have limitations in emotion detection due to susceptibility to artifacts and noise, as well as challenges in interpretation and equipment requirements. Similarly, HR measurements may lack specificity in distinguishing between emotional states. At the same time, HRV signals can be affected by various physiological factors, leading to potential ambiguity in emotional state classification. SKT measurements may exhibit limited sensitivity to rapid changes in emotional arousal, and BVP signals may face challenges in capturing nuanced emotional responses. Additionally, ECG signals may suffer from noise and artifacts, complicating accurate feature extraction for emotion classification. In contrast, EDA stands out for its high sensitivity to emotional arousal, rapid response time, and non-invasiveness. Leveraging the strengths of EDA alongside complementary modalities can enhance ER systems, enabling more effective applications in various fields.

Electrodermal Activity in Emotion Recognition

EDA serves as a pivotal real-time indicator of emotional arousal, reflecting the activity of the sympathetic division of the autonomic nervous system (Asahina, Poudel, and Hirano 2015), (Benedek and Kaernbach 2010b), (Y. Hu et al. 2018). Research demonstrates that EDA conductance sensitively responds to emotional states, with elevated levels observed during instances of happiness compared to sadness, attributed to sweat secretion by eccrine sweat glands (Khalfa et al. 2008) (Quazi 2012). This characteristic underscores the significant utility of EDA in ER tasks due to its non-invasive nature and ease of recording, facilitated by straightforward and cost-effective instrumentation (Boucsein 2012), (Greco, Valenza, Nardelli, et al. 2016).

The complexity of EDA signals, characterized by randomness, multiple components, and non-stationarity, with spectral variations over time due to anatomical, physiological, and methodological factors, highlights its intricate nature (Boucsein 2012). Notably, areas rich in sudomotor nerve activity, such as the palmar and plantar surfaces, exhibit heightened responsiveness during emotional stimuli, emphasizing the interplay between EDA, the autonomic nervous system, and psychological triggers (Edelberg 1993), (Dawson, Schell, Filion, et al. 2007).

EDA signals encompass two primary components: the tonic and phasic. The tonic component reflects baseline arousal levels and provides insights into overall sympathetic activity influenced by stress, anxiety, and arousal regulation (Posada-Quintero and Chon 2020). Meanwhile, the phasic component, also known as SCR, which mirrors changes in sweat gland activity regulated by sudomotor nerve bursts, emerges as pivotal in emotion detection. Studies indicate that the phasic component demonstrates rapid increases to peaks and gradual declines to baselines, effectively capturing fluctuations in sympathetic arousal and offering valuable insights into dynamic emotional responses (Boucsein 2012), (Posada-Quintero and Chon 2020). Therefore, the non-invasive nature of EDA measurements positions them as a versatile tool in psychophysiological research, facilitating the extraction of perceptual and emotional states from physiological parameters (Bach 2014), (Bach et al. 2011).

Decomposition Techniques for EDA Analysis

In the literature, many deconvolution schemes have been proposed for EDA to decompose into tonic and phasic components. In 1997, Lim et al. introduced an early EDA model that addressed issues with traditional methods that underestimated SCR amplitudes due to short inter-stimulus intervals. This model decomposed each SCR into three components: residual, phasic, and tonic, focusing particularly on analyzing post-stimulus epochs where discrete SCR responses overlapped a sloping baseline. This method provided a framework for separating overlapping responses, enhancing the accuracy of SCR analysis .

Alexander et al. 2005 proposed a model for SCR dynamics based on discrete sudomotor nerve bursts. They utilized deconvolution with a bi-exponential function to obtain a driver function, facilitating the separation of SCRs from overlapping responses. By focusing on the underlying neurophysiological processes, this model improved the ability to extract meaningful information from skin conductance measurements.

Benedek and Kaernbach's Ledalab framework offered two methods: Continuous Decomposition Analysis (CDA) and Deconvolution-Based Decomposition Analysis (DDA). CDA optimized individualized analysis by defining the tonic component and improving parameter estimation (Benedek and Kaernbach 2010a). However, DDA addressed occasional estimation issues by employing nonnegative deconvolution, resulting in more accurate decomposition of SCRs and better interpretation of results (Benedek and Kaernbach 2010b).

Bach et al. 2011 utilized general linear modeling for event-related SCRs, integrating individual differences using a canonical response function derived from principal component analysis. Additionally, they proposed a Dynamic Causal Model that provided insights into the underlying mechanisms of different types of skin conductance responses, allowing for a more comprehensive understanding of SCR dynamics.

Greco, Valenza, Lanata, et al. 2015 proposed a method to decompose SCRs into tonic, phasic, and noise components using a Linear Time-Invariant (LTI) model. By incorporating regularization terms into the optimization process, this method aimed to balance accurately capturing physiological processes while avoiding overfitting, providing a systematic approach to SCR decomposition.

Hernando-Gallego, Luengo, and Artés-Rodríguez 2017 developed a method based on nonnegative sparse deconvolution for decomposing the skin conductance signal. Their approach optimized SCR and SCL coefficients with L1 regularization, refining results through post-processing to enhance the interpretability of the decomposition. This method provided a structured approach to decomposing skin conductance signals and extracting meaningful information.

M. R. Amin and Faghih 2020 proposed a sparse decomposition method involving curve-fitting with a sigmoid-exponential function for the phasic component and a constant term for the tonic component. Manual parameter adjustment based on visual inspection allows user input in modeling. In their recent work Amin and Faghih introduced a new model based on the poral valve model (R. Amin and Faghih 2022). They developed a three-compartment pharmacokinetic model to simulate the skin conductance response process, aiming for a more comprehensive understanding.

These models and methods offer diverse approaches to decomposing skin conductance responses, providing insights into underlying physiological mechanisms and enhancing analysis accuracy. Researchers can extract meaningful information from EDA data by leveraging mathematical modeling and signal processing techniques, contributing to a deeper understanding of human psychophysiology.

Impact of Segmentation/Windowing Techniques on Emotional Classification

Segmentation and windowing are essential methodologies in the field of emotional classification, especially when dealing with physiological signals. Various studies illustrate the application and effectiveness of these techniques in capturing and analyzing emotional states. Zhang et al. 2020 employed segments ranging from 0.125 seconds to 16 seconds to classify emotions into 2-class, 3-class, and 4-class valence and arousal categories. This broad range of segment lengths allows for a detailed analysis of how different temporal resolutions impact the accuracy of emotional classification.

Fiorini et al. 2020 used longer window durations of 120 to 180 seconds to classify emotions into relaxation, positivity, and negativity. The longer windows are beneficial in capturing the overall emotional trends over extended periods, which can be particularly useful in applications like stress monitoring or long-term emotional state assessment.

Polo et al. 2021 focused on the last 100 seconds of signals to classify emotions into amusement, relaxation, and fear (scary). This approach highlights the importance of recent emotional states in the analysis, ensuring that the most relevant and immediate responses are considered.

K. Yang et al. 2021 utilized 10-second signals with 2.5-second windows that had a 50% overlap to distinguish between positive and negative emotions. This method balances the need for temporal precision with sufficient overlap to ensure continuity and context in the data.

In another study, Zhang et al. 2022 explored segments from 0.5 seconds to 30 seconds for valence and arousal classification in 2-class and 3-class systems. This varied segmentation strategy underscores the necessity of evaluating multiple segment lengths to identify the most effective temporal resolution for different classification tasks.

Finally, García-Martínez et al. 2022 concentrated on the second part of EEG signals to classify emotions into high valence high arousal (HVHA), high valence low arousal (HVLA), low valence low arousal (LVLA), and low valence high arousal (LVHA) states. Focusing on specific parts of the signal can enhance the detection of subtle emotional shifts that may not be evident in the entire signal.

These studies show that the way to divide and analyze physiological signals, either segmentation or windowing approaches, makes a big difference in how well emotional classification systems work. It's important to choose the right methods based on the specific emotions you want to identify and the type of physiological signals you are analyzing to get accurate and reliable results.

Feature Extraction of EDA in Emotional Classification

In the literature, researchers explored the diverse methods used to extract features from EDA signals for emotional classification. Traditional approaches rely on temporal statistical features, while newer techniques integrate advanced time-frequency analyses to capture the dynamics of EDA signals. Time-domain features provide a broad view over time, while frequency-domain analysis dissects spectral composition. Time-frequency methods allow for detailed examination across both axes, enhancing the richness of information for classification algorithms. These methodologies extend beyond ER to stress detection and human-computer interaction, highlighting their versatility.

In a study, authors have extracted various features from EDA signals for emotion analysis. Temporal statistical features like average skin resistance, derivatives, local minima/maxima count, mean rising time, spectral power across frequency bands, zero crossing rate, and peak magnitude have been explored (Koelstra et al. 2011), (Soleymani et al. 2011).

Posada-Quintero, Florian, et al. 2018 utilized EDA features to assess cognitive stress in subjects submerged underwater. The extracted features included skin conductance level, non-significant SCRs, EDA sympathetic nervous system component, and time-varying sympathetic nervous system component. These features provided valuable insights into cognitive stress responses in underwater environments, contributing to our understanding of physiological reactions to stressors in challenging conditions.

Zangróniz et al. 2017 utilized a variety of features, including temporal, morphological, and time-domain features, for classifying stress and calm states. The temporal features consisted of metrics such as mean, median, standard deviation, area, maximum peak, minimum peak, dynamic range, mean of the first derivative, mean of the second derivative, standard deviation of the first derivative, and standard deviation of the second derivative. Morphological features included arc length, integral, average power, root mean square, area-perimeter ratio, energy-perimeter ratio, variance, skewness, kurtosis, and spectral bandwidth. This comprehensive feature set was instrumental in distinguishing between stress and calm states, providing valuable insights into classifying emotional states based on physiological signals.

Shukla et al. 2019 utilized time, frequency, and time-frequency features for ER, examining SCR, statistical, Hjorth, and higher-order crossing features. Statistical parameters and wavelet coefficients were investigated for frequency and time-frequency domains. This comprehensive approach enhances ER accuracy by analyzing EDA responses across different temporal and spectral granularities.

Kołodziej et al. 2019 focus on detecting arousal-related emotions by analyzing phasic EDA signals. They employ a diverse set of 20 features extracted from EDA signals, encompassing statistical, frequency domain, and Hjorth features. Among them, features such as MaxAmpPeak (highest value of the determined maxima), VarAmpPeak (Variance of amplitude values calculated for local extremes), StdAmpPeak (standard deviation calculated for local extremes), MaxAbsAmpPeak (maximum value of modules of amplitudes of local extremes), VarSC (variance calculated for skin conductance signal samples), StdSC (standard deviation calculated for skin conductance signal samples), ActivitySC (Hjorth activity that represents the signal power), MaxDeltaForward (Maximum value of the difference between the amplitudes of the local extrema and the amplitude values of the signal samples measured a second later), MaxDeltaBack (Maximum value of the difference between the amplitudes of the local extrema and the amplitude values of the signal samples measured a second earlier), KurtosisAmpPeak (Kurtosis calculated for skin conductance signal samples), and SkewnessAmpPeak (skewness calculated for skin conductance signal samples) emerge as key indicators, reflecting aspects like maximum values, energy, and statistical properties of the phasic component.

Rahman, M. Z. Hossain, and Gedeon 2019 analyzed a comprehensive set of 16 features extracted from EDA signals to identify seven emotional categories. These features include mean, root mean square, the variance of EDA signals, integrated EDA, simple square integral, average amplitude change, Hjorth feature (specifically mobility), Hurst exponent, entropies (Sample entropy, approximate entropy, Shannon's entropy, permutation entropy, fuzzy entropy), log detector, difference absolute standard deviation value, and detrended fluctuation analysis.

Ganapathy, Veeranki, and Swaminathan 2020 classify two classes of arousal and valence by extracting features from time, frequency, and time-frequency domains. They meticulously analyze various features, such as the 21 time domain and morphological features, as well as metrics like area, mean, variance, kurtosis, and skewness, to provide insights into the temporal characteristics of emotional responses. In the frequency domain, features such as mean, median, maximum, and minimum frequency contribute to understanding the spectral composition of emotional signals. Additionally, 11 time-frequency domain features, including flux, flatness, energy concentration measure, normalized Renyi entropy, Shannon entropy, and features based on instantaneous frequency, are scrutinized to uncover time-frequency dynamics in emotional expression.

Ghiasi et al. 2020 employed features from EDA, including tonic and phasic components, along with HRV, to analyze emotions. These features encompassed various aspects of EDA signals. They included metrics such as the number of significant phasic driver peaks, the sum of SCR amplitudes, maximum phasic activity, and mean tonic activity. They also utilized features related to integrating time-frequency planes within specific frequency bands. Additionally, they examined ratios between different components of EDA and HRV, along with median values and area under the curve measurements derived from phasic decompositions. This comprehensive analysis offers valuable insights into the relationship between physiological signals and emotional states.

Lutin et al. 2021 explored various methods for detecting stress by extracting different features. These methods include through-to-peak features, which measure the number of SCRs, the total magnitude of SCRs, the summed duration of SCRs, and the overall area under the SCRs. They also analyze decomposition-based features, encompassing statistical features like mean, maximum, minimum, and standard deviation. Frequency features, such as EDASymp and EDASympn, along with power spectral density, are also investigated, shedding light on different aspects of EDA related to stress. Moreover, time-frequency features like TVSymp are examined to capture nuanced temporal patterns in stress response.

Rao Veeranki, Ganapathy, and Swaminathan 2021 focused on classifying arousal and valence using time-frequency domain features extracted from phasic EDA signals. Specifically, they extracted features such as mean, variance, skewness, and kurtosis from the time-frequency spectrogram generated by STFT, smoothed pseudo-WVD, and CWD methods. These spectrograms comprehensively represent the signal's frequency content over time, capturing transient and sustained changes in the EDA signal associated with emotional arousal and valence. By analyzing these features extracted from the time-frequency spectrograms, Veeranki et al. aimed to develop a robust classification model capable of accurately distinguishing between different emotional states.

Cecchi et al. 2020 explored the emotional reaction of individuals to physical sensations. This study extracted both time and frequency-based features. Time-based features include average value, variance, and standard deviation, while frequency-based features look at signal strength and energy. By examining both, they aim to understand emotional patterns in response to different stimuli.

Greco, Valenza, Lázaro, et al. 2021 utilized various features to classify stress and non-stress conditions. These features included MedianPhasic (median value over time of the phasic component), MedianTonic (median value over time of the tonic component), AUCPhasic (area under the curve of the phasic component, normalized by the length of the session), AUCTonic (area under the curve of the tonic component, normalized by the length of the session), maxPhasic (maximum peak value of the curve of the phasic component), maxTonic (maximum peak value of the curve of the tonic component), stdPhasic (standard deviation of the phasic component), stdTonic (standard deviation of the tonic component), NS-EDR-freq (number of significant skin-motor nerve activity peaks, normalized by the length of the session), and AmpSum (mean value of the amplitude of significant skin-motor nerve activity peaks). Additionally, they utilized EDASymp (power spectral density of EDA signal in the frequency range of 0.045-0.25Hz). These features provided valuable insights into stress classification, contributing to our understanding of physiological responses in stress-inducing situations.

Time-Frequency Representation of EDA in Emotional Classification

Recent advancements in affective computing have highlighted the significance of time-frequency representation techniques in understanding emotional states. Ganapathy, Veeranki, and Swaminathan 2020 utilized time-frequency methods, such as STFT, to analyze phasic signals and classify emotional states based on valence and arousal. Similarly, Rao Veeranki, Ganapathy, and Swaminathan 2021 employed STFT, smoothed pseudo-WVD, and CWD to convert phasic signals into time-frequency spectrograms for emotional state classification, particularly focusing on arousal and valence categories. These studies underscore the effectiveness of time-frequency analysis in extracting pertinent features for emotional classification, contributing to the evolving landscape of affective computing methodologies. In a recent development, Veeranki, Diaz, et al. 2024 utilized STFT to categorize discrete emotions such as amusing, boring, relaxing, scary, and neutral states, further illustrating the utility of STFT in accurately discerning and categorizing various emotional states.

In literature, a range of image features such as the GLCM, GLRLM, ZM, HM, FDTA, and FOS are extracted to capture detailed information. Öztürk and Akdemir 2018 demonstrated the utility of GLCM and GLRLM for histopathological image analysis, highlighting their robustness in texture classification. Haralick et al. 1973 established GLCM as a fundamental method for texture feature extraction, which has been widely adopted in various imaging domains. Tang 1998 provided insights into applying run-length matrices in texture analysis. Furthermore, advanced feature extraction methods such as Zernike

Moments and Hu Moments have been effectively used in robust video watermarking and vehicle detection, as illustrated by Chen et al. 2023 and Anandhalli et al. 2022, respectively. FDTA and FOS features are also significant in texture analysis and have been compared for their effectiveness in different applications, including the classification of medical images (Tsiaparas et al. 2010).

Exploring Machine Learning Classifiers in Emotion Recognition

In the literature, ML classifiers in ER are pivotal in classifying the individuals' emotional states from physiological signals. These algorithms, including SVM, RF, Neural Networks, KNN, Decision Tree, Naïve Bayes, and CNN, are trained on features extracted from signals like EDA and others. Feature extraction techniques capture essential characteristics of these signals, while feature selection methods identify the most discriminative features. The performance of these classifiers is evaluated using metrics such as accuracy, sensitivity, specificity, F1-score, and area under the curve through cross-validation techniques to ensure robustness and generalizability.

In a study conducted by Katsis, Katertsidis, and Fotiadis 2011, the focus was on discerning an individual's affective state across five distinct classes: relaxed, neutral, startled, apprehensive, and very apprehensive. This classification task relied on physiological signals, including EDA, HR, BVP, and RSP. Eight features were then extracted from these signals. To effectively categorize these affective states, the researchers explored the utility of four classification algorithms: artificial neural networks, SVM, RF, and a Neuro-Fuzzy System. The study's results found that the Neuro-Fuzzy System emerged as the most effective classifier, achieving an overall classification accuracy of 84.3%.

Jang et al. 2015 conducted a comprehensive study that classified emotions, namely boredom, pain, and surprise, utilizing various physiological signals, including ECG, SKT, EDA, and PPG. Through meticulous analysis, they identified 27 distinct physiological features from these signals. The objective was to develop effective classification models to discern between these emotions. The authors employed a combination of statistical methods and ML algorithms to achieve this. Among the statistical techniques utilized was Discriminant Function Analysis (DFA), a robust method commonly employed in pattern recognition and classification tasks. Additionally, they explored the efficacy of five ML models: Linear Discriminant Analysis, Classification and Regression Trees, Self-Organizing Map, Naïve Bayes algorithm, and SVM. Through experimentation, the study revealed that DFA yielded the highest recognition accuracy, achieving an impressive rate of 84.7%. This finding underscores the effectiveness of DFA in discerning the nuanced physiological patterns associated with different emotional states.

Khezri, Firoozabadi, and Sharafat 2015 employed three-channel forehead biosignals, including EEG, EMG, and EOG, alongside peripheral physiological measurements such as EDA, BVP, and interbeat intervals, to investigate and classify the six fundamental emotions: anger, sadness, fear, disgust, happiness, and surprise. To extract meaningful insights from this multifaceted data, a total of 192 features were meticulously computed from the signals. A sequential forward floating selection methodology was adopted in the quest for optimal feature selection. This approach facilitated the identification of the most pertinent features while discarding those deemed non-relevant, thus refining the dataset for subsequent analysis. For classification purposes, the researchers employed SVM and KNN classifiers to discern between the various emotional states. Through experimentation, 84.7% and 80% classification accuracies were achieved using the SVM and k-NN classifiers, respectively.

A studyAnderson, Hsiao, and Metsis n.d. utilized physiological signals (EDA, ECG, EOG, EEG, PPG) to classify emotions based on the valence-arousal model, distinguishing between low and high arousal states. They extracted 98 features, including time and frequency domain features, and applied feature selection, reducing to 21 and 26 features for arousal and stimulus classification, respectively. Four classification algorithms (Medium KNN, complex tree, medium Gaussian SVM, and bagged trees) were employed. Models underwent training and validation using a 5-fold leave-one-subject-out cross-validation scheme, yielding an 88.9% accuracy for arousal classification and 100% accuracy in discriminating between relaxing and exciting videos. Stimulus classification achieved an accuracy of 80.6%, with music classified at 100%, videos at 58.33%, and games at 83.33%. This study highlights the potential of biosignal-based emotion classification for future human-computer interaction research (Anderson, Hsiao, and Metsis n.d.).

Kim et al. 2018 aimed to classify emotions across five distinct tasks: baseline, mental arithmetic, recovery from stress, relaxation, and recovery from relaxation. They utilized cvxEDA to extract phasic EDA signals, deriving four key features. Subsequently, they employed SVM-Recursive Feature Elimination (SVM-RFE) to identify the most informative features for emotion classification. The study utilized SVM, decision tree, and KNN algorithms, supplemented by 5-fold cross-validation to ensure robustness and generalizability. Analysis of results revealed notable outcomes, with the SVM decision tree approach achieving an overall accuracy of 74%, sensitivity of 74%, and specificity of 71%.

Pinto, Fred, and Silva 2019 study analyzed the exploration of emotion classification by utilizing EDA, ECG, RSP, and BVP signals, with each signal type contributing a unique set of physiological and statistical features to the analysis. They extracted five physiological features and two statistical features from EDA signals, six physiological features and four statistical features from ECG signals, one physiological feature and two statistical features from RSP signals, and two physiological features and five statistical features from BVP signals. The study employed nested 4-fold cross-validation, and SVM models were utilized to classify emotions into two distinct classes: arousal and valence. Through results analysis, the study achieved outcomes, with SVM classification yielding accuracies of 69.13% for arousal and 67.75% for valence.

Machot et al. 2018 conducted a study focusing on emotion classification across four classes: HAHA, HVLA, LVLA, and LVHA, based on arousal (neutral or aroused) and valence (attractiveness or averseness of an event). Utilizing EDA signals, the study extracted 12 features in both the time and frequency domains. Classification models were built using SVM and KNN classifiers. Results demonstrated that the ML models accurately recognized various emotions within the four-class dimensional framework, including anger, sadness, fear, disgust, happiness, and surprise.

In their subsequent study, Al Machot et al. 2019 focused on HVHA, HVLA, LVHA, and LVLA. A deep-learning model was employed, using raw EDA signals from the MAHNOB and DEAP datasets, bypassing traditional feature engineering methods. ML classifiers, including CNN, KNN, Naive Bayes, RF, and SVM, were utilized for classification tasks. The CNN model exhibited superior performance, particularly in subject-dependent emotion classification, showcasing notable improvements across various evaluation metrics. Subject-independent classification accuracy rates were recorded at 78% and 82% for the MAHNOB and DEAP datasets, respectively, while subject-dependent rates reached 81% and 85%. This research underscores the significance of generalized models for ER across diverse lab settings, emphasizing the efficacy of deep-learning approaches in overcoming traditional feature engineering constraints for EDA-based ER.

Fiorini et al. 2020 undertook emotion classification utilizing EDA, ECG, and EEG signals with the aim of classifying three distinct emotional states: relaxation, positivity, and negativity. A total of 42 features were extracted from these physiological signals, sampled across three different time-window frames (180s, 150s, and 120s), enabling a thorough exploration of temporal dynamics in emotional processing. A meticulous analysis was conducted to identify the optimal time-window frame. Subsequently, the classification task was approached using a combination of unsupervised and supervised ML techniques. Three unsupervised approaches, including K-Means, K-Medoids, and Self-Organizing Maps, were employed alongside four commonly used supervised techniques: SVM, decision tree, and KNN. In the best performance of unsupervised learning, achieved an accuracy of 77% with K-Means, while in the supervised setting, the best-performing accuracy was 85% with KNN.

Cavallo et al. 2021 employed EDA, ECG, and EEG signals to categorize emotions into four classes based on valence and arousal dimensions: HVHA, HVLA, LVHA, and LVLA. They extracted 42 features from these signals, applying the Kruskal–Wallis test to identify statistically significant features. Three supervised machine learning classifiers were used, including SVM with different kernels, decision trees, and KNNs. To assess classifier performance, a 5-fold cross-validation technique was implemented. The KNNs classifier achieved the highest accuracy of 87.7% in emotion classification.

Veeranki, Ganapathy, et al. 2024 conducted a study focusing on classifying two-class arousal-valence emotions using phasic EDA signals. The study applied the cvxEDA method to decompose EDA signals into tonic and phasic components, subsequently extracting 38 features from the phasic component across time, frequency, and time-frequency domains. Employing a variety of ML algorithms, including SVM, linear discriminate analysis, decision tree, RF, and extreme learning machines, the classifiers were trained on the extracted features to discern arousal and valence states based on EDA signals. Notably, the decision tree classifier emerged as the top performer, achieving the highest F1-score for both arousal and valence scales, with a maximum performance of 79.30% in the arousal state.

In their subsequent research, the team advanced their approach by employing a multiscale strategy for EDA signals. These EDA signals were decomposed into multiple scales using the coarse-grained method, and they processed these multiscale signals with a multiscale CNN for feature extraction and classification. Results showcased a classification accuracy of 69.33% for valence states and 71.43% for arousal states. Remarkably, the classifier's performance was significantly influenced by the number of layers and signal length, with the multiscale CNN approach outperforming single-layer CNNs. Overall, the study underscored the effectiveness of utilizing EDA signals and multiscale convolutional neural networks for ER, achieving promising accuracies for valence and arousal states Ganapathy, Veeranki, and Swaminathan 2020.

Rao Veeranki, Ganapathy, and Swaminathan 2021 examined two-class arousal and Valence from the phasic EDA signals, utilizing three time-frequency domain representations to capture spectrogram information. Spectrogram-based features were extracted and employed for classification using Naive Bayes, LR, SVM, and RF machine learning algorithms. Notably, when applied with the smoothed pseudo-WVD representation, the RF classifier achieved the highest performance, attaining an F1-score of 68.74% and an area under the curve of 71.30%. These findings highlight the importance of thoughtful classifier and time-frequency representation selection for optimizing model performance, especially in emotion categorization based on physiological signals like EDA.

Mirzaeian and Ghaderyan 2023 conducted a study focusing on classifying three cognitive workloads using EDA signals. The EDA signals underwent a transformation into two-dimensional spectrograms utilizing smoothed pseudo-WVD. These spectrograms extracted features based on GLCM, including energy, correlation, homogeneity, and contrast. The classification was performed using various machine learning algorithms such as SVM, KNN, cascade forward neural network, and RNN. The study demonstrated significant discriminatory power, achieving a maximum performance accuracy of 97.71% using the contrast feature to distinguish between the three cognitive workload levels.

2.1 Motivation

The motivation for this thesis is as follows:

- Emotion detection permeates critical sectors such as healthcare, well-being, human-computer interaction, and safety, yielding profound practical implications across diverse fields.

- Utilizing EDA signals to characterize discrete emotional states fosters the creation of sophisticated and personalized wearable devices catering to individual needs and preferences.

- Fine-tuning EDA signal components and segmentation strategies elevates the precision and utility of emotion detection systems, ensuring more accurate and insightful analyses of human emotions.

- Pioneering advancements in feature extraction methodologies are indispensable for capturing the nuanced intricacies of EDA signals, driving forward the precision and reliability of emotion detection technologies.

2.2 Objectives

The main objectives of this study are:

- Identify the most effective method for decomposing EDA signals into phasic and tonic components.
- Explore optimal segmentation strategies for phasic EDA signals to classify emotional states effectively.
- Analyze the effectiveness of various windowing approaches on segmented phasic signals to enhance ER accuracy.
- Identify the optimal 2-dimensional representation technique for effective emotion classification.
- Determine the significant set of features from the 2-dimensional representation of EDA signals to improve emotion classification accuracy.