



Review Article

A Systematic Review on Emotion Recognition System Using Physiological Signals: Data Acquisition and Methodology

Tawsif K. ^{1*}, Nor Azlina Ab. Aziz ^{1*}, J. Emerson Raja ¹, J. Hossen ¹,
Jesmee M. Z. H. ¹

¹ Faculty of Engineering & Technology, Multimedia University, 75450, Melaka, Malaysia.

Abstract

Emotion recognition systems (ERS) have become a popular research field to contribute to human-machine interaction in different areas. Different kinds of applications on ERS can serve different purposes. Artificial intelligence (AI) and the internet of things (IoT) are the technologies behind such applications. The main objective of this study is to enable researchers and developers to search for the most suitable options to develop an emotional state recognition system. More specifically, this paper presents work on ERS, which is built using physiological signals extracted from biosensors. It also presents details of how the extracted physiological signals are used to identify the user's emotional state. In this review, the sensors are categorized based on their modality: contact-based sensors and contactless sensors. Next, the ERS process is presented together with the reported results for each described technique. Articles from four different research databases were reviewed, of which 147 articles from 2009 to 2021 were referred to that are related to ERS using physiological signals. This paper should be significant for researchers developing systems that integrate human emotion recognition capability. The findings reported here can guide them in choosing suitable methods for their systems.

Keywords:

Emotion Recognition System;
Biosensors;
Physiological Sensors;
Physiological Signals.

Article History:

Received:	03	March	2022
Revised:	04	June	2022
Accepted:	19	July	2022
Available online:	17	August	2022

1- Introduction

In engineering and computer science, one of the main objectives of recognizing human emotions is to provide a personalized and interactive human-machine interface. The potential applications of such technology are vast. Among many applications of emotion recognition (ER) are advanced driver-assistance systems [1], health care [2], social security [3], digital multimedia entertainment [4, 5], and other fields. However, recognizing an emotional state is complex and challenging due to the many interrelated, complex, objective, and subjective factors. Emotions are complex biological phenomena, which are referred to by the subject's physiological system's neural structures. Emotion is also affected by external factors such as culture, age, and numerous other social and external matters.

A machine can understand human emotions through various emotion elicitation protocols, feature extraction techniques, and classification methods [5]. An emotion can be recognized by extracting relevant data such as audio/visual (voice, facial expression, etc.), physiological measurements (respiration, skin temperature, etc.), and appropriate human activities (gesture, posture) [6]. In the past few years, many researchers working in ER have concentrated on analyzing data obtained from only one sensor, for example, video (facial expression) / audio (speech) data [7, 8]. These systems are known as single-modal ERSs. On the other hand, many other researchers combine the data from different sensors to develop multimodality ERSs. These systems' central task is to fuse the data obtained via various sensors attached to a human body and detect emotional states [6]. Multimodal ERs are claimed to be more robust [6, 9].

* **CONTACT:** 1191402716@student.mmu.edu.my; azlina.aziz@mmu.edu.my

DOI: <http://dx.doi.org/10.28991/ESJ-2022-06-05-017>

© 2022 by the authors. Licensee ESJ, Italy. This is an open access article under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Among all the modalities, physiological-based is attractive as no matter how hard a person tries to mask his/her emotions, their bodies usually experience a series of involuntary physiological changes, such as deeper respiratory breathing, faster heartbeats, sweating, and other chains of reactions. This is in contrast to facial-based ERS, where the recognition system fails when a person is showing a lack of facial expressions or chooses to mask the reaction. Physiological data can be obtained from bio-sensors. These sensors are widely available on the market with good accuracy. Wearable devices with bio-sensors are also beneficial for obtaining and monitoring different physiological measurements like blood pressure, heart rate, movements, temperature, etc. [1].

The main objective of this work is to investigate the options available for uncontrollable physiological information for the automated ER. Moreover, this work also aims to summarize the state-of-the-art techniques and information for further research [1]. The data obtained from one or many sensors needs to be deciphered to identify the emotion. The emotional state must first be discretized. It is also challenging to extract valuable features to train the learning models. There are several techniques for this purpose. Thus, this aspect is also reviewed in this work. Additionally, deep learning research is gaining momentum in this area. Unlike machine learning, deep learning allows learning without feature extraction and selection while offering outstanding efficiency and accuracy. Selecting the best model with tuned hyperparameters is also very important. Hence, this review also focuses on those learning approaches and their advantages and limitations. The review is conducted using the systematic literature review method. The main contributions of this study can be listed as follows:

- Review on the fundamental of ERS which are the emotional models to indicate the emotional states of a subject.
- The physiological modalities as input data of ERS are presented and discussed.
- The learning models approaches and the required pre-processing steps which are important in building an ERS are reviewed.

Existing ERS systematic reviews focus on a single emotion induction such as music by Joseph and Lekamge (2019) [10] or a single signal retrieval option such as EEG [11, 12] and or ECG [13]. This work provides a more comprehensive review.

The rest of this paper is organized as follows. Section II presents the systematic literature review methodology, followed by the review method. The emotion models are presented in section III; these models are commonly used for emotion classifications. Further, in Section IV, the widely used modality of ERS based on physiological signals is discussed. The common framework used by other researchers for ERS is reviewed in Section V. It describes the pre-processing steps, the machine learning approach (Feature Extraction, Feature Selection, and Emotion Classification), the deep learning (DL) approach, and the performance measurement tools. The reviewed works on ERS using physiological signals are summarized in Section VI, and finally, the conclusion and future work are in Section VII.

2- Methodology: Systematic Literature Review

2-1- Research Questions

The key research focus for this study is how an emotion can be automatically recognized from individuals' physiological signals from data collection to the analysis. Based on this, the research questions are as follow:

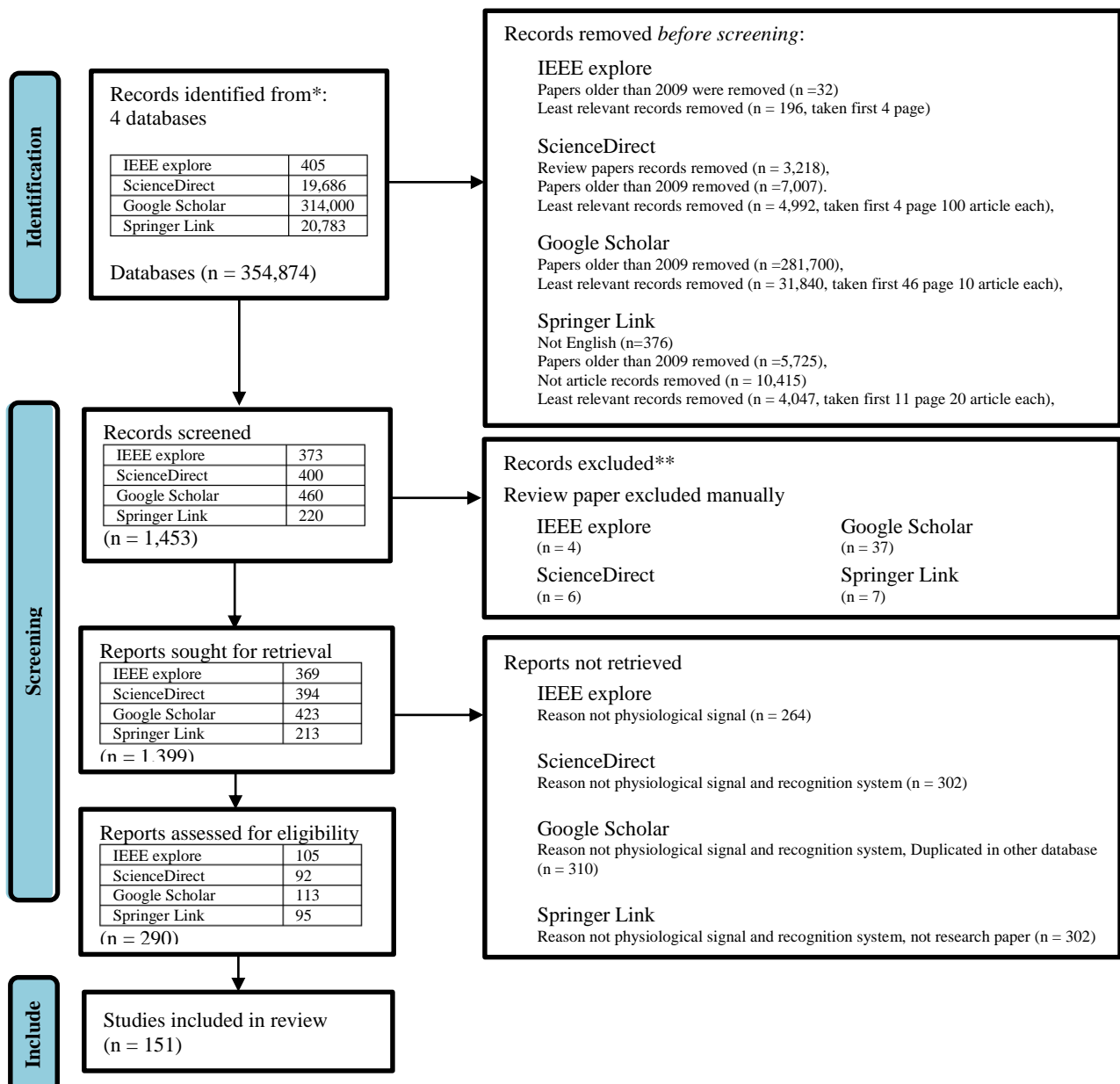
- RQ1:** What is the most common identification of emotional state in the selected research paper?
- RQ2:** What are the type of bio-sensors that can be used to recognize emotions?
- RQ3:** What is the most used feature extraction and selection approach to obtain features from the physiological signal?
- RQ4:** What are the machine learning algorithms frequently used for emotional recognition?
- RQ5:** What are the deep learning algorithms frequently used for emotional recognition?
- RQ6:** What are the most common methods to evaluate the developed and trained ERS model?

2-2- Literature Search

A systematic review is conducted to review the topic presented in this paper using the criteria and databases shown in Table 1. The articles published between 2009 to 2022 are researched with specific search terms. However, most articles are taken from the most recent research from 2016 to 2022. The search terms mentioned in Table 1 are used to search the papers in the four databases; IEEE Xplore, ScienceDirect, Google Scholar, and Springer link. IEEE Xplore provided the most relevant articles related to the search, with 373 articles for screening. However, these papers are not all precisely related, the papers are filtered further after reading the abstract and conclusion, obtaining 105 papers eligible to add in this paper. Google Scholar is also searched for works from 2009 to 2021 with the same search terms, and 460 articles were obtained, but these papers contain duplicate articles from other three journals which was later filtered. Finally, the Springer link database was searched with the selected filtered year, English written, only articles, most relevant; and 220 papers were retrieved. The ScienceDirect database provided 400 papers. However, there are few papers that do not have evaluated results of accuracy; for example, the work only has target arousal and valence value. After filtering the articles, a total of 151 articles are referred. Initially, all related papers' abstracts were reviewed with the search term "emotion recognition". Then the search is narrowed down to only articles containing physiological signal input for ERS. For example, a work on facial-based ERS is excluded due to the modality. A summary of the search is presented in Figure 1, a PRISMA flow diagram for systematic reviews for the four databases search.

Table 1. Literature Search Strategy with Inclusion and Exclusion Criteria for Datasets

Language	English
Year	2009-2021
Search terms	["emotion" OR "emotional"] AND ["recognition" OR "classification" OR "Feature Extraction" OR "Feature Selection" OR "model"] AND ["physiological signals" OR "ECG" OR "EEG" OR "EMG" OR "HRV" OR "PPG" OR "EOG"]
Survey Resources (digital libraries)	IEEE explore ScienceDirect Google Scholar Springer Link
Inclusion criteria	papers that uses machine learning model for emotion recognition papers that use deep learning model for emotion recognition Papers involving physiological data input The accuracy of the proposed ERS is presented clearly Relevant to research and published between 2009 and 2021 Research paper Peer reviewer journals
Exclusion criteria	Exclude papers that use facial data as input for emotion recognition Exclude papers that use speech data as input for emotion recognition Review papers Proceedings

**Figure 1. A PRISMA flow diagram**

2-3- Limitations of the Search

A limited number of the database are available, and even from the available database many of the research papers, only the abstracts are accessible due to subscription issues. Therefore, only the open access and articles from subscribed sources can be properly analysed. This is the main challenge faced in this study.

2-4- Common Limitations of Existing Systematic Reviews

Most systematic reviews were focused on a single emotion induction (such as music [10]) or a single signal retrieval option (such as EEG [11, 12]). Hence, the existing works did not consider comparing between different index tests.

3- Emotion Models

The RQ1 is answered in this section, where a detailed description of the emotional models commonly used among researchers is presented. The emotional states must be distinct and assessed quantitatively for effective emotional recognition. The primary emotions were defined decades ago by psychologists. Researchers widely adopted two models of emotions: a) discrete and b) multi-dimensional emotional state.

In order to build an ERS, emotional data need to be generated where a group of participants is usually provided with a sequence of emotionally evocative materials to induce their emotional state. These materials can be pictures, movies, music, situational simulation, computer games, or recollection. The collected data is later used to train and build ERS model.

3-1- Discrete Emotion Models

The discrete emotion models provide fundamental emotion theories as stated by Ekman (1992) [14], Izard (2007) [15], and Plutchik (2007) [16]. No one emotional state is more essential than another state. Two or more emotional states might be triggered together in some situations. For instance, the emotional state of admiration and joy can be activated as a person falls in love, whereas sadness can be felt once a person is hurt, and hurt might trigger anger. Hence, it is possible to get a mixture of emotional feelings.

In early 1992, Ekman (1992) [14] stated six standard emotion states (i.e., happy, sad, anger, fear, surprise, and disgust). He also mentioned other emotional states as producing responses and mixtures of these standard emotional states. In 1980 [16], a wheel model was proposed as a discrete emotion model. The wheel of emotions has primary emotional states of anger, anticipation, joy, trust, sadness, surprise, fear, and disgust (Figure 2). It also contains the weaker and stronger intensity of emotions, where the center includes a more vigorous intensity, and at the flower blooms are the weaker feelings.

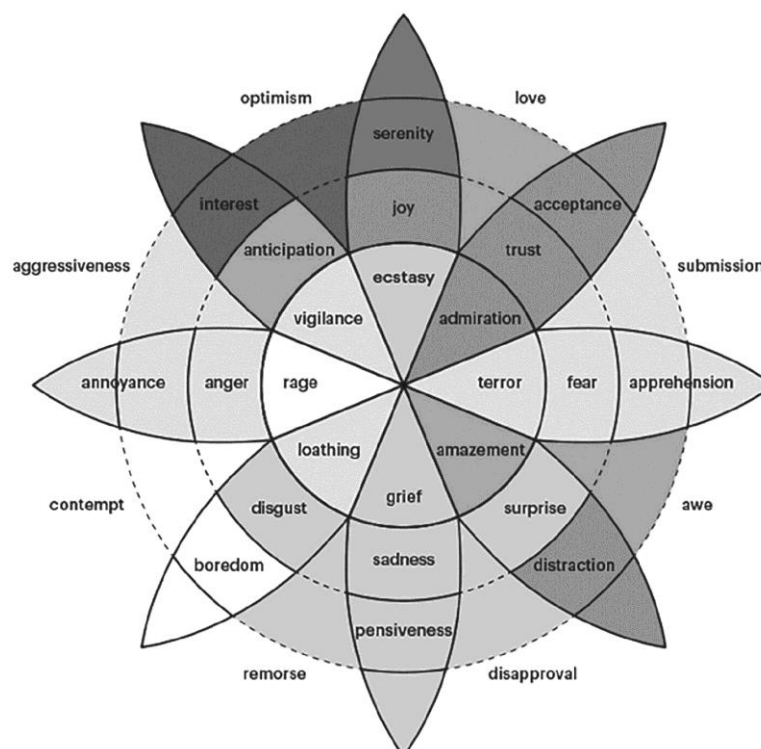


Figure 2. Plutchik's Wheel of Emotions [17]

Izard (2007) [15], also defined a set of basic emotions. According to the author, basic emotions are part of the human evolution course, and every essential emotional state is related to simple brain activity. Moreover, he categorized ten primary emotional states (i.e., interest, joy, surprise, sadness, fear, shyness, guilt, anger, disgust, and contempt).

Plutchik (2001) [16] stated 24 pairs of separate work, i.e., feelings that combined two different emotional states. Finally, the wheel of emotions can be paired into four classes, as follows:

Primary pair: e.g., alarm = fear+ surprise (one petal apart)

Secondary pair: e.g., Envy = Sadness+Anger (two petals apart)

Tertiary pair: e.g., Delight = Joy+ surprise (three petal apart)

Opposite pair: e.g., Conflict = distraction+ interest (opposite petal apart)

Another similar emotion model presented based on Robert Plutchik's model is Hourglass of Emotions (shown in Figure 3) Plutchik and Hourglass model uses a similar method, with the difference in the number of variables. There is more axis along which the emotion is modeled. On top of it, authors explicitly assigned specific labels to certain regions, providing some sort of compatibility with categorical models. These models are more advanced in terms of complexity and the ability to express complex emotions. Despite their complexity, the mentioned models do not take into account the context and personal typology of the reader or author.

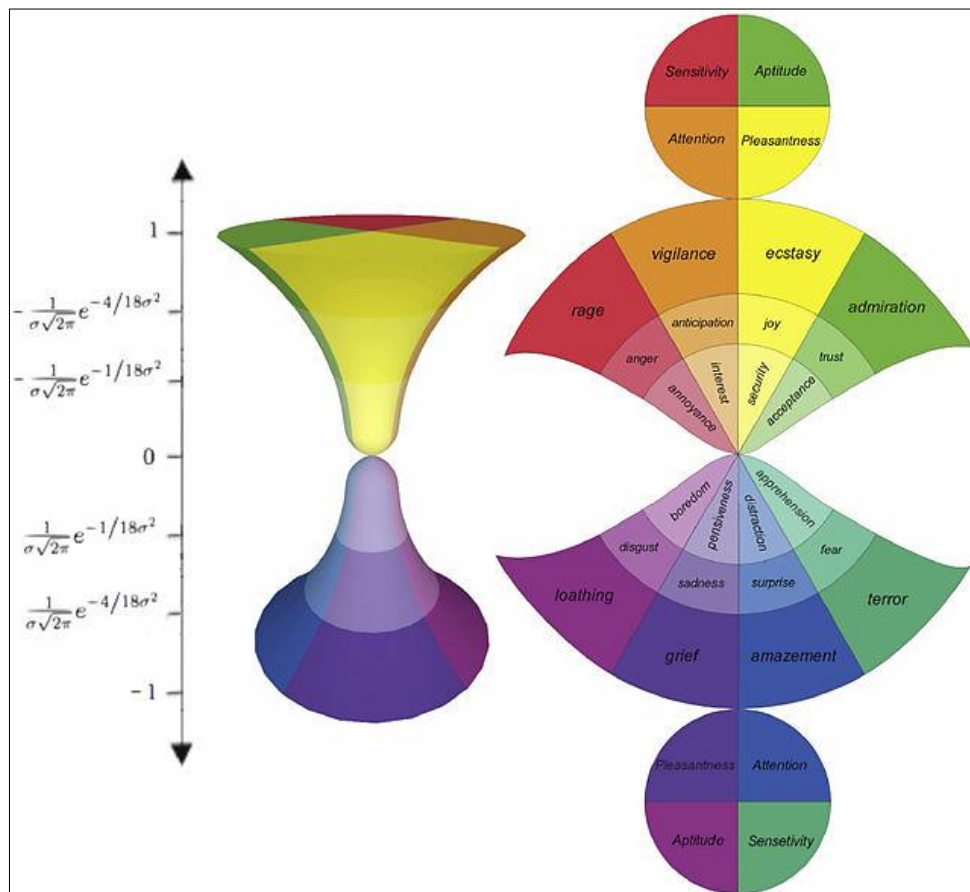


Figure 3. Hourglass of Emotions [18]

3-2- Multidimensional Emotion Space Model

An emotional state may contain different concentrations; for instance, sadness can be at a different level, such as very sad or moderately sad. Therefore, psychologists proposed multi-dimensional emotion space models to differentiate the different levels of emotions.

Cambria (2012) [18] suggested that these emotional levels can be divided into two categories, i.e., 2D dimension, valence, and arousal. In this theory, to specify human feeling, negative valence indicates unpleasant and positive valence indicates pleasant, while low arousal indicates passive and high arousal indicates active. For example, in Russell's Circumflex Model (presented in Figure 4), 'happy' has positive valence and high arousal, whereas 'sad' has negative valence and low arousal.

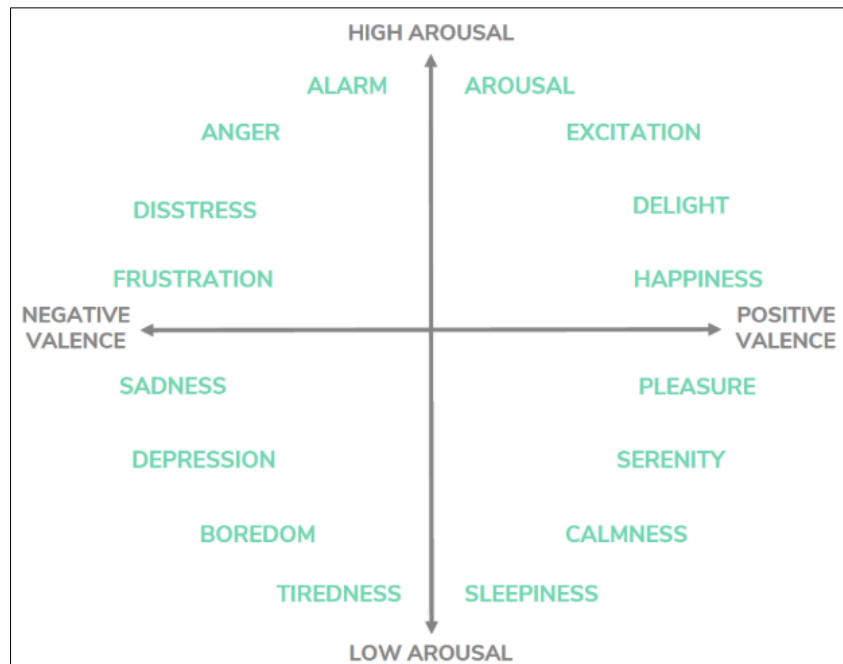


Figure 4. Russell's Circumflex Model of Affect [19-21]

Although the positive and negative emotional states can be easily differentiated using Russell's Circumflex Model, it failed to recognize individual emotional states. To overcome this issue, a 3D model has been proposed (see Figure 5). Dominance is the additional dimension axis; it ranges from submissive to dominant. Dominance helps reflect the specific emotional state to make the outcome more exact and clear. For example, both anger and fear in the distress quadrant for the 2D model can now be easily differentiated due to the dominance axis, where anger is dominant and fear is submissive.

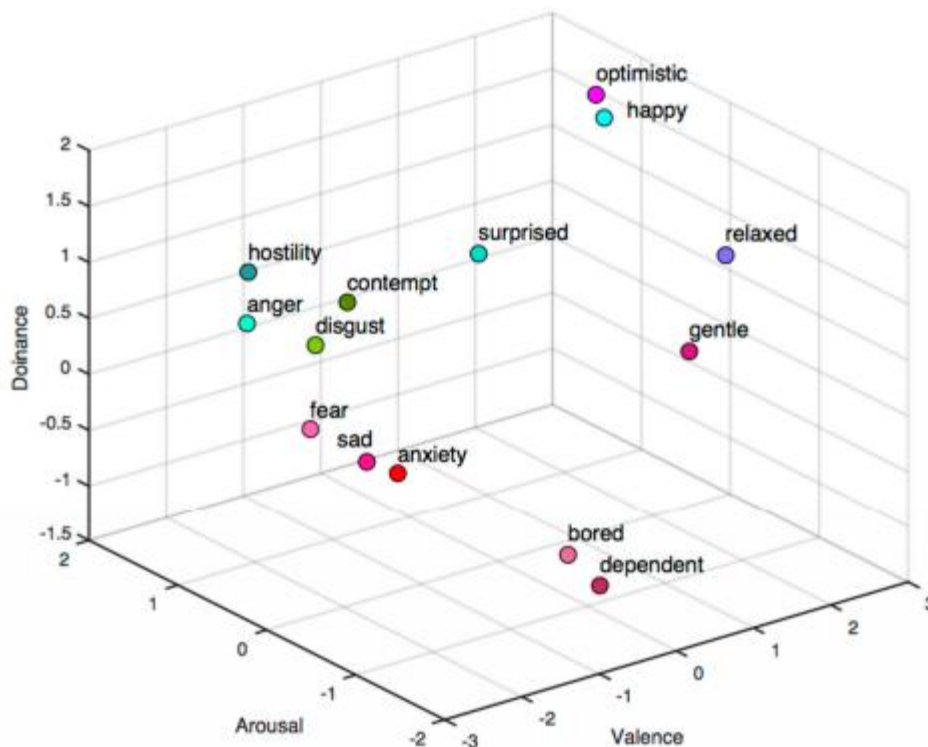


Figure 5. 3D emotion space model [22]

4- Modality of Physiological Based Emotions Recognition System

In this section, RQ2 is answered. Physiological measurements are produced from the central and autonomic nervous systems. These measurements can be obtained from positioning the biosensors, as indicated in Figure 6.

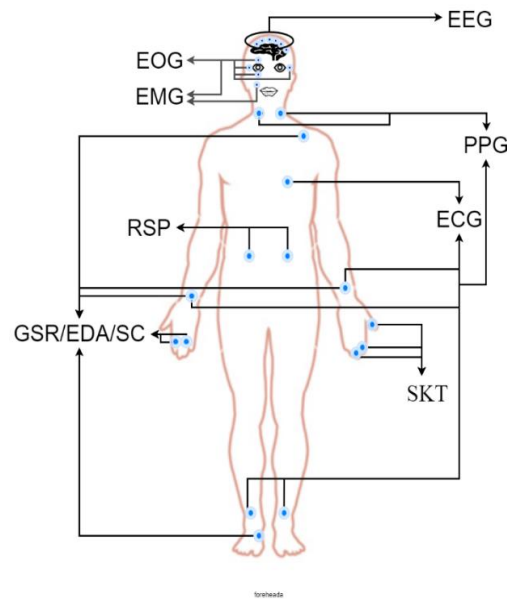


Figure 6. Physiological signals and position of sensors

4-1- Electroencephalogram (EEG)

An EEG signal is from electrical activities in the brain and can be extracted using electrodes attached to the scalp. It is one of the best inputs to detect emotions with high precision. The standard methods use 8, 16, or 32 pairs of electrodes which are attached to four specific places on the scalp (i.e., nasion, inion, and right and left preauricular points) by using the frontal, temporal, parietal, and occipital lobes (as shown in Figure 7-a) [23]. These electrodes are typically attached to the scalp using headsets or electrically conductive adhesive [24]. An example of placing electrodes on the scalp for recording EEG signals is presented in Figure 7-b and proposed by Heng et al. (2020) [25]; the signals were obtained and represented in a graph. However, EEG based system is best for clinical use, as it is very time-consuming to set the electrodes and overcome the sensible noise characteristics.

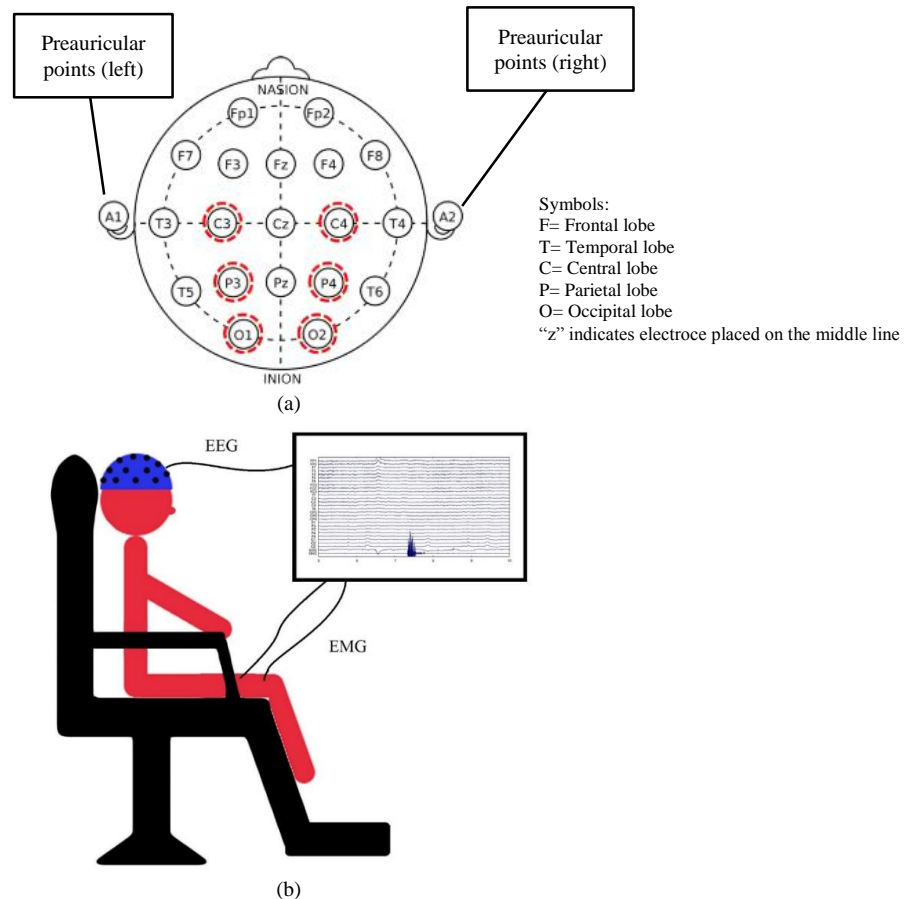


Figure 7. EEG signal reading: (a) EEG electrodes distribution on the scalp [26] and (b) An illustration of EEG recording [25]

4-2- Electrocardiogram (ECG)

Electrocardiogram (ECG) signals are frequently used for the valuation of heart functionality. Like EEG, electrical activities of the heart are measured. The muscular contraction and relaxation are from the heart. It is obtained using electrodes placed on a subject's body [27]. The heart is a vital organ in the human body, and a person's feeling influences its rhythm. Therefore, ECG signals can be beneficial for ERS [28].

There are various methods to obtain ECG; one standard method is the 12-lead ECG technique. Typically, ten electrodes are used to receive 12-lead ECG signals. The ten electrodes are set on limb leads and chest leads, as shown in Figure 8-a. Here, four limb sensors are placed on the flat space of the lower innermost legs area and lower innermost forearm area, or higher innermost arms area and higher innermost thighs area, or higher innermost arms area and innermost abdomen area. Figure 8-b shows the placement of the sensor on the chest area as presented by Tada et al. (2015) [29].

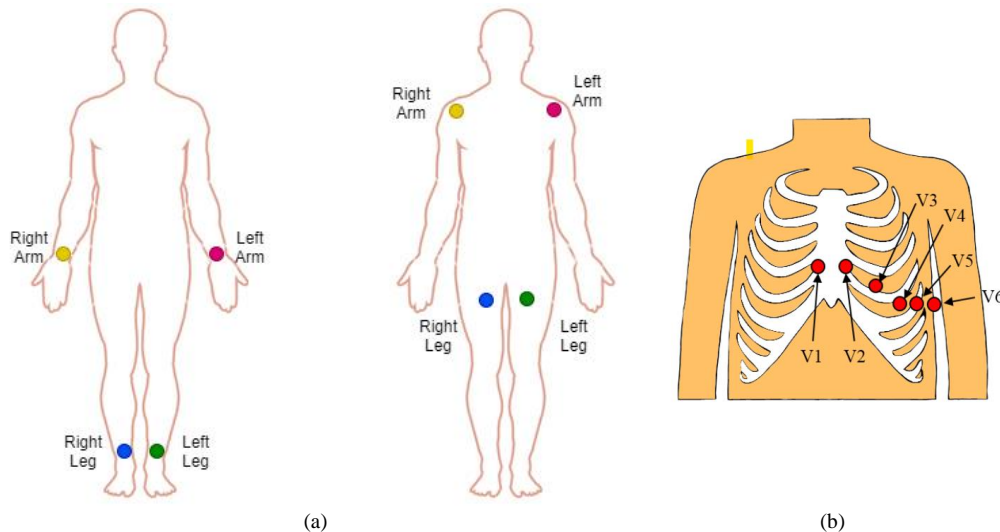


Figure 8. Schematic representation of electrocardiography (ECG) [30]: (a) 4 limb sensors of ECG: Right Arm, Left Arm, Left Leg, Right Leg; (b) Position of a higher resolution, six sensors (V1-V6) [29]

4-3- Electromyogram (EMG)

Electromyography (EMG) is an evaluation and recording method for electrical potential generated by muscle cells. In different areas, these signals can be used for various purposes. In the ERS case, the signals can be used to find perceptive emotions based on physiological reactions. Ekman and Friesen first stated the connection between actions, muscles, and emotions in 1978. The selection of muscles depends on the type of analysis; the most common muscles used are corrugator supercilii, occipitofrontalis, levator labii superioris, orbicularis oculi zygomaticus major [31-33].

Figure 9 shows the electrode placement to obtain emotional expression using EMG. The responses from four selected muscles are received using Ag/AgCl miniature surface electrodes by electrolyte gel. Here, the emotional expression can be measured by the activity of the Zygomaticus major for smiling, which is connected to happiness; the Corrugator supercilii measures wrinkle of one's forehead, which is related to fear, sadness, anger, etc.; the Levator labii Superioris measures upper lip, which is connected to worried, nervous, upset, etc.; and the Lateral frontalis measures eyebrow-raising, which is related to fear, surprise, etc.

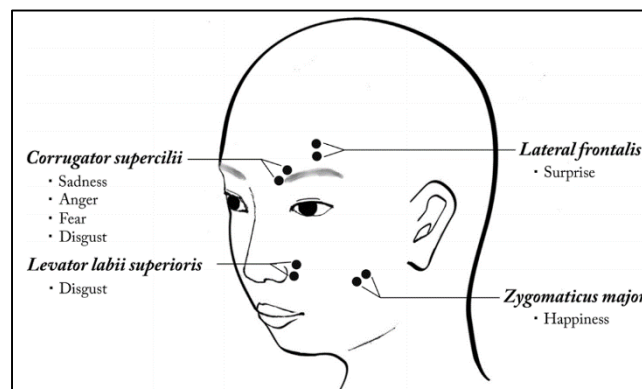


Figure 9. EMG electrode placement for emotional expression using Corrugator supercilii, Levator labii Superioris, Lateral frontalis and Zygomaticus significant measurement [34]

Other possible measurements of EMG for emotional expression are presented in Table 2.

Table 2. Relations between emotions and facial EMG [35, 36]

Emotion	Involved Muscles	Actions
Happiness	Zygomaticus Major and Orbicularis oculi	Closing eyelids, Raising and lowering eyebrows, raising and lowering lip corners
Surprise	Frontalist and Levator palpebrae superioris	Raising eyebrows, raising upper eyelid for eyes open and jaw drops down
Fear	Frontalist and Corrugator supercilii	Raising eyebrows, lowering eyebrows, and raising the upper eyelid
Anger	Corrugator supercilii, Levator palpebrae superioris and Orbicularis oculi	Lowering and raising eyebrows, raising and closing upper eyelid
Sadness	Frontalist, Corrugator supercilii and Depressor angulioris	Raising and lowering eyebrows, raising and lowering lip corners
Disgust	Levator labii superioris and alaeque nasi	Raising upper lip, lips pressed, and wrinkly skin of nasal

4-4- Heart Rate Variability (HRV)

Heart rate variability (HRV) is the Heart Rate (HR) weight by average beats per minute. It records the individual variations in time (or variability) among sequential heartbeats. HRV measurement provides data about heart variability, which can be used to predict emotional states. HRV data can be examined at a particular time in every heartbeat cycle and its regularity [37]. An unusual heartbeat can be due to emotional level changes. The sympathetic and parasympathetic autonomous nervous system in the human body regulates variability in HR. The parasympathetic nerves slow the heartbeat, and the sympathetic nerves accelerate the heartbeat. The variations in emotional levels, such as anxiety and physical pressure, influence HR [38, 39]. A low HRV could indicate relaxation, whereas a high HRV indicates frustration.

Furthermore, HRV measurements are also dependent on age, gender, physical condition, mental stress, eating habits, weight, blood pressure, and glucose level. Inherited genetic factors also affect HRV.

The standard method to measure HRV is by using an ECG [37], which gets a reading of the primary electro biological signal of heart activity. HRV can be calculated using the RR pulse intervals of the ECG signal. PPG is another method to get readings of HRV. PPG can be obtained by placing the probes on the brachial artery, radial or ulnar artery, and tibial artery for both hemibodies [40]. Two sensors are used to obtain PPG, the first sensor (emitter) emits light to the skin, and the second sensor (detector) detects light emitted back from artery pulse wave. The difference in light emitted and reflected is due to the person's blood volume (BV). The BV is produced by capillary dilation and constriction and can be used to estimate HR. Moreover, IBI (inter-beat-interval) from the PPG signal is calculated to obtain the HRV reading. The electrodes/probes placement to obtain HRV is presented in Figure 10-a. Figure 10-b shows ECG illustration and photoplethysmography (PPG) signals. In Figure 11-a, a more detailed operation of PPG sensors is presented. In the image, the reflection is from narrow arteries, where the pressure is lower with greater reflection, causing diastolic value. Whereas if the reflection is from wider arteries, it will absorb more, and this indicates higher pressure, creating systolic value (shown in Figure 10-b).

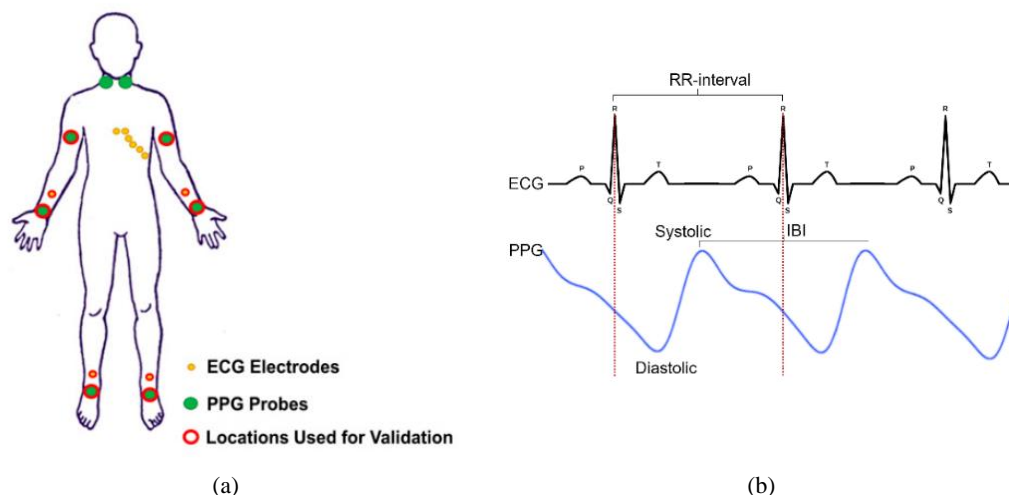


Figure 10. Schematic representation of the acquirable ECG and PPG (a) locations to set Electrodes and Probes reported on a body template [40]. (b) Schematic illustration of ECG and PPG signals over three heart cycles

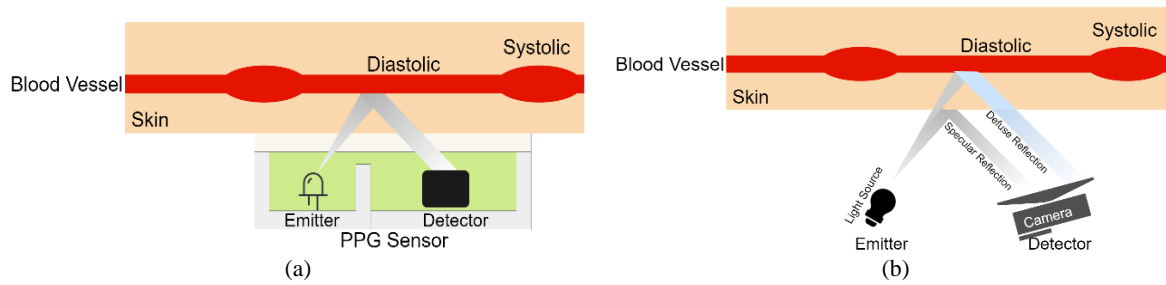


Figure 11. Representative of PPG and rPPG (a) Optical HR sensing. (b) rPPG illustration for obtaining PPG measurement

Remote photoplethysmography (rPPG) makes it possible to measure the cardiovascular pulse wave via a contactless approach by recording different back-scattered light levels remotely using ambient light and vision systems [41]. The rPPG measurements read human cardiac activity using a video camera. The process is similar to PPG, but it is a contactless process. Figure 11-b shows rPPG that gets reading from different light levels (i.e., red, green, blue) reflected from the skin. This processing was used by Benezheth et al. (2018) [39], to detect emotions without contacting the device with a human to obtain the measurements.

4-5- Electrooculography (EOG)

EOG is one of the standard eye movement measurement techniques. It is processed by assessing corneo-retinal polarity. The cornea is set in front of the human eye, which has a positive polarity, while the retina is positioned in the back of the human eye, which has negative polarity. Furman and Wuyts (2012) [23] presented a primary implementation of EOG for ophthalmological diagnosis. For measuring eye movement, electrode pairs are typically set either on the left or right side of the eye (Horizontal) or above and below the eye (vertical), as shown in Figure 12-a. When the eyeball moves from the centre towards any one of the electrodes set between the eye's position, a voltage spike is produced (Figure 12-b) [42]. The basic concept of using EOG for ERS is to detect eye-blinking, which is beneficial for detecting emotional states, such as surprise/ stress [43]. Besides emotions, EOG is also helpful in detecting fatigue, concentration, and drowsiness [44].

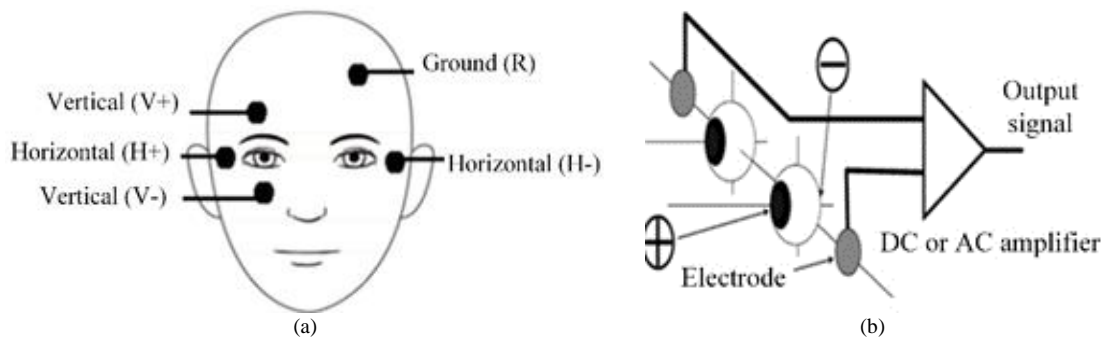


Figure 12. Schematic representation of EOG: [45] (a) Location to set electrodes to obtain EOG; (b) measurement principle [46]

EOG extraction can be via contact or contactless methods. Contact measurements can be implemented by similar instruments used for EMG signal retrieval. On the other hand, the contactless technique can be accomplished using video oculography (VOG) camera systems or using infrared oculography (IROG) cameras [42].

4-6- Galvanic Skin Response (GSR)/ Electrodermal Activity (EDA) / Skin Conductance (SC)

The GSR/EDA/SC is an uninterrupted raw data of electrical factors from the human skin. Here, skin conduction is taken as the main factor, where sweat response produces a different amount of salt on the skin and consequently the electrical resistance alteration from the skin to skin [45]. Sweat causes moisture on the skin's surface and brings fluctuations towards the stability of positive and negative ions in electrodes [47]. Sweat is produced due to the activation of the sweat gland. It is the unconscious reaction of a human [9] and a reflection of changes in the sensitive nervous system [48]. Some of the emotional responses cause sweat reactions, mostly on hand palms, fingers, and feet soles.

SC can be obtained by placing electrodes in different positions on the human body. The commonplace of attaching sensors is on the palmar surface of the hands/fingers or the foot's plantar surface. van Dooren et al. (2012). [49] had tested the emotional sweating across the body on 13 parts and found that attaching sensors on fingers, feet, forehead, shoulders, neck, and chest is the best position to get high SC responsiveness from emotional state changes. These positions to obtain SC measurements for recognizing emotional states are indicated in Figure 13.

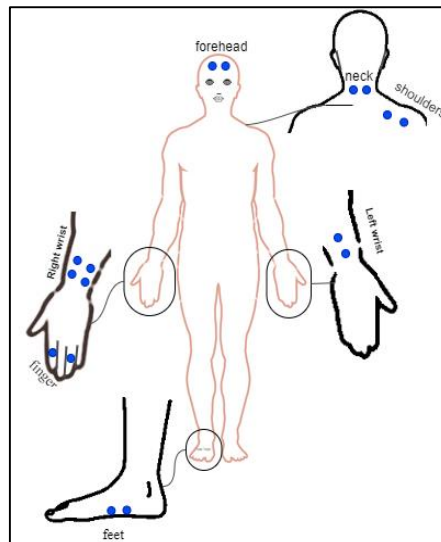


Figure 13. Skin conductance measurement locations: 1) fingers, 2) wrist, 3) chest, 4) foot (instep), 5) forehead, 6) neck, and 7) shoulders

4-7- Skin Temperature (SKT)

SKT is one of the best parameters for automatic ERS. SKT is the unconscious reaction of a human related to SC and also HR. SKT measurement is determined by the thermal radiation of the skin's surface, and it is a valuable indicator of emotional states, which is reflected by the Autonomic Nervous System (ANS) activity.

Similarly, SC, SKT is obtained from palmar surfaces and plantar surfaces (such as nose-tip, and fingertip), as it is influenced by endothelial cells. Basically, broken blood vessels present on the skin's surface will increase the temperature while relaxing and decrease the temperature if in a state of stress or anxiety [50].

The contactless measurement can be processed using electromagnetic radiation released on the skin's surface. It allows the measurement of SKT from a distance. However, there are challenges related to the sensor's accuracy and its coverage. A contact-less system was developed by Kosonogov et al. (2017) [50] to obtain SKT measurements. The measurement is obtained by using an infrared thermal imaging camera FLIR; an example output is presented in Figure 14. SKT data can also be obtained by using contact-based SKT sensors in the finger. Ayata et al. (2017) [51] used a Fingertip temperature (FTT) sensor for emotion state monitoring and recognition. When the person is relaxed, the fingertip is warm than in a tense state of emotion. In a relaxed mood, vessels are dilated, and the fingertip becomes warmer.

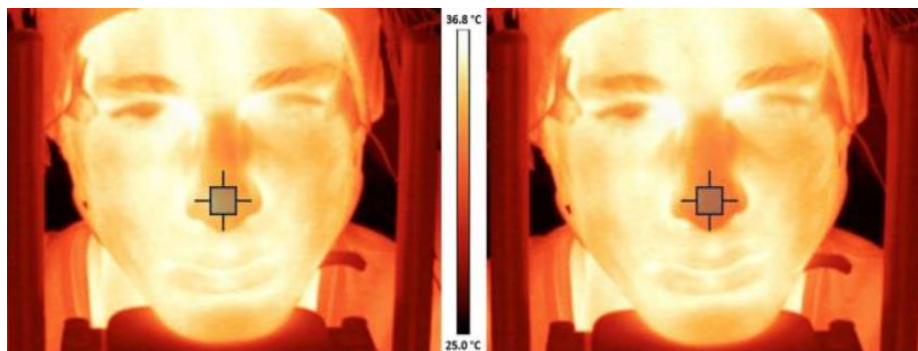


Figure 14. Example of obtaining skin temperature measurement using camera [50]

4-8- Respiration (RSP)/ Respiratory Belt (RB)

RSP biosensor or RB is commonly a stretchable band made of latex rubber; it is used for recording human breathing activity. The RSP is usually worn over the abdomen. The elastic band stretch data is measured as a change in voltage level. The standard recording is RSP rate and breathing depth. In Figure 15, a pressure sensor (EMFit) is shown, where the sensor is attached to the skin using a belt. Here, the human's ribcage volume changes due to respiration compressing the attached sensor [52].

RSP measurement is closely related to other cardiac measures; a deep breath can affect RSP, EMG, and SC measurements. The rate of RSP is typically reduced while relaxing, whereas when the person is stressed out, it results in momentary RSP interruption. In the case of the emotion level, negative emotions contribute to abnormal RSP patterns.

However, irregular RSP cycles can be caused while talking too [53]; hence it is essential to keep this in context while classifying RSP signals. Another method to obtain RSP is by measuring the carbon dioxide (CO₂) contents of inhaling and exhaling air. It is known as capnography or measuring chest cavity expansion [53]. Moreover, it is also possible to obtain RSP by using EMG data acquired from the respiratory muscles [54].

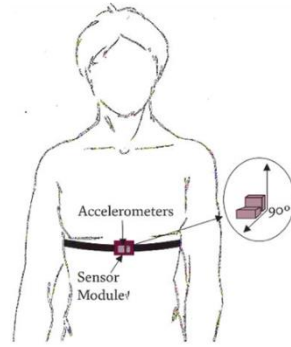


Figure 15. Wearable pressure sensors [52]

4-9- Discussion

There are many different ways to obtain data from various sensors and methods. It is essential to select the best approach. The selection also depends on the requirement of the user and the application area of the ERS. Here, the basic categorization clarifies the standard physiological sensors used for ERS.

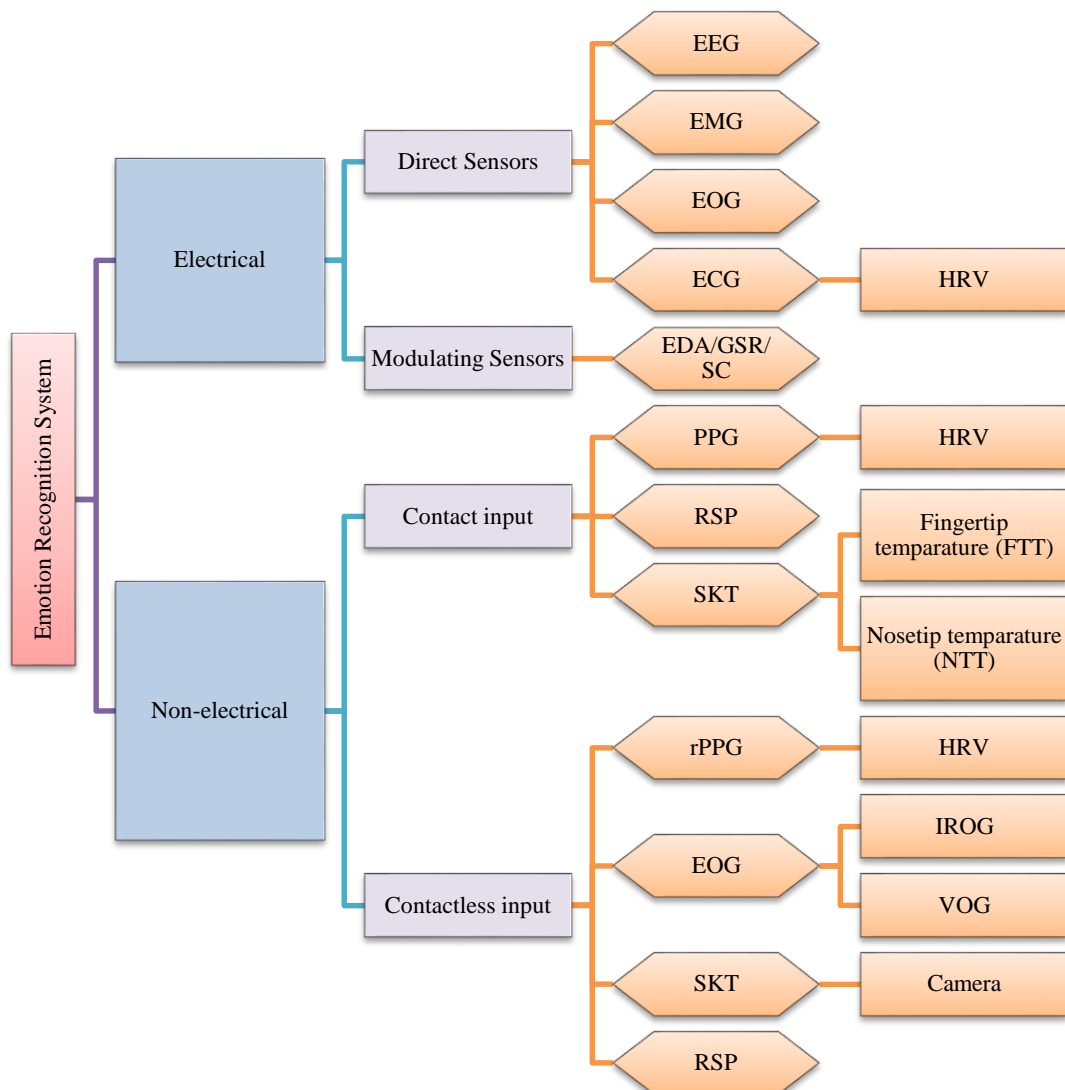


Figure 16. Selection of Unconscious Measurement Methods and Input Sensor for emotions recognition

Figure 16 groups the ERS measurement techniques, which help the selection of the appropriate measurement sensors for ERS. This selection is based on unconscious data collection. As in the conscious process, emotion identification is made using self-evaluation questionnaires, which is less reliable as the participant may not identify the real emotions or provide inaccurate answers due to the inability to understand the questions. However, the unconscious process has many varieties for measurement parameters. These physiological signals also are divided into electrical and non-electrical categories. It can be stated that these electrical parameters are the main signals, which provide a maximum precise outcome, whereas the non-electrical measurements provide human body responses strained by electric signals. The electrical parameters measurements consist of two structures a) direct (self-generating) sensors (such as EEG, ECG, HRV, EMG, EOG) and b) modulating sensors (such as GSR). Moreover, the non-electrical signal can be obtained using contact and contactless approaches like the thermal camera, rPPG, IROG, and VOG.

An engineer that builds ERS for advanced driver-assistance systems might be interested in a non-invasive approach. Hence contactless sensors like the rPPG and video oculography systems are better suited for the project. Whereas accuracy is essential in building an ERS-based healthcare system, thus in this case, direct electrical sensors like EEG or ECG are the better option.

5- Emotion Recognition System (ERS)

In this section, methods for physiological signal-based ERS (RQ3-5) are discussed. The system can be trained using two different ways. The most common way is to use traditional machine learning (ML) models, and another option is DL training approaches. The conventional method required feature extractions and feature selection techniques, whereas DL is free from these methods [55]. DL can train the model without feature extraction due to the capability of inherent data information and extracting features by design. The overall ERS process is presented in Figure 17.

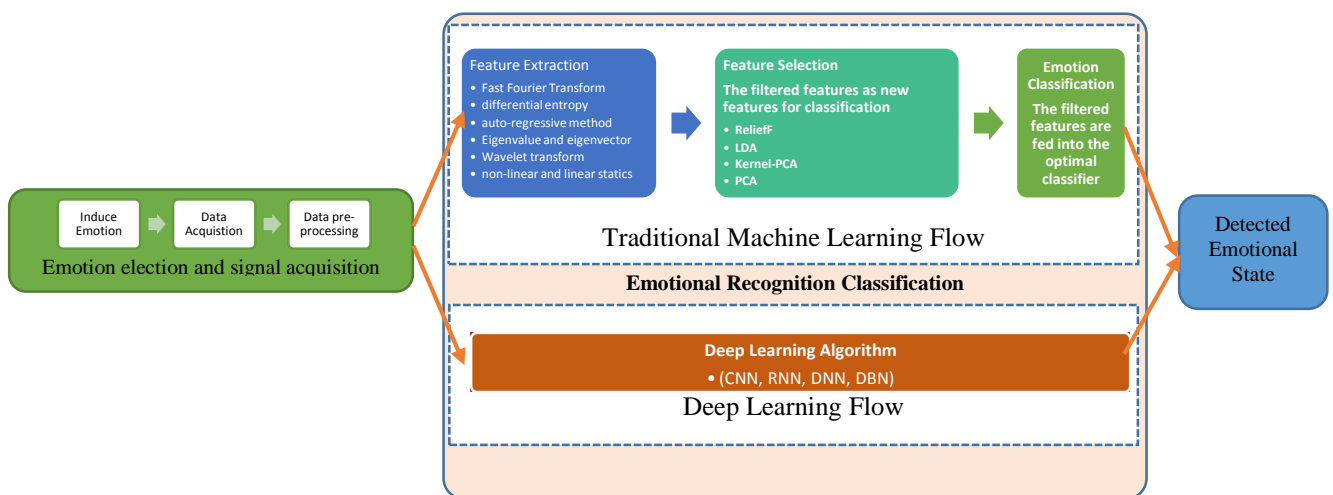


Figure 17. Emotion recognition process using physiological signals under target emotion stimulation

5-1- Pre-Processing

Preprocessing of physiological signals is essential for both traditional and DL approaches, where noise due to crosstalk, electromagnetic interference, etc., was eliminated.

Baseline wander one of the typical noises presented in ECG and PPG recording, which might be caused usually by motion, respiration, and skin-electrode impedance, and so on. For instance, to solve this issue of noisy data, the moving average preprocessing technique can be implemented to make the signal smoother. Another standard method is wavelet transform (WT), which was also used to eliminate Baseline wander and detection of wave characteristics [56]. The most common form of filtering is the "low-pass filter" technique [57, 58], and the "High pass filters" technique [59]. There are several other filter methods, such as independent component analysis (ICA), empirical mode decomposition (EMD), and discrete wavelet transform (DWT). The EMD is WT; it is used for filtering and noise cancellation, an entirely signal-dependent and adaptive approach suitable for real-time applications [60]. Lahmiri and Boukadoum (2015) [61] stated that DWT is more effective than Fourier-based filtering in the case of filtering noise due to multi-scale approximation (MSA). ICA, on the other hand, can extract and decompose multichannel signals from different sources into independent components with mixed signals in linear combination [62].

In general, in the signal collection, the abnormal data due to noise cannot be avoided. It requires some expertise to remove the artifact components through visual observation. Standard original signals contain interference signals.

Different approaches (filtering, DWT, ICA, EMD) are needed on distinctive bio-signals and various sensors of intrusions to remove the noise related to frequency domain and time domain characteristics.

5-2- Traditional Machine Learning Method Flow

There are three main steps in the ML, i.e., feature extraction, feature selection, and emotion classification. In this section, RQ3 and RQ4 are answered.

5-2-1- Feature Extraction

Feature extraction is a significant step in the ER system. Several feature extraction methods have been popular among researchers of ERS. For example, in Hindarto and Sumarno (2016) [63], the Fast Fourier Transform (FFT) was used to extract features from EEG signals; FFT lost the little amount of signal data during transformation. However, without having enough samples, it could not extract enough frequencies.

Time-frequency distributions (TFD) method works on obtaining features from time-frequency with stationary principle. The features contain both the time and frequency domains data from the provided signal. However, the main issue of TFD is it requires noiseless signals to perform effectively, and the data needs to be pre-processed correctly to remove the noise from raw data. Another issue is the performance is slow because of the gradient ascent computation [64].

Meanwhile, Eigenvalue and Eigenvector methods [65] are another option for feature extraction methods. Eigenvector only obtains frequency data from the sinusoid signal. Unlike TFD, it can process signals buried with noise by possible Eigen decomposition for correlation. The disadvantage of this technique is Eigenvector might produce false zeros for the Pisarenko approach of the lowest eigenvalue.

WT and phase space reconstruction has also been used for feature extraction. WT is known to work better for sudden and transient changes in the signal. It also works well in irregular data patterns. It can be used to extract both time and frequency features from the data with linear features. However, it is challenging to select a proper mother wavelet for extracting valuable features. The WT method can be categorized into continuous wavelet transform (CWT), and discrete wavelet transform (DWT). In CWT, the disadvantage is the continuous change in scaling and translation parameters. Another WT method approach is the dual-tree complex wavelet transform (DT-CWT) [66]. DT-CWT uses both real and imaginary tree wavelet filters to obtain complex shifted and dilated mother wavelets. Unlike DWT, DT-CWT has the projecting characteristics of estimated changing value of invariance and higher anti-aliasing [67].

Auto-regressive method (AR) [64] is used to obtain a frequency feature from the sharp spectral signal. AR works to reduce the loss of signal spectral issues and enhance frequency determination. It performs well even if short data segments are provided, as it does not depend on the size of data for the infinite AR spectrum. However, it is challenging to estimate model order in the spectral. If the model order is selected wrongly, the AR model will provide a deprived spectral value.

There are other feature extraction methods available. EEG feature extraction using differential entropy (DE) is discussed in Chen et al. (2019) [11]. Empirical mode decomposition (EMD) can be used to obtain different time and frequency data and obtain a series of intrinsic mode functions (IMFs) data [68]. Researchers used it for extracting physiological signals for emotion state recognition, such as ECG [69] and EEG [68]. However, it also suffers from unavoidable limitations in some domains. The issues stated by Liu et al. (2018) [70] are the frequency resolution stopping criterion, complementary ensemble empirical mode decomposition, and sampling frequency influence. There is no specific proven statement for stopping criterion threshold setting for the frequency resolution [70].

Deep learning is also applied for extracting valuable features. For example, the deep Belief Network (DBN) can be used to extract features for the physiological signal as presented in Hassan et al. (2019) [71] and Asghar et al. (2019) [72]. The main benefit of extracting features using DBN is that it is an unsupervised process and can handle many unlabeled data [73]. Moreover, DBNs can calculate the necessary variables' output weight by integrating the inference procedure approximation. However, few limitations due to the inference procedure of DBNs are only restricted to bottom-up pass. It contains a greedy layer that only learns a single layer's features and does not adjust again with the rest of the layers [74].

The features extracted depend and vary based on the type of the physiological signals and also the features extraction method used. Some of the features are presented in Table 3 according to the physiological signal.

Table 3. Type of features based on physiological signals

Types of Signal	Features
w, W	Energy density
N, D	Demagnetizing factor
ECG	Mean Amplitude Rate, Mean, Root Mean Square, Max, Average and Standard Deviation of HRV [75, 76] Mean of Absolute values of first differences, Mean Frequency, Median Frequency [77]
EDA	Mean Amplitude Rate, Conductance Responses[78], Rate of Skin Conductance [51], Mean of Absolute values of first differences [79], Mean rise duration of skin [78]
RSP	Respiratory Rate, Average level of Breathing, the Longest time between breaths, the Shortest time between breaths, RSPmax, RSPmin [76]
EMG	Mean Value, Root Mean Value, Standard Deviation
EEG	GFP, Zscore, log(Zscore), Interest Index (II), Memorization Index (MI), Power Spectral Density (PSD), Brainrate [76], mean absolute value (MAV) [66]; Pleasantness Index (PI), Fpz Alpha Band, Fpz Beta Band, F3/F4 Alpha Band, Fpz Beta Frequency. F3/F4 Beta Power or Power Ratio, Fpz Alpha and Beta band Power, F3/F4 Alpha and Beta Band Power [66, 80]
EDA/GSR	Decomposition algorithm as a sum of tonic component and phasic component [81], statistical properties such as Variance of amplitude, Standard deviation, Skewness, Kurtosis etc. [82]
SKT	SKTavg, SKTslope, and SKTstd [83]

5-2-2- Feature Selection & Reduction

Once features are extracted from the raw data, it is essential to find the quality and informative features that might be correlated with each other and remove the features which might be unrelated. Useless features may cause the following issues:

- It can cause data analytics overfitting issues and weak outcomes, which leads to low prediction accuracy.
- It might take a long time to process data with useless features.

Hence, feature optimization is a necessity. The popular techniques for feature selections are reliefF, linear discriminant analysis (LDA), principal component analysis (PCA), and kernel-PCA.

ReliefF algorithm is a filter method that ponders each feature according to its relevance to each class. For example, Gómez-Lara et al. (2019) [84] used a relief algorithm to filter each feature according to its relevance to each class from EEG signals. Another researcher proved that the feature selection of EEG signal data with reliefF algorithm could improve classification accuracy [85]. The reliefF algorithm can also select ECG signals [86].

The LDA and PCA are traditional linear techniques and technically are features reduction methods. Zhang et al. (2018) [56] recommended feature reduction rather than feature selection method. Goshvarpour et al. (2017) [87] applied PCA and obtained great accuracy of 100% for ECG data input for all levels of emotional states and dictionaries. The Kernel PCA feature selection uses a global nonlinear approach. However, with feature reduction techniques, a few important data might be lost in the reduction process.

Stepwise regression (SW) consists of regressing multiple variables by removing the most minor contributing predictors step by step. Only independent variables with non-zero coefficients are included in the final regression model. The Akaike information criterion (AIC) is used in SW as the stopping criterion. However, it also contains a few significant issues, as presented by Smith (2018) [88]. One of the issues is a local optimization calculated by involving parameters one-by-one is not guaranteed to be a global optimum. Another problem is it follows automatic rules that take in statistical correlations value without concerning whether or not it is sensible.

Genetic algorithm (GA) has also been applied to the features selection method. In GA, the selection is based on the natural biological evaluation. In nature, organisms have evolved over generations to better adapt to their environment. GAs can be used to maximize the performance of a predictive model on an unseen data set. GAs need a population of individuals and several generations to produce better approximations depending on some mutation and crossover probability parameters. At each generation, according to a fitness criterion, a new set of individuals, i.e., subsets of predictors, is created and also recombined using operators from natural genetics. As GA is processed using global optimization calculations, it is susceptible to being over-fitted, mainly when it is integrated with distance-based classifiers [89].

Random forest (RF) is a known Machine Learning algorithm that contains a feature selection function using the Gini value. Resampling methods such as cross-validation and bootstrap are helpful for feature selection during model building. These methods can maximize the model's performance but increase the computational cost. RF – recursive feature elimination (RF-RFE) provides a reliable assessment of predictors and presents a ranked set of the best predictors at the end.

5-2-3- Machine Learning Algorithms for Emotion Classification

After identifying the useful features through feature selection, the ML model must be trained using the selected input data, and the outcome is the class of each data; this is known as classifications. There are multiple classification techniques that are popular in ERS, namely; probabilistic neural network (PNN) [87], linear discriminant analysis (LDA) [28], recurrent neural network (RNN) [66], quadratic discriminant analysis (QDA) [66], K-nearest neighbor (KNN) [66], random forest (RF) [90] and support vector machine (SVM) [91].

Artificial Neural Network (ANN)

ANN based classifiers are commonly used in this field. A multilayer perceptron (MLP) is a class of feedforward artificial neural networks (FANN). The term MLP is used ambiguously, sometimes loosely to any feedforward ANN, sometimes strictly to refer to networks composed of multiple layers of perceptron (with threshold activation). For the introductory presentation of MLP, a diagram is presented in Figure 18 with one hidden layer, which is connected with every node of the input and output layer. The basic formula of the ANN hidden layer calculation is using Equation 1, and the output layer is calculated using Equation 2.

$$h_j = f(\sum_i w_{ji} x_i), \quad (1)$$

$$y_k = f(\sum_j w_{kj} h_j), \quad (2)$$

Here, the w_{ji} and w_{kj} are weight vectors. Input data vector is 'x', output data vector is 'y', and hidden data vector is 'h'. The input word vector position is 'i', the position of the hidden layer node is 'j', and the position of the output word vector is 'k'.

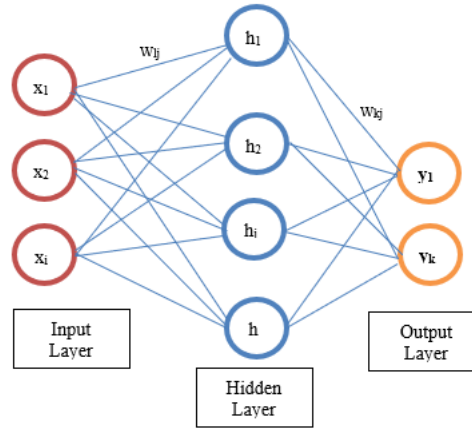


Figure 18. MLP

MLP can tolerate and be correctly trained even with missing data. It has fault tolerance capability; thus, a single node corruption does not prevent correct output calculation. However, it has a few limitations, such as relying on lag observations. Moreover, it is essential to map correctly between input and output to get good performance with static mapping function and fix output and input. Another variant of the ANN classifier is the recurrent neural network (RNN). In RNN, the current hidden state and current input data are calculated by the previous hidden state. For a clearer picture of RNN, an unfolded view is presented in Figure 19. The basic formula of RNN is presented in Equation 3.

$$h^t = f(h^{t-1}, x^t; \theta), \quad (3)$$

where, h^t presents hidden state at t. Where t is time-step. x^t is current input data. The final parameter for f is theta (which encapsulates the weights and biases for the network), for example, W_y and W_x .

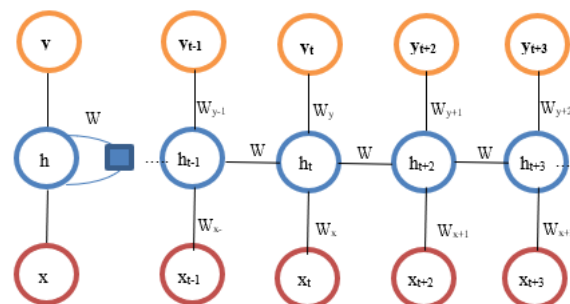


Figure 19. RNN Unfold

Here, W_y and W_x are weight vectors for output and hidden layer respectively. Where W value is weight vector for different time-step t . Input word vector is 'x' and output word vector is 'y'. Moreover, to calculate the activation function, Equation 4 is used, and Equation 5 is used to calculate the output:

$$h_t = \tanh(W h_{t-1} + W_x x_t), \quad (4)$$

$$y_t = W_y h_t, \quad (5)$$

RNN is capable of taking any dimension of input data. The trained model can remember every information throughout time, which makes it a valuable model for time-series data. It can share weights across the time-steps. Even if the training dataset is large, the model size is not large. An improved version of RNN, namely simple recurrent units (SRU), has been implemented by Wei et al. (2020) [66] as EEG-based ERS, and it is proven that SRU is better than MB and SVM. However, RNN has a few drawbacks. Due to the recurrent processing, the complete computation process is slow. It also does not work well in exploding or gradient vanishing data.

Support Vector Machine (SVM)

SVM performs classification by finding the best and optimal hyperplane for separating the data into two classes with maximum margin and no interior points. Hence to obtain the optimal hyperplane, Equation 6 is used.

$$H: l^t(p) + b = 0, \quad (6)$$

Here, b indicates value for Intercept and bias term of the hyperplane equation

The distance of a hyperplane presented in Equation 6, from a given point vector $p = \phi(x_o)$ can be calculated using Equation 7.

$$\text{dist}_{\text{Hyperplane}}(\phi(x_o)) = \frac{|l^t(\phi(x_o)) + b|}{\|l\|_2} \quad (7)$$

Here, for length (l), Euclidean norm is $\|l\|_2$, and is calculated using Equation 8:

$$\|l\|_2 = \sqrt{l_1^2 + l_2^2 + l_3^2 + \dots + l_n^2}, \quad (8)$$

SVM is one of the researchers' choices due to its effectiveness and memory efficiency. SVM is more effective in high-dimensional spaces. In Elsayyad et al. (2017) [86], SVM is used for classifying ECG signals. Another work that also used SVM is by Domínguez-Jiménez et al. (2020) [91]. They compared the classification using mean accuracy and ROC and concluded SVM with the linear kernel (SVML) provided good performance for identifying sadness and amusement. Nevertheless, there are a few downsides of SVM, such as not being able to make multi-class classification, SVM optimization, and correct probability estimates and parameter selection. It was then improved by Chang and Lin (2011) [92], and it was also implemented by a researcher for emotion recognition [93].

Random Forest (RF)

For each decision tree, calculating the importance of a node using Gini Importance as presented in Equation 9, here assumed there is only child nodes (binary tree):

$$\text{Imp}(m) = w_m G_m - w_{\text{left}(m)} G_{\text{left}(m)} - w_{\text{right}(m)} G_{\text{right}(m)}, \quad (9)$$

Here, w_m is weight of samples to reach node m . Where as G_m is calculated impurity for the node m . $\text{left}(m)$ and $\text{right}(m)$ is a child node from left and right split on node m , respectively.

And by involving the predictor variable, all splits are averaged in the forest, which can be used to average splits on variables contained in a group. If considering dataset C splatted into two class, T_1 and T_2 , with the amount of data N_1 and N_2 , respectively, then the Gini index for T is defined in Equation 10.

$$\text{Gini}_{\text{split}}(T) = \frac{N_1}{N} \text{Gini}(T_1) + \frac{N_2}{N} \text{Gini}(T_2), \quad (10)$$

The smallest split Gini (T) is selected to split the node, as it has lower impurities. The Gini index value of one class node will be 0. As mentioned in the previous section, RF calculates the importance of a node using Gini Importance.

Peker et al. (2015) [85] compared RF with feedforward neural network (FFNN), SVM, NB, C4.5 decision tree algorithm, and radial basis function (RBF) network of ANN and found that RF provides the best classification accuracy. The main reason RF gives good performance is the reduced overfitting issue in decision trees due to the learning process. It works for classification and regression cases and even for categorical and continuous data. It also overcomes the data quality issue of missing data. However, the number of many trees processing requires higher computational power and even higher processing time.

Naive Bayes (NB)

Bayes Theorem provides a way of calculating posterior probability $P(x|f)$ from $P(x)$, $P(f)$, and $P(f|x)$ using Equation 11.

$$P(x|f) = \frac{P(f|x) \cdot P(x)}{P(f)}, \quad (11)$$

Here, 'x' is class and 'f' is features for presenting the class. $P(x)$ is the prior probability of the class 'x' occurring independently. The prior probability of features 'f' is $P(f)$. $P(x|f)$ is the posterior probability of the class 'x' occurring in a given 'f'. Whereas $P(f|x)$ is the likelihood, which is the probability of features occurring, given class 'x'.

One of the known classification algorithm is NB. It is used commonly for text classification spam filtering, multiclass prediction, etc. However, the main issue of NB is it provides output considering the input features as independent. It limits the use of the learning algorithm in a few cases. The algorithms set the probability to zero for the unknown test data set. Hence, it is essential to process the smoothing technique before training the model.

Adaboost

Adaboost works as a boosting algorithm for decision trees on binary classification cases. It basically strengthens up a group of weak classifiers to achieve reasonable accuracy. Every weak value is assigned with a weight to boost up the learning. Initially, the weight (w) is set to each instance by using Equation 12.

$$w_{x_i} = \frac{1}{l}, \quad (12)$$

where, x_i is the input data of position i , and the size of data used to train is l .

Next, a weak classifier will be trained using the weighted samples by using the training dataset. Then the misclassification rate is obtained using Equation 13.

$$adaboost_{error} = \frac{adaboost_{correct} - n}{n}, \quad (13)$$

where, $adaboost_{correct}$ is the total value of correctly predicted output by the trained model.

Equation 8 is then modified with the weighting value of the training input and obtained in Equation 14:

$$adaboost_{error} = \frac{\sum(w_i t_{error_i})}{\sum(w)}, \quad (14)$$

Here, w_i is weight calculated and t_{error_i} is prediction error for input value i . The t_{error_i} value is 1 if misclassified and t_{error_i} value is 0 if correctly classified.

Another necessary calculation is to obtain a stage value; this can be done using Equation 15. It provides the weight of any predictions made by the model.

$$adaboost_{stage} = \ln\left(\frac{1 - adaboost_{error}}{adaboost_{error}}\right) \quad (15)$$

Here, \ln is the natural logarithm and $adaboost_{error}$ is misclassification rate calculated by Equation 14.

Adaboost classification is known to be very simple for implementation. However, the algorithm only supports binary classification. Adaboost is also not tolerable to outliers and noisy data.

5-3- Deep Learning Algorithms for Emotion Classification

In this section, the RQ5 is answered. Five deep learning algorithms that are used in ERS are discussed.

5-3-1- Convolutional Neural Network (CNN)

CNN is a class of DL algorithms and a type of FFNNs. Its architecture contains shared weight values with convolution invariance features. CNN is famous in many different domains, and it is also currently gaining popularity for classifying physiological input signals (such as ECG, and EEG) of ERS.

Martinez et al. (2013) [94] used CNN for classifying mental states (i.e., excitement, relaxation, fun, and anxiety) using SC and BV pulse signals. In another research, different statistical features were obtained from the benchmark dataset DEAP and passed to the CNN model for emotional state classification [95]. The combination of CNN with the ability for dynamic learning of a new system was proposed by Song et al. (2020) [96] as dynamical graph convolutional neural networks (DGCNN). The dynamical architecture helps to learn information connections between various EEG channels. The EEG channel data can be presented in an adjacency matrix in rows and columns to generate different features and classify them in the SEED dataset.

Santamaria et al. (2018) [97] implemented deep CNN (DCNN) on a dataset of bio-signals (ECG and GSR) from the AMIGOS database for emotional state detection by associating the ECG and GSR signal data with arousal and valence level. Al Machot et al. (2019) [98] implemented human ERS by proposing a CNN and proved that the system would perform effectively for both subject-dependent and independent human ER using DEAP and MAHNOB datasets. They had explicitly worked with the stress detection of an individual. CNN's main issue is it requires a large dataset to get an effective trained model, and hyper-parameter tuning is non-trivial [99].

5-3-2- Deep Belief Network (DBN)

The deep belief network (DBN) is another DL technique containing more straightforward Restricted Boltzmann machine (RBM) models. DBN can learn in-depth by inputting features through pre-training. From the input data, DBN extracts the deep features gradually.

Hassan et al. (2019) [71] implemented DBN for ERS using EDA, PPG, and zygomaticus EMG (zEMG) bio-sensors data. The proposed ERS classifies the signals into five basic emotional states (i.e., Neutral, Happy, sad, Relaxed, and Disgust). Chao et al. (2018) [100] proposed a combined DL architecture of DBNs with glia chains known as DBN-GCs. The experiment was conducted using the DEAP benchmarking dataset and achieved good performance for arousal and valence states classification. In the work of Huang et al. (2017) [101], before training DBN, they extracted DE features and mapped them, where they found outstanding multi-classification performance.

Deep learning does not require any feature extraction. DBN itself is capable of extracting high-level features from different types of data. Kawde & Verma (2017) [55] ignored the feature extraction step and directly entered the four bio-signal into DBN. These data (i.e., EMG, EEG, GSR, and EOG) were obtained from the DEAP dataset. They achieved greater than 70% accuracy for valence and arousal.

5-3-3- Probabilistic Neural Network (PNN)

Probabilistic neural network (PNN) is an FFNN based algorithm that follows the Bayesian approach. PNN is known as fast training capability due to the ANN structure. It is also proven to give a more accurate classification performance with greater noise tolerance due to the insensitivity to outliers. Nakisa et al. (2018) [102] applied PNN in an EEG-ERS and evaluated its performance using three datasets MAHNOB, DEAP, and a mobile EEG sensor. They had only used PNN and fed the feature vector for the emotional system classification using GSR and ECG signals. However, PNN takes relatively more time for classifying new cases and requires more excellent storage to store the trained model.

5-3-4- Long Short Term Memory (LSTM)

Another deep learning architecture of artificial recurrent neural network (RNN) is long short-term memory (LSTM). LSTM contains feedback connections, which can handle the vanishing issues of RNN. LSTM can process single data points and a sequence of data. Hence, it can process a sequence of data obtained from the physiological signals as implemented by Wöllmer et al. (2013) [103].

Li et al. (2016) [104] presented a framework involving LSTM as classifier and EEG signal as input data and found the classifier's performance was fitting and provided accurate output each time of prediction. Another research extracted rational asymmetry (RASM) features from EEG signals, trained LSTM to explore EEG signal correlations and obtained 76.67% classification accuracy [105]. In the research of Xing et al. (2019) [106], a framework was proposed; it used EEG signal and trained the classification of LSTM-RNN by integrating context relations between the feature sequences and achieved an enhanced performance. Another similar work was done by Alhagry et al. (2017) [107]; the researchers presented an end-to-end design involving LSTM classification by learning features for arousal, valence, and liking.

Basically, LSTM is designed to overcome the fundamental issue of RNN, i.e., the vanishing gradient problem. However, it does not entirely overcome the issue due to the transferring of data from one cell to another for evaluation. To get the architecture's effectiveness, it requires a large number of training data and higher configured hardware. Moreover, LSTM models are susceptible to overfitting issues.

5-3-4-1- SincNet - Customized Deep Learning

SincNet is a customized CNN-based algorithm designed by Ravanelli and Bengio (2018) [108], for a speech recognition system. A bandpass filtration parameterized sinc was used. The classifier is only trained by high and low cutoff frequencies in the raw dataset rather than training using each element from filters in a typical CNN algorithm. SincNet classifier was proven to be efficient for speaker recognition.

SincNet can also handle the classification of EEG measurement data. In Zeng et al. (2019) [109], and improved SincNet-based classification method SincNet-R was proposed. It contains three CNN and three DNN layers. The proposed technique was tested using EEG measurements and proven to perform better than other classification techniques, such as CNN, LSTM, and SVM.

5-4- Performance Evaluation

In this section, RQ6 is answered. The trained classification needs to be evaluated to make sure the classification prediction accuracy is accurate. The most common method to test the classification is by calculating the accuracy rate and error rate. For more details, precision rate, (P) and recall rate (R) are calculated. These rates can be defined using the confusion matrix, as shown in Table 4.

Table 4. The confusion matrix of classification

Symbol	Emotion Predicted	
	Positive	Negative
Positive	True Positive (TP) Predicting positive emotion state correctly	False Negative (FN) Predicting positive emotion state wrongly
Negative	False Positive (FP) Predicting negative emotion state wrongly	True Negative (TN) Predicting negative emotion state correctly

The accuracy rate (A) is the percentage of adequately classified output by the total input, and the calculation is defined by Equation 16. At the same time, the error rate (e) is the classification error and is calculated by the misclassified output to the total input by Equation 17.

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (16)$$

$$e = \frac{FP + FN}{TP + TN + FP + FN} \quad (17)$$

P and Rs are opposite, dependent on each other. They can be calculated using Equations 18 and 19, respectively. If the R output rises, the accuracy output will decrease and vice versa. F1 is the P's harmonic mean, and the R. An F1 score reaches its best value at 1 (perfect precision and recall). The formula for calculating F1 is given in Equation 20.

$$P = \frac{TP}{TP + FP} \quad (18)$$

$$R = \frac{TP}{TP + FN} \quad (19)$$

$$F1 = \frac{2 * P * R}{P + R} \quad (20)$$

Receiver Operating Characteristic (ROC) can be useful for selecting the best classifier. The ROC's horizontal and vertical axis is the false positive rate (FPR) and the true positive rate (TPR), respectively. The FPR and TPR can be calculated using Equations 21 and 22, respectively.

$$\text{false positive rate (FPR)} = \frac{FP}{TP + FP} \quad (21)$$

$$\text{true positive rate (TPR)} = \frac{TP}{TP + FN} \quad (22)$$

An example of ROC is presented in Figure 20, and the two lines represent two different classifiers. If the line is nearer to the upper left corner, then the trained classifier has a good performance. In the example given, the blue line is nearer to the upper left corner than the green line, which indicates the blue classifier is better than the green classifier. Additionally, the blue line contains a bigger area under the ROC (AUC). It can also indicate that the blue line's classifier works better than the classifier represented by the green line.

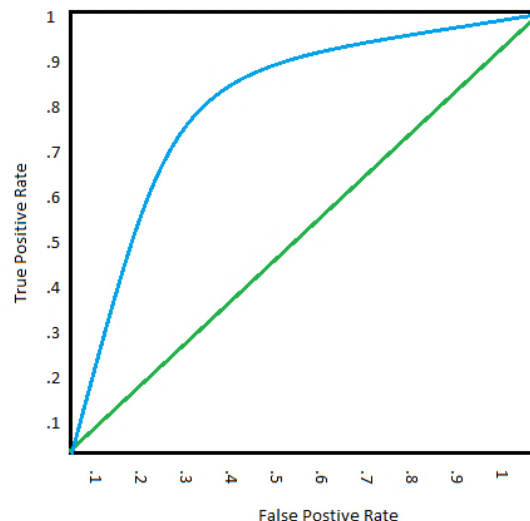


Figure 20. ROC example

6- Previous Work on ERS

This section summarized the reviewed ERS, which are tabulated in Table 5, where "Ref" contains a citation number of the referred work, "Signal" presents the type of input data the researcher used, "No. of participant/dataset" contains information for how many participants were involved in generating the dataset or which benchmark dataset was used, "Emotion stages" presents which classes of emotions were adopted. Next, "Inducement" contains the methods of inducing emotions in the participants. The "Technique and Features" column presents the techniques used for feature extraction and selection. Next, "Classification" contains information on the ML model used by the researchers. Finally, "Accuracy (%)" presents their outcome in percentage.

Table 5. Related Works on emotion recognition using physiological signals

Ref	Signal	No. of participant/dataset	Emotion states	Inducement	Technique and Features	Classification	Accuracy (%)
[31]	ECG, EMG, SC, RSP	MIT database	Joy, Pleasure, Anger, Sadness	Music	Hilbert Huang Transform (mean frequency)	Support vector Machine	76(Fission) 62 (Fusion)
[110]	ECG, BVP, SC, EMG, RSP	-	Valence, Arousal	Images (28 International Affective Picture System)	GA	kNN LDA QDA	78.04 (Arousal) 75.93 (Valence)
[111]	Blood Volume Pulse, EMG, SKT	10 (International Affective Picture System)	Fear, Disgust Contentment, Neutral, Sadness, Amuse	Images	Statistical features (Mean, Standard deviation, etc.)	Support vector Machine Fisher linear discriminant analysis	45 (subject independent) 100 (subject dependent)
[112]	ECG	31	Excitement, Erotica, Disgust, Fear, Gore, Neutral	Passive: IAPS Active: Video games	Instantaneous frequency, the local oscillation	Linear discriminants, leave-one-out cross-validation	52.41 % passive, 78.43 % active
[112]	ECG	44 subjects	Valence, Arousal	International Affective Picture System	Hilbert-Huang transform Empirical Mode Decomposition (EMD)	LDA	89
[113]	ECG	60	Basic Emotions except for Anger	Audio Visual	Rescale Range Statistics, Finite Variance Scaling, Higher-order Statistics (Hurst, Skewness, Kurtosis)	Finite Variance scaling and Higher-order statistics	92.87 (random) 76.45 (subject independent)
[65]	SC, SKT, and BVP	four	Fear, Disgust, Joy, and Neutrality	IAPS image slide-show	Kernel Principal Component Analysis (KPCA), eigenvector	SVM	SC-SKT 60-100
[114]	EMG	12	Joy, Affection and Humor	Pictures	ICA	ANN	>98 (Precision) >97 (recall)
[115]	EMG	Inter National Affective Picture system	Fear, Disgust, Sadness	Images	Wavelet Transform	Support Vector Machine	81.82 (subject Dependent)
[69]	ECG	30	Basic emotions except for anger	Audiovisual	Hilbert Huang Transform and Discrete Fourier Transform	K-Nearest Neighbour Linear discriminant analysis	48.53(DFT) 52.11 DFT
[57]	EMG	15	Sad-Happy, Afraid, Disgust, and Neutral	Audiovisual	Higher-order statistics	K Nearest Neighbour K Nearest Neighbour, Principal component analysis	64.89 69.5
[116]	ECG, EDA, SKT	217	Boredom, Pain, Surprise	Audio	HR, HRV, Ratio of low to high-frequency power	Discriminant function analysis	84.7
						Linear discriminant analysis	74.9
						Support vector machine	62
						Classification and regression tree	67.8
						Naive Bayes	71.9
						Self-organizing map	61.5
[117]	EEG facial expressions	MAHNOB-HCI	Valence, Arousal	Videos	Power spectral density (PSD)	LSTM-RNN	-
[118]	ECG	25	Sad, Angry, Fear, Happy, Relax	Movie clips	Time, frequency, Poincare, statics	Support vector Machine	56.9
[119]	ECG, RSP	11	Joy, Sadness, Anger, and Pleasure	Video	Low-pass filter, Genetic algorithm	SVM	Joy: 81.82, Sadness: 63.64, Anger: 54.55, Pleasure: 30.00
[120]	EEG	DEAP	Valence, Arousal	Videos	Minimum-Redundancy-Maximum-Relevance (mRMR)	SVM	Arousal: 60.7 Valence: 62.33

[121]	EEG	21	Sad, Scared, Happy, Calm	IAPS	Frequency bands (delta, theta, alpha, beta, gamma)	KNN, SVM	55 (KNN) 58 (SVM)
[122]	ECG	Augusberg database	Joy, Anger, Sadness, Pleasure	Music	Db5 Multi-scale wavelet Decomposition (maximum and standard deviation)	Neural Network	91.67 (RBF) 87.5(BP)
[123]	EDA	25	Valence, Arousal	IAPS picture	Leave-One-Subject-Out procedure (LOSO)	K-NN	77.33(Arousal) 84(Valence)
[124]	EEG	DEAP	Valence, Arousal	Videos	-		86.67
[125]	EEG	DEAP	Valence, Arousal	Videos	-	Ensemble deep learning model	84.63(Arousal) 88.54(Valence)
[126]	ECG, EEG, RSP	6	Negative Neutral	Videos	Hilbert Huang Transform	SVM	92.5(subject dependent)
[127]	EEG	DEAP	Valence, Arousal, and Liking levels	Film clips	Liquid State Machines	ANN, SVM, K-NN, DT, LDA	84.63(Arousal) 88.54(Valence) 87.03(Liking)
[128]	EEG, BVP	MAHNOB	Valence, Arousal	Videos	Particle Swarm Optimization, Simulated Annealing, Random Search and Tree-of-Parzen-Estimators (TPE)	LSTM	77.68
[129]	EEG	10	Valence, Arousal	Videos	Radial Basis Function	SVM	60
[130]	GSR, PPG	DEAP/32	Valence, Arousal	Videos	Time domain, Statistics	decision tree, RF, SVM, K-NN	GSR: 71.53 (Arousal) GSR: 71.04 (Valence) PPG: 70.92 (Arousal) PPG: 70.76 (Valence)
[68]	EEG	10	Happy, Calm, Sad, and Fear.	Videos	Empirical Mode Decomposition (EMD) , eigenvectors	Deep Belief Network and SVM	83.34
[131]	EEG	DEAP	Arousal, Valence	Videos	Power Spectral Density (PSD)	Deep Learning	93.6
[56]	ECG, GSR, EMG, PPG	29	Happy, Fear, Sad, Anger	Video	Augsburg Biosignal Toolbox (AuBT): Statistical features	SVM, KNN and Gaussian Naive Bayes	93.42
[72]	EEG	DEAP/32 SEED	Arousal, Valence	Music videos	Time and Frequency domain (AlexNet tool): Deep neural network	SVM/K-NN	DEAP 93.8 (SVM) 91.4 (k-NN) SEED 77.4 (SVM) 73.6 (k-NN)
[132]	ECG	DECAF/10	Happy, Sad	Movie clips	Discrete wavelet transform and Hilbert Haung transform	SVM, k-NN	db7 (DW T) 75 (subjective) HHT 65 (Non-subjective) ST 65 (Non-subjective)
[133]	ECG, EEG	SEED, DEAP	Arousal, Valence	Music video	Wavelet Transform	RF, SVM	83.33 (SEED) 59.06 (DEAP)
[134]	EDA, HR	MAHNOB	Valence, Arousal	Videos	Multi- Dimensional Dynamic Time Warping (MD-DTW)	k-NN	Arousal: 93.6 valence: 94
[135]	EEG	39	Valence, Arousal	Videos	EEGLAB software	Liblinear, REPTree, XGBoost, MultilayerPerceptron, RandomTree, and RBFNetwork	Arousal: 97.79 valence: 96.79
[136]	ECG, GSR, HR, GSR, SKT, etc.	32	Anger, Happy, Sad, Joy	Videos	ANN	neural networks (NN)	75.38
[90]	EEG, GSR, RSP, blood volume pressure (BVP), EMG, EOG	DEAP	Valence, Arousal, Dominance, and Liking	Movie	Bag of Words (BoW) framework	Naive Bayes (NB), the linear SVM, RF	MI-SVM 68 (Valence) EMDD-SVM 69.1 (Arousal)
[71]	EDA, PPG, zEMG	DEAP	Happy, Relaxed, Disgust, Sad and Neutral	Video	Deep Belief Network (DBN)	Gaussian Support Vector Machine	89.53

[11]	EEG	SEED	Positive, Neutral, and Negative	film clips	Differential entropy with Linear Discriminant Analysis (LDA)	k-NN, LR, MLP, RF, and SVM.	68
[71]	EDA, PPG and zEMG	DEAP/ 32	Happy, Relaxed, Disgust, Sad and Neutral	Video	Deep Belief Network (DBN)	SVM	89.53
[137]	EEG	20	Happy, Fear, Sad, and Relax		Variational Mode Decomposition (VMD)	SVM	90.63
[138]	EEG	DEAP	Valence, Arousal,	Video	Variational Mode Decomposition (VMD)	Deep Neural Network	Arousal-61.25 Valence-62.50
[139]	HRV	25	Happy, Neutral, and Sad	Videos	Statistical method, PCA (principal component analysis)	KNN	70.4
[140]	EEG	SEED, DEAP	Arousal, Valence, Dominance and Liking	Video	CNN Fourier transform (FT)	deep model transfer learning	90.59
[141]	EEG	DEAP	Valence, Arousal,	Video	-	CNN	82.84
[93]	EEG	20	Valance, Arousal	videos	Empirical Mode Decomposition(EMD)	LIBSVM	Arousal- 96.13 Valence- 96.79
[142]	RB, PPG, and fingertip temperature (FTT)	DEAP	Valence, Arousal,	Video	Statistical functions, mRMR (minimum redundancy maximum relevance) algorithm	RF,SVM,LR	Arousal: 82.63 Valence: 74.88
[143]	EEG	SEED, DEAP	Valence-Arousal-Dominance	Video	DWT	Echo State Network (ESN): a recurrent neural network with intrinsic plasticity	Arousal: 69.86-73.08 Valence: 69.53-72.18
[144]	EEG	DEAP, DREAMER	Valence, Arousal and Dominance	Video	-	CNN	DE: 83.98 Raw EEG: 89.01
[66]	EEG	SEED/15	Negative, Neutral, Positive	Film clips	Dual-tree Complex Wavelet Transform (DT-CWT)	Recurrent Neural Network (RNN)	DEAP Valence: 97.97 Arousal: 98.31 Dominance: 98.32
[91]	PPG, HR, and GSR/ EDA	DEAP	Amusement (joy, happy, fun) and Sadness (depressing, sadness)	Video	Short-time Fast Fourier Transform (ST-FFT), Random-forest recursive feature elimination, Genetic Algorithm	SVM	DREAMER Valence: 94.59 Arousal: 95.26 Dominance: 95.13
[145]	EEG , BVP	MAHNOB/ 17	Arousal, Valence	Video	Sliding window strategy	CNN, LSTM	83.13
[146]	EEG	DEAP/ MAHNOB-HCI	Arousal, Valence	Video	LRFS	LSSVM, NB	97.78
[147]	EEG, ECG,GSR	AMIGOS	Arousal, Valence	Video	-	LSTM-RNN	71.61 \pm 2.71
[148]	ECG	AMIGOS, DREAMER, WESAD, SWELL.	Arousal, Valence, Stress	Video	Self-supervised network	Self-Supervised CNN	Arousal: 0.65/0.67 Valence: 0.68/0.70
[149]	EEG	DEAP	Valence, Arousal, Dominance, and Liking	Video	Eigenvector matrix linear discriminant analysis	Adaboost	Arousal: 83.3 Valence: 79.4
[150]	EEG	DEAP	High and Low arousal	Video	Empirical Mode Decomposition	SVM and Multilayer Perceptron	AMIGOS Arousal: 79.6 Valence: 78.3
[151]	EEG	DEAP/32	Happy, Pleased, Relaxed, Excited, Neutral, Calm, Distressed, Miserable, and Depressed	Video	time, wavelet, and frequency	SVM	DREAMER Arousal: 77.1 Valence: 74.9
							WESAD Affect State (95.0)
							SWELL: Arousal: 92.6 Valence: 93.8 Stress: 90.2

The systematic review shows that the researchers' most common biosensor data are ECG and EEG signals. Some of them preferred multimodality data. However, most of them used a single modality. The most common single modality data used is EEG from the DEAP dataset.

In recent work, even though deep learning models are becoming popular among researchers working with classifications, including ERS, ML continues to be chosen by researchers. From Table 3, it can be seen that ML such as SVM, RF, and NN are still commonly used in classifying emotions, where SVM classification accuracy is comparatively better than other ML models. In ML-based ERS, feature extraction is necessary to form the data signal as most physiological data is in time-series format. Deep learning requires less pre-processing compared to regular ML model training. That is the reason that motivates researchers to adopt deep learning so that the feature extraction and selection process can be avoided. Due to the capability of the deep learning feature extraction and selection process in hidden layers, few researchers also used it as a pre-processing technology.

The dataset used for training the learning models is developed by inducing emotion in individuals. Unconscious emotion inducement is commonly used rather than conscious. The most common materials used are music, movie clips, audio, images, and video games. The most common pictures/images used to induce emotion are obtained from the International Affective Picture System (IAPS) [152].

Both discrete and arousal/valence are popular in this field. However, only arousal and valence are not enough for a user to understand the individual's exact feelings. The rule of thumb for ML classification states that training samples should be significant to get the best classification result. The minimum number of participants used is six from the work tabulated above, and the maximum is 60. Different researchers obtained different accuracy and performance levels for different input data. It shows that there is no fixed method for the required case. A multimodal model, as researched by Jang et al. (2015) [116], was found to be the best model for the input and output data. Many physiological signals are non-stationary and chaotic. Commonly, time and frequency data can be extracted from these non-stationary physiological signals and reduce the impact of non-stationary characteristics on subsequent processing.

7- Conclusion and Future Work

In this paper, an initial systematic review was presented to answer research questions for developing ERS. It was concluded that the discrete and multi-dimensional emotional state models help distinguish between different emotional values, which are usually used as output classes of ERS. Input data can be obtained from various physiological signals (ECG, EEG, EMG, HRV, EOG, EDA/GSR/SC, SKT, RSP) by using biosensors. The ERS can be built using one source or a fusion of sources of signal data. However, single-modality is preferred due to reducing hardware cost. This paper describes the whole framework of ERS. Physiological measurement processing and data exploration methods play an essential role in selecting the best classification method and biosensors. Emotion classification can be implemented using traditional ML and deep learning models. ML model requires feature extraction and feature selection. Deep learning models automatically extract features instead of manual extraction. It is crucial to evaluate the trained classifiers' performance before integrating the model into the actual system. The evaluation can be processed using the confusion matrix value. Current research indicates that the most effective ERS methods are Deep Learning Models (CNN, DBN, PNN, LSTM, and SincNet). However, some traditional ML models (such as SVM and RF) can also provide classification with reasonable accuracy. According to the accuracy obtained by different researchers, there is room to improve the model performance. In particular, work with single modality needs to improve, as single input will reduce overall system expense.

The findings from this systematic review are going to be used as guidelines in building and designing the final ERS. In the future, the presented system will be implemented to detect emotions in real-time.

8- Declarations

8-1- Author Contributions

Conceptualization, T.K., and N.A.A.A.; methodology, T.K.; formal analysis, N.A.A.A.; investigation, N.A.A.A.; resources, T.K.; writing—original draft preparation, T.K., and J.M.Z.H; writing—review and editing, T.K., N.A.A.A., J.E.R, J.M.Z.H, and J.H.; visualization, T.K.; supervision, N.A.A.A., J.E.R, and J.H.; project administration, N.A.A.A.; funding acquisition, N.A.A.A. All authors have read and agreed to the published version of the manuscript.

8-2- Data Availability Statement

Data sharing is not applicable to this article.

8-3- Funding

This project is funded by TM Research & Development Grant (RDTC/190988), which is awarded to the Multimedia University.

8-4- Institutional Review Board Statement

Not applicable.

8-5- Informed Consent Statement

Not applicable.

8-6- Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

9- References

- [1] Mühlbacher-Karrer, S., Mosa, A. H., Faller, L. M., Ali, M., Hamid, R., Zangl, H., & Kyamakya, K. (2017). A Driver State Detection System - Combining a Capacitive Hand Detection Sensor with Physiological Sensors. *IEEE Transactions on Instrumentation and Measurement*, 66(4), 624–636. doi:10.1109/TIM.2016.2640458.
- [2] Ali, M., Al Machot, F., Mosa, A. H., & Kyamakya, K. (2016). A novel EEG-based emotion recognition approach for e-healthcare applications. *Proceedings of the 31st Annual ACM Symposium on Applied Computing*. doi:10.1145/2851613.2851916.
- [3] Ambach, W., & Gamer, M. (2018). *Physiological Measures in the Detection of Deception and Concealed Information. Detecting Concealed Information and Deception: Recent Developments*. Academic Press, Massachusetts, United States. doi:10.1016/B978-0-12-812729-2.00001-X.
- [4] Kang, S., Kim, D., & Kim, Y. (2019). A visual-physiology multimodal system for detecting outlier behavior of participants in a reality TV show. *International Journal of Distributed Sensor Networks*, 15(7). doi:10.1177/1550147719864886.
- [5] Kumar, A., Garg, N., & Kaur, G. (2019). An emotion recognition based on physiological signals. *International Journal of Innovative Technology and Exploring Engineering*, 8(9S), 335–341. doi:10.35940/ijitee.II054.0789S19.
- [6] Gravina, R., & Li, Q. (2019). Emotion-relevant activity recognition based on smart cushion using multi-sensor fusion. *Information Fusion*, 48, 1–10. doi:10.1016/j.inffus.2018.08.001.
- [7] Joseph, A., & Geetha, P. (2020). Facial emotion detection using modified Eyemap–Mouthmap algorithm on an enhanced image and classification with Tensorflow. *Visual Computer*, 36(3), 529–539. doi:10.1007/s00371-019-01628-3.
- [8] Ko, B. C. (2018). A brief review of facial emotion recognition based on visual information. *Sensors (Switzerland)*, 18(2). doi:10.3390/s18020401.
- [9] Udovičić, G., Derek, J., Russo, M., & Sikora, M. (2017). Wearable Emotion Recognition system based on GSR and PPG signals. *Proceedings of the 2nd International Workshop on Multimedia for Personal Health and Health Care*. doi:10.1145/3132635.3132641.
- [10] Joseph, C., & Lekamge, S. (2019). Machine Learning Approaches for Emotion Classification of Music: A Systematic Literature Review. *2019 International Conference on Advancements in Computing (ICAC)*. doi:10.1109/ICAC49085.2019.9103378.
- [11] Chen, D. W., Miao, R., Yang, W. Q., Liang, Y., Chen, H. H., Huang, L., Deng, C. J., & Han, N. (2019). A feature extraction method based on differential entropy and linear discriminant analysis for emotion recognition. *Sensors (Switzerland)*, 19(7), 1631. doi:10.3390/s19071631.
- [12] Bhise, P. R., Kulkarni, S. B., & Aldhaheri, T. A. (2020). Brain Computer Interface based EEG for Emotion Recognition System: A Systematic Review. *2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*. doi:10.1109/ICIMIA48430.2020.9074921.
- [13] Hasnul, M. A., Aziz, N. A. A., Alelyani, S., Mohana, M., & Aziz, A. A. (2021). Electrocardiogram-based emotion recognition systems and their applications in healthcare—a review. *Sensors*, 21(15). doi:10.3390/s21155015.
- [14] Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, 6(3–4), 169–200. doi:10.1080/02699939208411068.
- [15] Izard, C. E. (2007). Basic Emotions, Natural Kinds, Emotion Schemas, and a New Paradigm. *Perspectives on Psychological Science*, 2(3), 260–280. doi:10.1111/j.1745-6916.2007.00044.x.
- [16] Plutchik, R. (2001). The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American Scientist*, 89(4), 344–350. doi:10.1511/2001.4.344.
- [17] Acheampong, F. A., Wenyu, C., & Nunoo-Mensah, H. (2020). Text-based emotion detection: Advances, challenges, and opportunities. *Engineering Reports*, 2(7). doi:10.1002/eng2.12189.

- [18] Cambria, E., Livingstone, A., Hussain, A. (2012). The Hourglass of Emotions. Cognitive Behavioural Systems. Lecture Notes in Computer Science, vol 7403. Springer, Berlin, Germany. doi:10.1007/978-3-642-34584-5_11.
- [19] Lang, P. J. (1995). The Emotion Probe. American Psychologist Association, 50(5), 372–385. doi:10.1037/0003-066X.50.5.372.
- [20] Bailon, C., Damas, M., Pomares, H., Sanabria, D., Perakakis, P., Goicoechea, C., & Banos, O. (2018). Intelligent Monitoring of Affective Factors Underlying Sport Performance by Means of Wearable and Mobile Technology. Proceedings, 2(19), 1202. doi:10.3390/proceedings2191202.
- [21] Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. Development and Psychopathology, 17(3), 715–734. doi:10.1017/S0954579405050340.
- [22] Shu, L., Xie, J., Yang, M., Li, Z., Li, Z., Liao, D., Xu, X., & Yang, X. (2018). A review of emotion recognition using physiological signals. Sensors (Switzerland), 18(7). doi:10.3390/s18072074.
- [23] Furman, J. M., & Wuyts, F. L. (2012). Front matter. Aminoff's Electrodiagnosis in Clinical Neurology (6th Ed.). Saunders, Philadelphia, United States. doi:10.1016/b978-1-4557-0308-1.00041-8.
- [24] Khalifa, W. H., Roushdy, M. I., Abdel-badeeh, M. S., & Revett, K. (2015). AIS Inspired Approach for User Identification Based on EEG Signals. Recent Advances in Information Science AIS, 84-89.
- [25] Heng, H. M., Lu, M. K., Chou, L. W., Meng, N. H., Huang, H. C., Hamada, M., Tsai, C. H., & Chen, J. C. (2020). Changes in balance, gait and electroencephalography oscillations after robot-assisted gait training: An exploratory study in people with chronic stroke. Brain Sciences, 10(11), 1–13. doi:10.3390/brainsci10110821.
- [26] Alyasseri, Z. A. A., Khader, A. T., Al-Betar, M. A., Abasi, A. K., & Makhadmeh, S. N. (2020). EEG Signals Denoising Using Optimal Wavelet Transform Hybridized with Efficient Metaheuristic Methods. IEEE Access, 8(1), 10584–10605. doi:10.1109/ACCESS.2019.2962658.
- [27] Imad, A., Malik, N. A., Hamida, B. A., Seng, G. H. H., & Khan, S. (2022). Acoustic Photometry of Biomedical Parameters for Association with Diabetes and Covid-19. Emerging Science Journal, 6, 42-56. doi:10.28991/esj-2022-SPER-04.
- [28] Goshvarpour, A., Abbasi, A., & Goshvarpour, A. (2017). An Emotion Recognition Approach based on Wavelet Transform and Second-Order Difference Plot of ECG. Journal of AI and Data Mining, 5(2), 211–221. doi:10.22044/JADM.2017.887.
- [29] Tada, Y., Amano, Y., Sato, T., Saito, S., & Inoue, M. (2015). A smart shirt made with conductive ink and conductive foam for the measurement of electrocardiogram signals with unipolar precordial leads. Fibers, 3(4), 463–477. doi:10.3390/fib3040463.
- [30] Khan, G. M. (2015). A new electrode placement method for obtaining 12-lead ECGs. Open Heart, 2(1), e000226. doi:10.1136/openhrt-2014-000226.
- [31] Zong, C., & Chetouani, M. (2009). Hilbert-Huang transform based physiological signals analysis for emotion recognition. 2009 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT). doi:10.1109/isspit.2009.5407547.
- [32] Ladakis, I., & Chouvarda, I. (2021). Overview of biosignal analysis methods for the assessment of stress. Emerging Science Journal, 5(2), 233-244. doi:10.28991/esj-2021-01267.
- [33] Matzke, B., Herpertz, S. C., Berger, C., Fleischer, M., & Domes, G. (2014). Facial reactions during emotion recognition in borderline personality disorder: A facial electromyography study. Psychopathology, 47(2), 101–110. doi:10.1159/000351122.
- [34] Murata, A., Saito, H., Schug, J., Ogawa, K., & Kameda, T. (2016). Spontaneous facial mimicry is enhanced by the goal of inferring emotional states: Evidence for moderation of “automatic” mimicry by higher cognitive processes. PLoS ONE, 11(4). doi:10.1371/journal.pone.0153128.
- [35] Turabzadeh, S., Meng, H., Swash, R., Pleva, M., & Juhar, J. (2018). Facial Expression Emotion Detection for Real-Time Embedded Systems. Technologies, 6(1), 17. doi:10.3390/technologies6010017.
- [36] Huang, Y., Chen, F., Lv, S., & Wang, X. (2019). Facial expression recognition: A survey. Symmetry, 11(10), 1189. doi:10.3390/sym11101189.
- [37] Hsieh, P. Y., & Chin, C. L. (2011). The emotion recognition system with Heart Rate Variability and facial image features. 2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011). doi:10.1109/FUZZY.2011.6007734.
- [38] Huang, C., Liew, S. S., Lin, G. R., Poulsen, A., Ang, M. J., Chia, B. C., ... & Foo, K. (2019). Discovery of irreversible inhibitors targeting histone methyltransferase, SMYD3. ACS Medicinal Chemistry Letters, 10(6), 978-984. doi:10.1021/acsmedchemlett.9b00170.
- [39] Benezeth, Y., Li, P., Macwan, R., Nakamura, K., Gomez, R., & Yang, F. (2018). Remote heart rate variability for emotional state monitoring. 2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI). doi:10.1109/bhi.2018.8333392.

- [40] Perpetuini, D., Chiarelli, A. M., Maddiona, L., Rinella, S., Bianco, F., Bucciarelli, V., ... & Fallica, G. (2019). Multi-site photoplethysmographic and electrocardiographic system for arterial stiffness and cardiovascular status assessment. *Sensors*, 19(24), 5570. doi:10.3390/s19245570.
- [41] Wang, W., Den Brinker, A. C., Stuijk, S., & De Haan, G. (2017). Algorithmic Principles of Remote PPG. *IEEE Transactions on Biomedical Engineering*, 64(7), 1479–1491. doi:10.1109/TBME.2016.2609282.
- [42] Lord, M. P., & Wright, W. D. (1950). The investigation of eye movements. *Reports on Progress in Physics*, 13(1), 1–23. doi:10.1088/0034-4885/13/1/301.
- [43] R.Aguíñaga, A., Lopez Ramirez, M., Alanis Garza, A., Baltazar, R., & M. Zamudio, V. (2013). Emotion analysis through physiological measurements. *Workshop Proceedings of the 9th International Conference on Intelligent Environments J.A. Botía and D. Charitos (Eds.)*, 97–106. doi:10.3233/978-1-61499-286-8-97.
- [44] Picot, A., Charbonnier, S., & Caplier, A. (2011). EOG-based drowsiness detection: Comparison between a fuzzy system and two supervised learning classifiers. *IFAC Proceedings Volumes*, 44(1), 14283–14288. doi:10.3182/20110828-6-it-1002.00706.
- [45] Ramkumar, S., Sathesh Kumar, K., Dhiliphan Rajkumar, T., Ilayaraja, M., & Shankar, K. (2018). A review-classification of electrooculogram based human computer interfaces. *Biomedical Research (India)*, 29(6), 1078–1084. doi:10.4066/biomedicalresearch.29-17-2979.
- [46] Siddiqui, U., & Shaikh, A. N. (2013). An overview of electrooculography. *International Journal of Advanced Research in Computer and Communication Engineering*, 2(11), 4328-4330.
- [47] Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30(3), 261–273. doi:10.1111/j.1469-8986.1993.tb03352.x.
- [48] Wu, G., Liu, G., & Hao, M. (2010). The Analysis of Emotion Recognition from GSR Based on PSO. *2010 International Symposium on Intelligence Information Processing and Trusted Computing*. doi:10.1109/iptc.2010.60.
- [49] van Dooren, M., de Vries, J. J. G. G. J., & Janssen, J. H. (2012). Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. *Physiology and Behavior*, 106(2), 298–304. doi:10.1016/j.physbeh.2012.01.020.
- [50] Kosonogov, V., De Zorzi, L., Honoré, J., Martínez-Velázquez, E. S., Nandrino, J.-L., Martinez-Selva, J. M., & Sequeira, H. (2017). Facial thermal variations: A new marker of emotional arousal. *PlosOne*, 12(9), e0183592. doi:10.1371/journal.pone.0183592.
- [51] Ayata, D., Yaslan, Y., & Kamaşak, M. (2017). Emotion recognition via galvanic skin response: Comparison of machine learning algorithms and feature extraction methods. *IU-Journal of Electrical & Electronics Engineering*, 17(1), 3147-3156.
- [52] Daiana da Costa, T., de Fatima Fernandes Vara, M., Santos Cristino, C., Zoraski Zanella, T., Nunes Nogueira Neto, G., & Nohama, P. (2019). Breathing Monitoring and Pattern Recognition with Wearable Sensors. *Wearable Devices - the Big Wave of Innovation*, IntechOpen, London, United Kingdom. doi:10.5772/intechopen.85460.
- [53] Kim, J., & André, E. (2008). Emotion recognition based on physiological changes in music listening. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(12), 2067–2083. doi:10.1109/TPAMI.2008.26.
- [54] Norali, A. N., Abdullah, A. H., Zakaria, Z., Rahim, N. A., & Nataraj, S. K. (2017). Human breathing assessment using Electromyography signal of respiratory muscles. *2016 6th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*. doi:10.1109/ICCSCE.2016.7893596.
- [55] Kawde, P., & Verma, G. K. (2017). Deep belief network based affect recognition from physiological signals. *2017 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON)*. doi:10.1109/UPCON.2017.8251115.
- [56] Zhang, X., Xu, C., Xue, W., Hu, J., He, Y., & Gao, M. (2018). Emotion recognition based on multichannel physiological signals with comprehensive nonlinear processing. *Sensors (Switzerland)*, 18(11), 1–16. doi:10.3390/s18113886.
- [57] Jerri, S., Murugappan, M., Wan, K., & Yaacob, S. (2014). Emotion recognition from facial EMG signals using higher order statistics and principal component analysis. *Journal of the Chinese Institute of Engineers, Transactions of the Chinese Institute of Engineers, Series A*, 37(3), 385–394. doi:10.1080/02533839.2013.799946.
- [58] Gong, P., Ma, H. T., & Wang, Y. (2016). Emotion recognition based on the multiple physiological signals. *2016 IEEE International Conference on Real-time Computing and Robotics (RCAR)*. doi:10.1109/RCAR.2016.7784015.
- [59] Izard, C. E. (2009). Emotion theory and research: Highlights, unanswered questions, and emerging issues. *Annual Review of Psychology*, 60(60), 1–25. doi:10.1146/annurev.psych.60.110707.163539.
- [60] Suchetha, M., Kumaravel, N., Jagannath, M., & Jaganathan, S. K. (2017). A comparative analysis of EMD based filtering methods for 50 Hz noise cancellation in ECG signal. *Informatics in Medicine Unlocked*, 8, 54–59. doi:10.1016/j.imu.2017.01.003.

- [61] Lahmirmi, S., & Boukadoum, M. (2015). Physiological signal denoising with variational mode decomposition and weighted reconstruction after DWT thresholding. 2015 IEEE International Symposium on Circuits and Systems (ISCAS). doi:10.1109/ISCAS.2015.7168756.
- [62] Mannan, M. M. N., Kamran, M. A., & Jeong, M. Y. (2018). Identification and removal of physiological artifacts from electroencephalogram signals: A review. IEEE Access, 6, 30630–30652. doi:10.1109/ACCESS.2018.2842082.
- [63] Hindarto, H., & Sumarno, S. (2016). Feature Extraction of Electroencephalography Signals Using Fast Fourier Transform. CommIT (Communication and Information Technology) Journal, 10(2), 49. doi:10.21512/commit.v10i2.1548.
- [64] Al-Fahoum, A. S., & Al-Fraihat, A. A. (2014). Methods of EEG Signal Features Extraction Using Linear Analysis in Frequency and Time-Frequency Domains. ISRN Neuroscience, 2014, 1–7. doi:10.1155/2014/730218.
- [65] Vanny, M., Park, S.M., Ko, K. E., Sim, K.B. (2013). Analysis of Physiological Signals for Emotion Recognition Based on Support Vector Machine. Robot Intelligence Technology and Applications 2012. Advances in Intelligent Systems and Computing, 208. Springer, Berlin, Germany. doi:10.1007/978-3-642-37374-9_12.
- [66] Wei, C., Chen, L. Lan, Song, Z. Zhen, Lou, X. Guang, & Li, D. dong. (2020). EEG-based emotion recognition using simple recurrent units network and ensemble learning. Biomedical Signal Processing and Control, 58, 101756. doi:10.1016/j.bspc.2019.101756.
- [67] Bal, U. (2012). Dual tree complex wavelet transform based denoising of optical microscopy images. Biomedical Optics Express, 3(12), 3231. doi:10.1364/boe.3.003231.
- [68] Chen, T., Ju, S., Yuan, X., Elhoseny, M., Ren, F., Fan, M., & Chen, Z. (2018). Emotion recognition using empirical mode decomposition and approximation entropy. Computers and Electrical Engineering, 72, 383–392. doi:10.1016/j.compeleceng.2018.09.022.
- [69] Jerri, S., Murugappan, M., Wan, K., & Yaacob, S. (2014). Electrocardiogram-based emotion recognition system using empirical mode decomposition and discrete Fourier transform. Expert Systems, 31(2), 110–120. doi:10.1111/exsy.12014.
- [70] Liu, T., Luo, Z., Huang, J., & Yan, S. (2018). A comparative study of four kinds of adaptive decomposition algorithms and their applications. Sensors (Switzerland), 18(7), 1–51. doi:10.3390/s18072120.
- [71] Hassan, M. M., Alam, M. G. R., Uddin, M. Z., Huda, S., Almogren, A., & Fortino, G. (2019). Human emotion recognition using deep belief network architecture. Information Fusion, 51, 10–18. doi:10.1016/j.inffus.2018.10.009.
- [72] Asghar, M. A., Khan, M. J., Fawad, Amin, Y., Rizwan, M., Rahman, M., ... Mirjavadi, S. S. (2019). EEG-Based Multi-Modal Emotion Recognition using Bag of Deep Features: An Optimal Feature Selection Approach. Sensors, 19(23), 5218. doi:10.3390/s19235218.
- [73] Huang, C., Gong, W., Fu, W., & Feng, D. (2014). A research of speech emotion recognition based on deep belief network and SVM. Mathematical Problems in Engineering, 2014, 1-7. doi:10.1155/2014/749604.
- [74] Bengio, Y., Lamblin, P., Popovici, D., & Larochelle, H. (2006). Greedy layer-wise training of deep networks. Advances in Neural Information Processing Systems, NeurIPS Proceedings, 153–160.
- [75] Dissanayake, T., Rajapaksha, Y., Ragel, R., & Nawinne, I. (2019). An ensemble learning approach for electrocardiogram sensor based human emotion recognition. Sensors (Switzerland), 19(20), 1–24. doi:10.3390/s19204495.
- [76] Colomer Granero, A., Fuentes-Hurtado, F., Naranjo Ornedo, V., Guixeres Provinciale, J., Ausín, J. M., & Alcañiz Raya, M. (2016). A comparison of physiological signal analysis techniques and classifiers for automatic emotional evaluation of audiovisual contents. Frontiers in Computational Neuroscience, 10, 1–14. doi:10.3389/fncom.2016.00074.
- [77] Deb, S., Rabiul Islam, S. M., Johura, F. T., & Huang, X. (2018). Extraction of Linear and Non-Linear Features of Electrocardiogram Signal and Classification. 2017 2nd International Conference on Electrical & Electronic Engineering (ICEEE). doi:10.1109/ICEEE.2017.8412857.
- [78] Shukla, J., Barreda-Angeles, M., Oliver, J., Nandi, G. C., & Puig, D. (2021). Feature Extraction and Selection for Emotion Recognition from Electrodermal Activity. IEEE Transactions on Affective Computing, 12(4), 857–869. doi:10.1109/TAFFC.2019.2901673.
- [79] Konar, A., & Chakraborty, A. (2015). Emotion recognition: A pattern analysis approach. John Wiley & Sons, Hoboken, United States. doi:10.1002/9781118910566.
- [80] Mehmood, R. M., & Lee, H. J. (2015). Exploration of prominent frequency wave in EEG signals from brain sensors network. International Journal of Distributed Sensor Networks, 11(11). doi:10.1155/2015/386057.
- [81] Kyriakou, K., Resch, B., Sagl, G., Petutschnig, A., Werner, C., Niederseer, D., ... & Pykett, J. (2019). Detecting moments of stress from measurements of wearable physiological sensors. Sensors, 19(17), 3805. doi:10.3390/s19173805.

- [82] Kołodziej, M., Tarnowski, P., Majkowski, A., & Rak, R. J. (2019). Electrodermal activity measurements for detection of emotional arousal. *Bulletin of the Polish Academy of Sciences: Technical Sciences*, 67(4), 813–826. doi:10.24425/bpasts.2019.130190.
- [83] Cho, D., Ham, J., Oh, J., Park, J., Kim, S., Lee, N. K., & Lee, B. (2017). Detection of stress levels from biosignals measured in virtual reality environments using a kernel-based extreme learning machine. *Sensors (Switzerland)*, 17(10). doi:10.3390/s17102435.
- [84] Gómez-Lara, J. F., Ordóñez-Bolaños, O. A., Becerra, M. A., Castro-Ospina, A. E., Mejía-Arboleda, C., Duque-Mejía, C., ... Peluffo-Ordóñez, D. H. (2019). Feature Extraction Analysis for Emotion Recognition from ICEEMD of Multimodal Physiological Signals. *Lecture Notes in Computer Science*, 351–362. doi:10.1007/978-3-030-14799-0_30.
- [85] Peker, M., Arslan, A., Sen, B., Celebi, F. V., & But, A. (2015). A novel hybrid method for determining the depth of anesthesia level: Combining ReliefF feature selection and random forest algorithm (ReliefF+RF). *2015 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA)*. doi:10.1109/inista.2015.7276737.
- [86] Elsayyad, A., Al-Dhaifallah, M., & Nassef, A. M. (2017). Features selection for arrhythmia diagnosis using Relief-F algorithm and support vector machine. *2017 14th International Multi-Conference on Systems, Signals & Devices (SSD)*. doi:10.1109/SSD.2017.8166920.
- [87] Goshvarpour, A., Abbasi, A., & Goshvarpour, A. (2017). An accurate emotion recognition system using ECG and GSR signals and matching pursuit method. *Biomedical Journal*, 40(6), 355–368. doi:10.1016/j.bj.2017.11.001.
- [88] Smith, G. (2018). Step away from stepwise. *Journal of Big Data*, 5(1). doi:10.1186/s40537-018-0143-6.
- [89] Xu, L., Redman, C. W. G., Payne, S. J., & Georgieva, A. (2014). Feature selection using genetic algorithms for fetal heart rate analysis. *Physiological Measurement*, 35(7), 1357–1371. doi:10.1088/0967-3334/35/7/1357.
- [90] Romeo, L., Cavallo, A., Pepa, L., Bianchi-Berthouze, N., & Pontil, M. (2022). Multiple Instance Learning for Emotion Recognition Using Physiological Signals. *IEEE Transactions on Affective Computing*, 13(1), 389–407. doi:10.1109/TAFFC.2019.2954118.
- [91] Domínguez-Jiménez, J. A., Campo-Landines, K. C., Martínez-Santos, J. C., Delahoz, E. J., & Contreras-Ortiz, S. H. (2020). A machine learning model for emotion recognition from physiological signals. *Biomedical Signal Processing and Control*, 55, 101646. doi:10.1016/j.bspc.2019.101646.
- [92] Chang, C. C., & Lin, C. J. (2011). LIBSVM: A Library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3), 1–39. doi:10.1145/1961189.1961199.
- [93] Chen, T., Ju, S., Ren, F., Fan, M., & Gu, Y. (2020). EEG emotion recognition model based on the LIBSVM classifier. *Measurement*, 164(108047). doi:10.1016/j.measurement.2020.108047.
- [94] Martinez, M., Schauerte, B., Stiefelwagen, R. (2013). “BAM!” Depth-Based Body Analysis in Critical Care. *Computer Analysis of Images and Patterns. CAIP 2013. Lecture Notes in Computer Science*, 8047. Springer, Berlin, Germany. doi:10.1007/978-3-642-40261-6_56.
- [95] Salari, S., Ansarian, A., & Atrianfar, H. (2018). Robust emotion classification using neural network models. *2018 6th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)*. doi:10.1109/CFIS.2018.8336626.
- [96] Song, T., Zheng, W., Song, P., & Cui, Z. (2020). EEG Emotion Recognition Using Dynamical Graph Convolutional Neural Networks. *IEEE Transactions on Affective Computing*, 11(3), 532–541. doi:10.1109/TAFFC.2018.2817622.
- [97] Santamaria-Granados, L., Munoz-Organero, M., Ramirez-Gonzalez, G., Abdulhay, E., & Arunkumar, N. (2019). Using Deep Convolutional Neural Network for Emotion Detection on a Physiological Signals Dataset (AMIGOS). *IEEE Access*, 7(1), 57–67. doi:10.1109/ACCESS.2018.2883213.
- [98] Al Machot, F., Elmachot, A., Ali, M., Al Machot, E., & Kyamakya, K. (2019). A deep-learning model for subject-independent human emotion recognition using electrodermal activity sensors. *Sensors (Switzerland)*, 19(7), 1–14. doi:10.3390/s19071659.
- [99] Pandey, S. K., & Janghel, R. R. (2019). Recent deep learning techniques, challenges and its applications for medical healthcare system: a review. *Neural Processing Letters*, 50(2), 1907–1935. doi:10.1007/s11063-018-09976-2.
- [100] Chao, H., Zhi, H., Dong, L., & Liu, Y. (2018). Recognition of Emotions Using Multichannel EEG Data and DBN-GC-Based Ensemble Deep Learning Framework. *Computational Intelligence and Neuroscience*, 2018, 1–11. doi:10.1155/2018/9750904.
- [101] Huang, J., Xu, X., & Zhang, T. (2017). Emotion classification using deep neural networks and emotional patches. *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. doi:10.1109/BIBM.2017.8217786.
- [102] Nakisa, B., Rastgoo, M. N., Tjondronegoro, D., & Chandran, V. (2018). Evolutionary computation algorithms for feature selection of EEG-based emotion recognition using mobile sensors. *Expert Systems with Applications*, 93, 143–155. doi:10.1016/j.eswa.2017.09.062.

- [103] Wöllmer, M., Kaiser, M., Eyben, F., Schuller, B., & Rigoll, G. (2013). LSTM-modeling of continuous emotions in an audiovisual affect recognition framework. *Image and Vision Computing*, 31(2), 153–163. doi:10.1016/j.imavis.2012.03.001.
- [104] Li, X., Song, D., Zhang, P., Yu, G., Hou, Y., & Hu, B. (2016). Emotion recognition from multi-channel EEG data through Convolutional Recurrent Neural Network. 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). doi:10.1109/bibm.2016.7822545.
- [105] Li, Z., Tian, X., Shu, L., Xu, X., & Hu, B. (2018). Emotion recognition from EEG using RASM and LSTM. *Communications in Computer and Information Science*, 819, 310–318. doi:10.1007/978-981-10-8530-7_30.
- [106] Xing, X., Li, Z., Xu, T., Shu, L., Hu, B., & Xu, X. (2019). SAE+LSTM: A new framework for emotion recognition from multi-channel EEG. *Frontiers in Neurorobotics*, 13(37), 1–14. doi:10.3389/fnbot.2019.00037.
- [107] Alhagry, S., Fahmy, A. A., & El-Khoribi, R. A. (2017). Emotion recognition based on EEG using LSTM recurrent neural network. *International Journal of Advanced Computer Science and Applications*, 8(10). doi:10.14569/ijacsa.2017.081046.
- [108] Ravanelli, M., & Bengio, Y. (2018). Speaker recognition from raw waveform with sincnet. 2018 IEEE Spoken Language Technology Workshop (SLT). doi:10.1109/SLT.2018.8639585.
- [109] Zeng, H., Wu, Z., Zhang, J., Yang, C., Zhang, H., Dai, G., & Kong, W. (2019). EEG emotion classification using an improved sincnet-based deep learning model. *Brain Sciences*, 9(11). doi:10.3390/brainsci9110326.
- [110] Gu, Y., Tan, S. L., Wong, K. J., Ho, M. H. R., & Qu, L. (2009). Using GA-based feature selection for emotion recognition from physiological signals. 2008 International Symposium on Intelligent Signal Processing and Communications Systems. doi:10.1109/ISPACS.2009.4806747.
- [111] Maaoui, C., & Pruski, A. (2010). Emotion Recognition through Physiological Signals for Human-Machine Communication. *Cutting Edge Robotics 2010*, 317–333. doi:10.5772/10312.
- [112] Agraftioti, F., Hatzinakos, D., & Anderson, A. K. (2012). ECG pattern analysis for emotion detection. *IEEE Transactions on Affective Computing*, 3(1), 102–115. doi:10.1109/T-AFFC.2011.28.
- [113] Selvaraj, J., Murugappan, M., Wan, K., & Yaacob, S. (2013). Classification of emotional states from electrocardiogram signals: A non-linear approach based on hurst. *BioMedical Engineering Online*, 12(1), 1–18. doi:10.1186/1475-925X-12-44.
- [114] Gruebler, A., & Suzuki, K. (2014). Design of a wearable device for reading positive expressions from facial EMG signals. *IEEE Transactions on Affective Computing*, 5(3), 227–237. doi:10.1109/TAFFC.2014.2313557.
- [115] Gouizi, K., Maaoui, C., & Bereksi Reguig, F. (2014). Negative emotion detection using EMG signal. 2014 International Conference on Control, Decision and Information Technologies (CoDIT). doi:10.1109/CoDIT.2014.6996980.
- [116] Jang, E. H., Park, B. J., Park, M. S., Kim, S. H., & Sohn, J. H. (2015). Analysis of physiological signals for recognition of boredom, pain, and surprise emotions. *Journal of Physiological Anthropology*, 34(1), 1–12. doi:10.1186/s40101-015-0063-5.
- [117] Soleymani, M., Asghari-Esfeden, S., Fu, Y., & Pantic, M. (2016). Analysis of EEG Signals and Facial Expressions for Continuous Emotion Detection. *IEEE Transactions on Affective Computing*, 7(1), 17–28. doi:10.1109/TAFFC.2015.2436926.
- [118] Guo, H. W., Huang, Y. S., Lin, C. H., Chien, J. C., Haraikawa, K., & Shieh, J. S. (2016). Heart Rate Variability Signal Features for Emotion Recognition by Using Principal Component Analysis and Support Vectors Machine. 2016 IEEE 16th International Conference on Bioinformatics and Bioengineering (BIBE). doi:10.1109/BIBE.2016.40.
- [119] He, C., Yao, Yj., Ye, Xs. (2017). An Emotion Recognition System Based on Physiological Signals Obtained by Wearable Sensors. *Wearable Sensors and Robots. Lecture Notes in Electrical Engineering*, 399. Springer, Singapore. doi:10.1007/978-981-10-2404-7_2.
- [120] Atkinson, J., & Campos, D. (2016). Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers. *Expert Systems with Applications*, 47, 35–41. doi:10.1016/j.eswa.2015.10.049.
- [121] Mehmood, R. M., & Lee, H. J. (2016). A novel feature extraction method based on late positive potential for emotion recognition in human brain signal patterns. *Computers and Electrical Engineering*, 53, 444–457. doi:10.1016/j.compeleceng.2016.04.009.
- [122] Park, Y. L. (2017). Soft wearable robotics technologies for body motion sensing. *Human Modelling for Bio-Inspired Robotics*, 161-184. Academic Press, Massachusetts, United States. doi:10.1016/B978-0-12-803137-7.00009-4.
- [123] Greco, A., Valenza, G., Citi, L., & Scilingo, E. P. (2017). Arousal and valence recognition of affective sounds based on electrodermal activity. *IEEE Sensors Journal*, 17(3), 716–725. doi:10.1109/JSEN.2016.2623677.
- [124] Chen, J., Hu, B., Wang, Y., Moore, P., Dai, Y., Feng, L., & Ding, Z. (2017). Subject-independent emotion recognition based on physiological signals: A three-stage decision method. *BMC Medical Informatics and Decision Making*, 17. doi:10.1186/s12911-017-0562-x.

- [125] Yin, Z., Zhao, M., Wang, Y., Yang, J., & Zhang, J. (2017). Recognition of emotions using multimodal physiological signals and an ensemble deep learning model. *Computer Methods and Programs in Biomedicine*, 140, 93–110. doi:10.1016/j.cmpb.2016.12.005.
- [126] Liu, X., Wang, Q., Liu, D., Wang, Y., Zhang, Y., Bai, O., & Sun, J. (2018). Human emotion classification based on multiple physiological signals by wearable system. *Technology and Health Care*, 26, 459–469. doi:10.3233/THC-174747.
- [127] Al Zoubi, O., Awad, M., & Kasabov, N. K. (2018). Anytime multipurpose emotion recognition from EEG data using a Liquid State Machine based framework. *Artificial intelligence in medicine*, 86, 1-8. doi:10.1016/j.artmed.2018.01.001.
- [128] Nakisa, B., Rastgoo, M. N., Rakotonirainy, A., Maire, F., & Chandran, V. (2018). Long short term memory hyperparameter optimization for a neural network based emotion recognition framework. *IEEE Access*, 6(1), 49325–49338. doi:10.1109/ACCESS.2018.2868361.
- [129] Kaur, B., Singh, D., & Roy, P. P. (2018). EEG Based Emotion Classification Mechanism in BCI. *Procedia Computer Science*, 132, 752–758. doi:10.1016/j.procs.2018.05.087.
- [130] Ayata, D., Yaslan, Y., & Kamasak, M. E. (2018). Emotion Based Music Recommendation System Using Wearable Physiological Sensors. *IEEE Transactions on Consumer Electronics*, 64(2), 196–203. doi:10.1109/TCE.2018.2844736.
- [131] Bagherzadeh, S., Maghooli, K., Farhadi, J., & Zangeneh Soroush, M. (2018). Emotion Recognition from Physiological Signals Using Parallel Stacked Autoencoders. *Neurophysiology*, 50(6), 428–435. doi:10.1007/s11062-019-09775-y.
- [132] Chettupuzhakkaran, P., & Sindhu, N. (2018). Emotion recognition from physiological signals using time-frequency analysis methods. *2018 International Conference on Emerging Trends and Innovations in Engineering and Technological Research (ICETIETR)*. doi:10.1109/ICETIETR.2018.8529145.
- [133] Gupta, V., Chopda, M. D., & Pachori, R. B. (2019). Cross-Subject Emotion Recognition Using Flexible Analytic Wavelet Transform from EEG Signals. *IEEE Sensors Journal*, 19(6), 2266–2274. doi:10.1109/JSEN.2018.2883497.
- [134] Albraikan, A., Tobon, D. P., & El Saddik, A. (2019). Toward User-Independent Emotion Recognition Using Physiological Signals. *IEEE Sensors Journal*, 19(19), 8402–8412. doi:10.1109/JSEN.2018.2867221.
- [135] Xing, B., Zhang, H., Zhang, K., Zhang, L., Wu, X., Shi, X., Yu, S., & Zhang, S. (2019). Exploiting EEG Signals and Audiovisual Feature Fusion for Video Emotion Recognition. *IEEE Access*, 7(1), 59844–59861. doi:10.1109/ACCESS.2019.2914872.
- [136] Tiwari, S., Agarwal, S., Adiyarta, K., & Syafrullah, M. (2019). Classification of physiological signals for emotion recognition using IoT. *2019 6th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*. doi:10.23919/EECSI48112.2019.8977062.
- [137] Taran, S., & Bajaj, V. (2019). Emotion recognition from single-channel EEG signals using a two-stage correlation and instantaneous frequency-based filtering method. *Computer Methods and Programs in Biomedicine*, 173, 157–165. doi:10.1016/j.cmpb.2019.03.015.
- [138] Pandey, P., & Seeja, K. R. (2022). Subject independent emotion recognition from EEG using VMD and deep learning. *Journal of King Saud University - Computer and Information Sciences*, 34(5), 1730–1738. doi:10.1016/j.jksuci.2019.11.003.
- [139] Shu, L., Yu, Y., Chen, W., Hua, H., Li, Q., Jin, J., & Xu, X. (2020). Wearable emotion recognition using heart rate data from a smart bracelet. *Sensors (Switzerland)*, 20(3), 1–19. doi:10.3390/s20030718.
- [140] Wang, F., Wu, S., Zhang, W., Xu, Z., Zhang, Y., Wu, C., & Coleman, S. (2020). Emotion recognition with convolutional neural network and EEG-based EFDMs. *Neuropsychologia*, 146, 107506. doi:10.1016/j.neuropsychologia.2020.107506.
- [141] Salama, E. S., El-Khoribi, R. A., Shoman, M. E., & Wahby Shalaby, M. A. (2021). A 3D-convolutional neural network framework with ensemble learning techniques for multi-modal emotion recognition. *Egyptian Informatics Journal*, 22(2), 167–176. doi:10.1016/j.eij.2020.07.005.
- [142] Ayata, D., Yaslan, Y., & Kamasak, M. E. (2020). Emotion Recognition from Multimodal Physiological Signals for Emotion Aware Healthcare Systems. *Journal of Medical and Biological Engineering*, 40(2), 149–157. doi:10.1007/s40846-019-00505-7.
- [143] Fourati, R., Ammar, B., Sanchez-Medina, J., & Alimi, A. M. (2022). Unsupervised Learning in Reservoir Computing for EEG-Based Emotion Recognition. *IEEE Transactions on Affective Computing*, 13(2), 972–984. doi:10.1109/TAFFC.2020.2982143.
- [144] Liu, Y., Ding, Y., Li, C., Cheng, J., Song, R., Wan, F., & Chen, X. (2020). Multi-channel EEG-based emotion recognition via a multi-level features guided capsule network. *Computers in Biology and Medicine*, 123(March), 103927. doi:10.1016/j.combiomed.2020.103927.
- [145] Nakisa, B., Rastgoo, M. N., Rakotonirainy, A., Maire, F., & Chandran, V. (2020). Automatic emotion recognition using temporal multimodal deep learning. *IEEE Access*, 8, 225463–225474. doi:10.1109/ACCESS.2020.3027026.

- [146] Yin, Z., Liu, L., Chen, J., Zhao, B., & Wang, Y. (2020). Locally robust EEG feature selection for individual-independent emotion recognition. *Expert Systems with Applications*, 162, 113768. doi:10.1016/j.eswa.2020.113768.
- [147] Li, C., Bao, Z., Li, L., & Zhao, Z. (2020). Exploring temporal representations by leveraging attention-based bidirectional LSTM-RNNs for multi-modal emotion recognition. *Information Processing and Management*, 57(3), 102185. doi:10.1016/j.ipm.2019.102185.
- [148] Sarkar, P., & Etemad, A. (2020). Self-supervised ECG representation learning for emotion recognition. *IEEE Transactions on Affective Computing*, 1(1). doi:10.1109/TAFFC.2020.3014842.
- [149] Chen, Y., Chang, R., & Guo, J. (2021). Emotion Recognition of EEG Signals Based on the Ensemble Learning Method: AdaBoost. *Mathematical Problems in Engineering*, 2021, 1-12. doi:10.1155/2021/8896062.
- [150] Salankar, N., Mishra, P., & Garg, L. (2021). Emotion recognition from EEG signals using empirical mode decomposition and second-order difference plot. *Biomedical Signal Processing and Control*, 65, 102389. doi:10.1016/j.bspc.2020.102389.
- [151] Khateeb, M., Anwar, S. M., & Alnowami, M. (2021). Multi-Domain Feature Fusion for Emotion Classification Using DEAP Dataset. *IEEE Access*, 9, 12134–12142. doi:10.1109/ACCESS.2021.3051281.
- [152] Lang, P., & Bradley, M. M. (2007). The International Affective Picture System (IAPS) in the study of emotion and attention. *Handbook of emotion elicitation and assessment*, 29, 70-73, Oxford University Press, Oxford, United Kingdom.