Iraqi Journal for Computer Science and Mathematics

Journal Homepage: http://journal.esj.edu.iq/index.php/IJCM
e-ISSN: 2788-7421 p-ISSN: 2958-0544

Emotion Recognition Using Various Measures and Computational Methods: Review

Abir Hamrouni[1*], Fatima Bendella[2]

[1]SIMPA Laboratory, Department of Computer Science,
Université des sciences et de la Technologie d'Oran Mohamed Boudiaf USTOMB, Oran, ALGERIA

SIMPA Laboratory, Department of Computer Science,
Université des sciences et de la Technologie d'Oran Mohamed Boudiaf USTOMB, Oran, ALGERIA

*Corresponding Author: Abir Hamrouni

DOI: https://doi.org/10.52866/ijcsm.2024.05.03.018
Received February 2024; Accepted May 2024; Available online July 2024

ABSTRACT: Emotion recognition has garnered significant attention as a burgeoning research domain, owing to its potential applications across diverse fields such as human-computer interaction, affective gaming, marketing, and human-robot interaction. Accurately interpreting and appropriately responding to human emotions remains a critical challenge in the development of systems. This obstacle necessitates a thorough understanding of emotions to enhance user experiences within such systems. This paper conducts a comprehensive review focusing on advancements in emotion recognition techniques, with an emphasis on leveraging a variety of sensors and computational methods. Our study findings highlight the significant improvement to emotion recognition accuracy when multiple measures and computational methods, rather than a single modality, is used. This article contributes to the field by thoroughly reviewing and comparing diverse measures and computational methods for emotion recognition. The study highlights the pivotal role of employing multiple modalities and advanced machine learning algorithms to achieve superior accuracy and reliability in emotion recognition. Furthermore, this research identifies potential avenues for further investigation and development, such as integrating multimodal data and exploring novel features and fusion techniques. The insights offered in this study provide valuable guidance for researchers and practitioners in the field, facilitating the advancement of technologies that adeptly understand and respond to human emotions.

Keywords: Emotions, Sensors, Emotions recognitions techniques, machine learning, deep learning, Computational methods

1. INTRODUCTION

Emotions play a crucial role in human interactions. They heavily influence our thoughts, behaviors, and overall experiences heavily. Nevertheless, there is no consensus on a definition of emotion in the literature yet "The struggle to define emotion in scientific terms is as old as the field of psychology" [1]. Many psychologists define emotion as a biological state that arise as a result of various internal and external stimuli, while others researchers define it as a brief conscious experience characterized by intense mental activity.

Emotion recognition, the ability to accurately perceive and interpret human emotions, has gained significant attention in recent years due to its potential applications in diverse fields, ranging from mental health and human-computer interaction to personalized experiences and artificial intelligence. Researchers have made substantial progress in developing and deploying various sensors and computational methods to enhance the accuracy and effectiveness of emotion recognition.

In the realm of sensor-based approaches, multiple modalities have been explored to capture different aspects of emotional expressions. Facial expression analysis, a widely studied modality, has been instrumental in recognizing emotions. Traditional approaches like the Facial Action Coding System (FACS) have provided valuable insights into facial muscle movements and their association with specific emotions [2]. Recent advancements in this area include the utilization of 3D facial models and deep learning techniques to extract fine-grained facial features for more accurate emotion recognition [3] [4].

In addition to facial expressions, researchers have incorporated other physiological signals to complement the understanding of emotions. Electrodermal activity (EDA), which measures the skin's electrical conductivity, and heart rate variability (HRV), reflecting changes in heart rhythm, have been shown to correlate with emotional states [5][6].

The integration of multiple modalities, such as facial expressions, speech analysis, and physiological signals, has demonstrated improved emotion recognition performance [7]. Moreover, the growing availability of textual data from social media platforms has led to the exploration of natural language processing techniques for sentiment analysis and emotion detection [8] [9].

While sensor-based approaches have provided valuable insights, computational methods, particularly machine learning and deep learning algorithms, have revolutionized emotion recognition. Traditional machine learning models, such as Support Vector Machines (SVM) and Random Forests, have been widely used for emotion classification [10]. However, the emergence of deep learning techniques has yielded substantial advancements in this domain. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have shown exceptional performance in capturing spatial and temporal dependencies in emotional data, leading to improved emotion recognition accuracy [11] [12]. Moreover, the integration of attention mechanisms and transfer learning strategies has further enhanced the robustness and generalizability of deep learning models in emotion recognition tasks [13] [14].

While considerable progress has been made in both sensor-based and computational approaches, it is important to acknowledge that challenges persist in achieving a perfect and comprehensive understanding of human emotions. The subjectivity and cultural variations in emotional expressions pose significant hurdles in developing universally applicable emotion recognition systems [15]. Additionally, ensuring privacy and ethical considerations in the collection and analysis of emotional data remains a critical concern [16].

Therefore, this article aims to provide a comprehensive review of the recent advancements in sensors and computational methods used in emotion recognition. By surveying recent research works, we aim to present an up-to-date overview of the state of the art in this rapidly evolving field. We discuss the strengths, limitations, and future directions of emotion recognition, with a focus on recent innovations and emerging trends. The insights gained from this review will contribute to further advancements in emotion recognition technology and foster its responsible and meaningful integration into various domains.

2. Emotions Recognition Methods and Sensors

Emotion assessment methods presented in the literature of psychology and computer science can be divided into two main groups according to the underlying measures used for emotion detection: subjective measures are mainly based on self-reporting techniques by using the self-assessment approaches [17] [18]; objective measures based on human body signal processing and analysis techniques. The subjective measures themselves are commonly classified into self-assessment techniques and behavioral modalities. The most popular techniques of physiological measures are: electrocardiogram (ECG) to measure the heart rate [19], galvanic skin response (GSR) for skin resistance measurements [20], electroencephalogram (EEG) for cerebral activity measurements [21], blood volume pulse (BVP) for blood pressure measurements [22] and electrooculography (EOG) for eye tracker [23][24]. The behavioral modalities are the using of physical signals such as facial expression [25], speech [26], posture [27], etc.

2.1 Subjective Measures

Subjective measures are methods based on conscious responses, which are relatively simple and has the advantage of easy collection. The first modality which is the self-assessment techniques, are not reliable and it is quite possibly that a person does not correctly recognize his feelings or gives inaccurate answers to unpleasant questions. Behavioral modalities like facial expressions, speech, text expressions and posture are easy to be controlled by people. For example, a person can appear neutral when feeling frustrated or angry. Therefore, the reliability cannot be guaranteed.

2.1.1 Self-Assessment techniques

Self-assessment techniques generally utilize rating scale or pictograms to facilitate emotion identification by users. These instruments can be verbal or non-verbal. Verbal instruments are often used by psychologist while non-verbal ones are used when user must feel free to assess and to identify his emotions. The instruments are provided in the form of questionnaire, which may have various form either with a radio button that allows this emotion to be represented in binary or using rating scale such as Likert Scale [28]. Self-Assessment Manikin (SAM) [29] and Emocards [30] which are pictorials instruments are often used for measuring pleasure, arousal and dominance. Other self-report instruments focus on distinct emotional states like Product Emotion Measurement (PrEmo) [31] and Positive and Negative Affect Schedule (PANAS) [32].

2.1.2 Behavioral Signals

Recognizing the emotional state of human from his facial expression, vocal tone or posture is the most common way that these signals are one of the most natural and powerful channels in interpersonal communication. Research indicate that verbal signals (speech) convey one third of interpersonal communication, while nonverbal signals (facial expressions, body posture and gestures) convey two-thirds. Therefore, it is natural that interest in emotion recognition methods based on the analysis of facial expressions, postures, and gestures has increased significantly.

2.1.2.1 Facial Expressions

Facial expression recognition (FER) is one of the most common and natural emotion recognition methods based on visual sensors. There is a list of sensors used in various FER systems such as cameras, eye trackers, electrocardiograms (ECG), electromyograms (EMG), electroencephalographs (EEG), etc. Nevertheless, the camera is the most common sensor due to its ease of use and its low cost. The FER process comprises three steps: first face detection then facial expression detection and finally emotion classification (Figure 1).

Figure 1. Facial expression recognition process

Most Facial Expression Recognition (FER) systems aim to identify the six fundamental emotional expressions introduced by Ekman, which include fear, disgust, anger, surprise, happiness, and sadness [33]. Some FER systems go beyond these basic emotions and strive to detect more intricate emotional states. Human FER systems can be classified into two main categories: spontaneous recognition systems, which analyze natural and uncontrolled expressions, and pose-based recognition systems, which focus on capturing expressions based on specific facial poses. [34]. Spontaneous expression recognition systems recognize expressions that appear explicitly on people's faces on a daily basis, such as when preserving or watching movies. Conversely pose-based expression recognition systems detect artificial expression which people produced when they are asked to do. Other classification of the FER systems groups facial expressions into micro-expressions and macro-expressions. Facial micro-expressions are brief, involuntary facial expressions that occur spontaneously in response to an emotional stimulus [35]. They are very subtle and typically last for only a fraction of a second (less than 0.5 seconds). However, Macro-expression are more prolonged and visible expressions that last longer than half a second. These expressions are consciously and voluntarily controlled, allowing individuals to convey their emotional states more explicitly [35].

2.1.2.2 Vocal Expressions

Speech as an acoustic signal is fundamental and rich source of information about speakers. Emotional information in vocal expressions is conveyed through several factors, including tone of voice, pitch, volume, tempo, and rhythm. However, understanding the emotional cues in speech is a complex process. In contrast to facial expression which have been standardized through the Facial Action Coding System (FACS) developed by Ekman [35], vocal expressions of emotion are culture-specific. Different expressivity styles of different people from different cultures lead to acoustic variability that directly affect speech emotion recognition process. Speech emotion recognition systems are generally based on two models: Linguistic-based models and Acoustic-based models. The Linguistic-based models analyze the linguistic content and language characteristics of speech to infer emotions. They consider features like word choice, sentence structure, intonation patterns, and semantic context [36]. The acoustic-based models focus on extracting spectral features from the speech signal to recognize emotions [37]. Many works have combined language and acoustic models to improve the performance of speech emotion recognition [38] [39].

2.1.2.3 Body Posture and Gesture

While emotions can be expressed through different modalities, Body posture and gestures have indeed been extensively studied in the field of nonverbal communication for recognizing human emotions. These nonverbal cues provide valuable information about a person's emotional state and intentions. In addition to advancing our fundamental understanding of human behavior, research on body posture and language has played an important role in the study of animated conversational agents (ACA) [40]. Many studies explored the relation between emotions and body posture and gestures (Table 1) [41] [42]. In [43], authors defined a Body Action and Posture (BAP) coding system to study 12 emotional activities performed by professional actors using wearable accelerometers.

Table 1. - Emotional Expressive elements of posture and gesture [41]

Emotion	Body Posture and Gesture
Joy	Body extended, shoulders up, Head backward, arms raised up or away from the body
Fear	Body muscles tense up, arms are raised forwards, shoulders forwards
Boredom	head backwards and bent sideway, collapsed body posture, Raising the chin
Disgust	Shoulders forwards, head downwards, upper body collapsed, hands close to the body
Surprise	Head backward, chest backward, abdominal twist, Right/left hand going to the head (covering the cheeks/mouth)
Anger	Head backward, arms raised forwards and upwards, shoulders squared

2.2 Objective measures

In addition to behavioral signals, there are indeed various physiological signals that can convey emotions. These physiological signals are bodily responses that occur as a result of emotional experiences and can provide valuable insights into a person's emotional state. These signals are generated through the natural functioning of the human body's physiological systems. Physiological signals are measured using specialized instruments and techniques. Electrodes, sensors, or other monitoring devices are applied to the appropriate areas of the body to capture the signals accurately. The signals are then amplified, filtered, and recorded for further analysis. The most commonly used biosensors which can be used for collection physiological signals are: EEG (electroencephalogram), ECG (electrocardiogram), GSR (galvanic skin response), PVB (blood volume pulse) and EOG (electrooculography).

EEG measures the electrical activity generated by the firing of neurons in the brain. Electrodes placed on the scalp pick up the tiny electrical signals produced by the brain's neural activity. As many studies have proved that prefrontal cortex, temporal lobe, and anterior cingulate gyrus are involved in the regulation and control of emotions in the human brain, so EEG measures hold information that can be extracted to gain insights into the emotional state of a human being. In Intelligence Artificial literature, there are several works devoted to develop EEG-based emotion recognition systems which has broad application prospects. For the evaluation of human emotions, brain electrical signals have been classified into five different frequency bands, known as the brain waves [44]: delta (0-4 Hz), theta (4-8 Hz), alpha (812 Hz), beta (12-30 Hz) and gamma (>30 Hz). These frequency bands are commonly observed in EEG recordings and represent different states of brain activity.

ECG device measure the electrical activities of the heart in different phases and perspectives based on the situation and configuration. The ECG signals are recording using electrodes over the skin in a period of time. Though the primary purpose of ECG in medicine is to detect pathological heart conditions, ECG signals can be used to evaluate heart activities changes caused by emotional experiences [45][46]. ECG-based emotion recognition systems are widely used and research has shown their efficiency [47].

Galvanic skin response GSR is one of several electrodermal signals that can measure electrical changes of human skin which are not under his conscious control [48]. These signals are the result of changes in the electrical conductivity of the skin in response to various stimuli or emotional arousal. It is a physiological measure that reflects the activity of the sweat glands in the skin. Indeed, when sweat glands becomes active, they secrete more moisture towards the surface of the skin. For instance, when an individual experiences anxiety or concern, their sweat glands tend to produce a higher amount of sweat, leading to more significant fluctuations in electrical current.

Clinically, Electrooculography (EOG) refers to the measurement and recording of electrical signals generated by the movement of the eyes. It involves placing electrodes near the eyes to detect and analyze the changes in electrical potentials that occur as the eyes move. EOG is commonly used in medical and research settings to assess eye movements, monitor sleep patterns, diagnose certain eye conditions, and study various aspects of visual perception and cognitive processes. The idea of applying EOG to emotion detection is based on the same hypothesis as EMG which evaluate and record the electrical potential generated by muscle cells [49]. EOG is based on eye blink detection in most cases and is useful for detecting emotions such as sadness, stress and surprise. Many studies used EOG for assessing fatigue (driving or studying) [50][51] and concentration.

Related physiological signals also include RSP, Skin Temperature Measurements (SKT), etc.

3. Computational emotion recognition methods

In this section, we focus on the computational methods used in emotion recognition process, which can be classified into two categories as shown in figure 2.

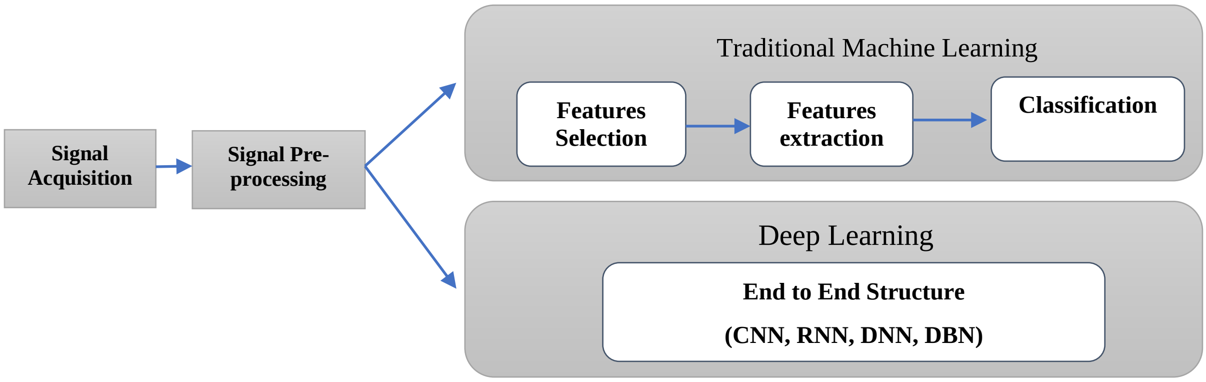

Figure 2. - Emotion recognition process

Signal pre-processing stage has traditionally been a preliminary necessary step for the emotion recognition process. Indeed, the majority of collected data from different sensors are highly susceptible to be noisy, inconsistent, and missing due to their heterogeneous origin, the complex and subjective nature of some signals, and sensitivity to noise from crosstalk and electromagnetic interference, etc. Signal pre-processing itself is divided into four steps: Data cleaning which refers to techniques to remove outliers and duplicate data, filling in missing values, and correcting inconsistent data [52]. Data integration step involves approaches that lead to merge the data collected from multiple sources into a single, larger, and consistent view of the data. Once data cleaning and integration have been done, data reduction step begins with the use of special techniques like principal component analysis to have a condensed representation of the data. Data transformation is the final step of data pre-processing that transform the data in appropriate forms suitable for data processing. This could include structuring unstructured data, combining key variables, and identifying key ranges.

A classical emotion recognition analysis with traditional machine learning methods usually employs three stages: Features extraction, Features optimization, and classification. The first stage i.e., feature extraction plays a crucial role in the emotion recognition process. This aims to summarize information contained in the original data and transformed into more manageable set of features. Therefore, feature extraction can reduce computational complexity, overcome the curse of dimensionality, and improve the generalization ability of models [53]. Due to their complexity and non-stationarity, the physiological signals often require feature extraction before being input into traditional classification models for emotion recognition. For example, speech signal features involve prosodic features, energy features, frequency spectral features and frequency spectral coefficients [54]. Among these features those that carry information and others that carry emotions. Hence, features extraction methods should be used to extract emotion features. In general, feature optimization step consists for minimizing the number of features to reduce computational complexity and improve the efficiency of machine learning models. There are several algorithms for feature selections which reduce the dimensionality of the features by rejecting redundant and irrelevant ones such as ReliefF ((Relief), Sequential Forward Selection (SFS), Sequential Backward Selection (BS) and Tree-Based Feature Selection (TS. The primary objective of emotion recognition systems revolves around the classification of input data to determine the expressed emotions. Classification involves training a machine learning model or employing other classification algorithms to learn patterns and relationships between the input features and corresponding emotional states. The trained model is then used to predict or classify the emotional state of new, unseen data. Classical machine learning offers a range of classification methods that are well-suited for the task of emotion classification. These methods include decision trees, logistic regression, support vector machines (SVM), k-nearest neighbors (KNN), naive Bayes classifiers, random forests, Gaussian Mixture Modelling (GMM), and Hidden Markov Model (HMM), among others. Traditional methods perform well for lab-pose datasets, but are less effective for truly complex and spontaneous scenes.

In deep learning methods, the model learns by itself discovering the relations between the features. Therefore, feature extraction, feature optimization and classification are combined into an overall step. Several deep learning methods such as CNN have adapted to be used in emotion recognition systems. In fact, CNN have been widely used to process physiological signals such as EEG, EMG and ECG. CNNs have proven to be effective in extracting meaningful features from these signals and have shown promising results in various applications [55][56]. In addition to convolutional neural networks (CNNs), other deep learning methods such as Long Short-Term Memory (LSTM) and

Deep Belief Networks (DBNs) have been utilized in emotion recognition tasks. LSTM is a type of recurrent neural network (RNN) that is well-suited for modeling sequential data, such as time series or speech signals, recognizing and classifying emotional states over time. [57]. Deep Belief Networks (DBNs) are a type of generative neural network that consists of multiple layers of hidden units. DBNs have been successfully applied in various tasks, including feature learning and representation. In the context of emotion recognition, DBNs can be used to automatically learn relevant features from the input data, allowing for more effective representation and classification of emotions. Moreover, Probabilistic neural network (PNN), in the context of EEG signal classification, have been explored for different tasks such as emotion recognition, mental state analysis [58].

4. Literature review

The following tables present recent works representative for each emotion recognition measurement method, which lists the relevant information of: number of participants were included, emotions measured and classification method, as well as the recognition rate. The main keywords used for the literature search are as follows: Emotion Recognition, Computational models of Emotion Recognition, Emotion Classification, Physiological Sensor of Emotion Recognition. Works with inaccurate information about equipment, features or calculation accuracy were excluded. Studies related to emotion recognition of participants with mental or organic disorders were also excluded from this review. We start our review with scientific researches focused on emotion recognition used single physiological signal (Table 2).

Table 2. - Review of previous researches focused on emotion recognition used single physiological signal

Ref	Signal	Year	Emotion	Classification Method	Recognition Accuracy
[59]	EEG	2020	-Negative, positive, and neutral. -Amusement, excitement, happiness, calmness, anger, disgust, fear, sadness, and surprise	CNN Dynamical graph	90.40%
[60]	EEG	2021	Valence, Arousal	Deep forest	Valence 97.69% Arousal 97.53 %
[61]	EEG	2021	happy, pleased, relaxed, excited, neutral, calm, distressed, miserable, and depressed.	QDC RNN	89.00%
[62]	EEG	2021	-Disgust, sadness, surprise and anger -Positive, negative, and neutral	LSTM	4 class: 94.12% 3 class: 92.66%
[63]	ECG	2022	happy, exciting, calm, and tense	SVM	90.51667%
[64]	ECG	2020	Happy, Sad, Pleasant and Angry	SVW, CART and KNN	Happy: 91% Sad: 90% Pleasant: 88% Angry: 97% The overall recognition rate is 92%
[65]	ECG	2023	surprise, sadness, anxiety, passion, joy, shame, hope, tired, fear, disgust, anger, gratitude, intimacy, trust, pain, confidence and relaxation	logistic regression (LR)	84.3%
[66]	GSR	2023	happiness, grief, fear, anger, and calm	SVM	66.67%
[67]	HRV	2019	High/low valence and arousal	CNN	Valence: 75.3% Arousal: 76.2%
[68]	EOG	2020	High arousal and low valence, low arousal and moderate valence, and high arousal and high valence	SVM	80%

Different from the above, many works attempted to combine and fuse signal features before feeding them into an emotion classifier (Table 3).

Table 3. - Review of previous researches focused on emotion recognition used a combination of physiological signals

Ref	Signal	Year	Emotion	Classification Method	Recognition Accuracy
[69]	RB, PPG, and fingertip temperature (FTT)	2020	Valence, Arousal	RF, SVM, LR	Arousal: 69.86 % - 73.08 % Valence: 69.53% - 72.18%
[70]	EOG, EMG	2021	happy, relaxing, angry and sad	SVM, Naïve Bayes, ANN	98%
[71]	EEG, ECG	2020	Positive, Negative	LSTM	EEG Signal: 76.67% ECG Signal: 75.00% EEG and ECG signals: 95.00%
[72]	EEG, GSR, PPG	2020	happy, relaxed, angry, and sad	KNN	79.76% for four emotions
[73]	EEG, EOG, EMG, GSR, RSP, BVP, and TMP	2020	valence, arousal	CNN	Valence: 93.06% Arousal: 91.95% (Combined signals)
[74]	GSR, ECG	2023	valence, arousal	CNN	InceptionResnetV2 CNN classifier: Valence: 91.27% Arousal91.45% MobileNet CNN classifier: Valence:99.19% Arousal: 98.39%
[75]	ECG, GSR, HR, GSR, SKT	2019	Anger, Happy, Sad, Joy	ANN	75.38 %

Emotion recognition techniques that analyze facial expressions, body postures, and gestures operate on the same assumption as physiological signals, which is that body postures and gestures are also implicated in emotional responses. Generally, the databases of facial expressions, postures, and gestures can come in various forms such as static images, videos, 3D models, or real-world recordings, depending on the purpose and techniques used to create them. Researchers in the field of facial expression recognition have access to a wide range of databases that can be utilized for their research. Most of these databases are annotated with the six basic emotions (anger, disgust, fear, happiness, sadness, and surprise) in addition to a neutral expression. Table 4 provide an overview of the most widely used FER databases.

Table 4. - An overview of facial expression recognition databases

Ref	Database	Samples	Types	Emotions
[76]	Affect Net	450,000 images	Posed and Spontaneous	(Anger, disgust, fear, happiness, sadness, surprise, and neutral
[77]	Emotio Net	1,000,000 images	Posed and Spontaneous	26 emotions: 6 basic emotion anger, disgust, fear, happiness, sadness, surprise) and 20 compound emotions (e.g., happy-surprise, sad-disgust, etc.).
[78]	MMI	2900 videos and 740 images	Posed	anger, disgust, fear, happiness, sadness, surprise, and neutral
[79]	FER-2013	35,887 images	Posed and Spontaneous	anger, disgust, fear, happiness, sadness, surprise, and neutral

[80]	CK+	593 images	Posed	anger, disgust, fear, happiness, sadness, surprise, neutral and contempt.
[81]	JAFEE	213 images	Posed	anger, disgust, fear, happiness, sadness, surprise, and neutral
[82]	RaFD	8040 images	Posed	anger, disgust, fear, happiness, sadness, surprise, neutral and contempt.
[83]	KDEF	4900 images	Posed	anger, disgust, fear, happiness, sadness, surprise, and neutral

Although facial expression databases are the most commonly used, body posture and gesture databases also are widely used to recognize human emotions. There are several databases available that contain data on body posture and gestures associated with specific emotions. These databases typically involve capturing data from human subjects using motion capture technology, which records the movements and positions of the body. Examples of such databases include:

- Chalearn Gesture Dataset: A large-scale dataset of RGB-D videos of people performing 249 gestures in front of a Kinect camera.
- NTU RGB+D Dataset: A large-scale dataset of RGB-D videos of 60 action classes, including various human actions and interactions.
- UT-Kinect Gesture Dataset: A dataset of 10 different hand gestures performed by 10 different subjects in front of a Kinect camera. These include gestures such as swipe left, swipe right, wave, circle, and others.
- MPII Human Pose Dataset: A dataset of human pose estimation with more than 25,000 images of people in various poses. This dataset has been widely used for human pose estimation, and various models and algorithms have been developed and evaluated using this dataset [83][84].

A summary of researches focused on emotions recognition using facial expressions, body posture and gestures with classification methods and recognition accuracy is provided in Table 5.

Table 5. - Review of previous studies focused on emotions recognition using facial expressions, body posture and gestures

Ref	Year	Methods	Emotion	Methods	Recognition Accuracy
[85]	2020	Facial expressions	happy, sad, anger, surprise and neutral	• Haar filter for extracting facial features • CNN for emotion recognition and classifying	• 65% for (FER)-2013 • 60% for personalized datasets
[86]	2020	Facial expression	• For JAFEE Dataset: Happy, Neutral, Fear, Sad, Disgust, Surprise and Anger • For CK+ Dataset: Contempt, Anger, Disgust, Fear, Sadness, Surprise and Happiness • For YALE FACE Dataset: Happy, Center-light, Left-light, W/no glasses, W/glasses, Sad, Right-light, Normal, Sleep, Surprised and Wink	• LBP for features extraction and SVM for classification • CNN	• LBP: ✓ CK+: 96.66% ✓ JAFEE: 76.23% ✓ YALEFACE:74% • CNN: ✓ CK+: 97.32% ✓ JAFEE:77.27% ✓ YALFACE:31.82 %
[87]	2022	Facial expression	• For CK+ dataset: anger, contempt, fear, disgust, happiness, surprise, and	• MLP • SVM • KNN	• MLP: ✓ JAFFE:90% ✓ CK+:94%

			sadness • For JAFEE Dataset: happiness, sadness, surprise, anger, disgust, and fear and neutral • For RAF dataset: anger, happiness, fear, surprise, disgust, sadness and neutral	• LR	✓ RAF:67% • SVM: ✓ JAFFE:88% ✓ CK+:94% ✓ RAF:67% • KNN: ✓ JAFFE:95% ✓ CK+:97% ✓ RAF:63% • LR: ✓ JAFFE:86% ✓ CK+:87% ✓ RAF:66%
[88]	2021	Facial expression	• For CK+ dataset: anger, contempt, fear, disgust, happiness, surprise, and sadness • For JAFEE Dataset: happiness, sadness, surprise, anger, disgust, and fear and neutral • FER2013: angry, disgust, fear, happy, sad, surprise, including neutral	• A feedforward learning model	• JAFFE:96.8% • CK:86.5% • FER2013:62.5%
[89]	2021	Body Movements	• Happiness, sadness, fear, anger, and neutral.	• Two-layer feature-selection process	• During walking: 90% • During setting: 96% • Action independent-scenario: 86.66%
[90]	2019	Skeletal Movement	• neutral state, sadness, surprise, fear, anger, disgust and happiness	• CNN • RNN • RNN-LSTM	Case of 4 emotions • CNN: 63.6% • RNN:80.8% • RNN-LSM:82.7% Case of 6 emotions: • CNN: 54.2% • RNN:59.2% • RNN-LSM:72%
[91]	2019	Fusion of Facial expression and Body Gesture	• disgust, anger, fear, neutral, surprise, sad, and happy	• Multi SVM	• Anger: 95% • Fear:91% • Happy:95% • Disgust:96% • Neutral:95% • Sad:94% • Surprise:93%
[92]	2021	Upper Body Movements and Facial Expressions	disgust, fear, happiness, sadness, surprise, and neutral	CNN	94.41%

5. Discussion

This paper extensively examined numerous recent publications focusing on emotion recognition and sentiment analysis within the field. Emotion recognition systems are indeed in high demand and have numerous applications in the fields of artificial intelligence (AI) and the Internet of Things (IoT). There are many reasons why emotions recognition systems are in demand and here we can cite a few examples:

- ✓ Human- Computer Interaction: Emotion recognition systems enhance human-computer interaction by enabling computers and machines to understand and respond to human emotions.
- ✓ Personalized Experiences: Emotion recognition can help tailor experiences based on individual emotions. For instance, it can be used in entertainment and gaming industries to adjust the content or difficulty level based on the user's emotional state, providing a more personalized and engaging experience.
- ✓ Healthcare and Well-being: Emotion recognition systems have potential applications in healthcare and mental well-being. They can be used for monitoring and diagnosing conditions like depression, anxiety, and stress, enabling timely interventions and personalized treatments.
- ✓ Market Research and Advertising: Emotion recognition can be utilized in market research and advertising to measure consumer emotional responses to products, advertisements, or user interfaces. This information can help companies to optimize their marketing strategies and improve user satisfaction.
- ✓ Security and Surveillance: Emotion recognition systems can enhance security and surveillance systems by identifying suspicious or threatening behaviors based on facial expressions or voice analysis. They can be employed in public spaces, airports, or high-security areas to improve safety measures.
- ✓ Education and Learning: Emotion recognition technology can assist in educational settings by gauging student engagement and emotional states during learning activities. This information can help educators customize teaching methods, provide additional support, and create a more effective learning environment.

Due to the high demand on emotion recognition systems, researchers have been actively working to improve the accuracy of these systems by focusing on several key areas such as multimodal approaches. Emotion recognition systems that incorporate multiple modalities, such as facial expressions, speech analysis, physiological signals (e.g., heart rate variability), and body language, tend to be more robust and accurate.

The statement that Physiological signals, being beyond conscious control, are considered highly dependable for recognizing emotions and are regarded as the most reliable signals in this context is partially true. Physiological signals, such as heart rate, skin conductance, and electroencephalography (EEG), can provide valuable insights into a person's emotional state. These signals are largely involuntary and can reflect underlying physiological processes associated with emotions. Unlike facial expressions or verbal cues, which can be consciously manipulated or masked, physiological signals are less prone to deliberate control. However, it is important to note that physiological signals are not entirely immune to influence. While individuals may not be able to directly control their physiological responses, various factors can still influence these signals. For example, physiological arousal can be affected by factors such as stress, physical exertion, medications, or certain health conditions. Additionally, individual differences in baseline physiological responses can also impact the interpretation of these signals.

Basically, it can be argued that psychological methods for recognizing human emotions are generally simpler compared to physiological methods. Psychological methods involve observing and interpreting observable behavioral cues and subjective self-report measures to infer emotional states. These methods rely on the analysis of facial expressions, body language, vocal tone, and verbal content to identify and categorize emotions. Psychological methods are relatively straightforward to implement as they do not require specialized equipment or complex data processing. They often involve trained observers or individuals themselves reporting their emotional experiences using standardized questionnaires or rating scales. These methods can provide valuable insights into emotions and are widely used in various fields. However, it is important to note that psychological methods have their limitations. They rely on external cues and self-report, which can be influenced by factors such as social desirability, individual differences, and cultural variations. Additionally, accurately interpreting behavioral cues and self-report measures requires expertise and training to minimize bias and ensure reliable results. On the other hand, physiological methods, as mentioned earlier, involve measuring physiological signals associated with emotions, such as heart rate, skin conductance, or brain activity. While more complex in terms of equipment and data analysis, physiological methods provide an additional layer of objective information and can capture emotional responses that may not be apparent through behavioral cues alone. Therefore, the choice between the two methods depends on the specific goals, context, and resources available for emotion recognition.

Recent studies conducted in the field of human emotion recognition indicate that there is no single method that can be considered ideal for emotion recognition in all situations, as explained in [93] and [94] combing physiological methods with psychological methods can provide a more comprehensive understanding of emotions. Furthermore, combining multiple methods can also provide opportunities for cross-validation and verification. While the availability and accessibility of services utilizing multimodal emotion recognition may vary, research and development in this field continue to advance. These findings align closely with the essence of fusion techniques in emotion recognition, which aim to integrate information from various sources or modalities ranging from facial expressions, voice, text, to physiological signals. The integration of these diverse sources serves to compensate for the limitations inherent in individual modalities, reflecting the idea that combining physiological and psychological methods yields a more comprehensive understanding of emotions. Moreover, the crux of fusion techniques lies in their objective to extract meaningful features from multiple data sources and effectively merge them to enhance the accuracy of emotion prediction.

Deciding how and when to combine information from different modalities in multimodal emotion recognition is a pivotal challenge. Various fusion techniques like early (feature-level) fusion, late (decision-level) fusion, and the more recent bilinear pooling fusion offer distinct approaches with their own trade-offs. Early (feature-level) fusion involves the combination of features from diverse modalities at the initial input level before feeding them into the model. This method offers the advantage of retaining the raw information of all modalities from the outset, using only one learning stage and facilitating a straightforward integration process [95]. However, early fusion encounters challenges in managing varying temporal resolutions among different modalities and may escalate the input data's dimensionality, potentially leading to computational complexities. In contrast, late (decision-level) fusion extracts distinct features from each modality independently and merges them at a higher processing stage, typically during decision-making phases (classification). This approach permits the individual processing of each modality's specific features and exhibits more flexibility in addressing temporal misalignments. The primary aim of this method is to leverage the redundancy present in a collection of independent classifiers, merging their results to enhance robustness and achieve higher accuracy in the classification process [96]. Nonetheless, late fusion risks the potential loss of certain contextual nuances or intermodal interactions and necessitates meticulous alignment of features derived from different modalities to be effective. Bilinear pooling fusion is an advanced feature-based fusion technique used in multimodal emotion recognition that focuses on combining information from different modalities at the feature level. Unlike early fusion that merges features directly, bilinear pooling computes a bilinear interaction between features from different modalities. This interaction captures complex cross-modal relationships and dependencies, allowing the model to learn sophisticated correlations between modalities. However, bilinear pooling requires careful design of the interaction function, and its computational demands might be higher compared to simpler fusion methods. Despite these challenges, this approach has gained attention for its ability to capture intricate intermodal interactions, potentially improving the model's understanding of multimodal data and enhancing overall recognition accuracy [97].

Such techniques, whether executed through early fusion, combining raw data at the input level, or late fusion, merging extracted features at higher processing stages, strive to encapsulate the intricacies of human emotional expression across different channels. By doing so, they aim to create a more holistic and accurate portrayal of emotions, aligning with the quest for a comprehensive understanding of an individual's emotional state.

Ultimately, the juxtaposition of psychological and physiological methods alongside fusion techniques signifies a pivotal juncture in the evolution of emotion recognition. This discourse underscores the critical need for comprehensive approaches that combine the strengths of diverse methodologies to capture the intricacies of human emotions. As research continues to explore innovative fusion techniques and integrate multimodal approaches, the pursuit of a nuanced and holistic understanding of human emotions remains at the forefront, fostering groundbreaking developments in the field of emotion recognition.

In conjunction with these fusion methodologies, the extraction of pertinent features from various data modalities significantly contributes to the effectiveness of the overall emotion recognition framework. Indeed, the success of the fusion techniques is intricately tied to the quality and relevance of the extracted features. The accuracy and informativeness of these features ensured a more robust and comprehensive representation of emotions, thereby augmenting the efficacy of the fusion process. For instance, in facial emotion recognition, features like facial landmarks, textures, or certain statistical measures from images or videos might be extracted. Extracting facial landmarks or key points (e.g., eye corners, mouth edges) using techniques such as the Constrained Local Model (CLM) [98] [99] or Active Appearance Models (AAM) [100] helps capture subtle facial muscle movements indicating emotions. However, in recent years, convolutional neural networks (CNNs) have been widely employed for end-to-end feature extraction from raw images, enabling the network to learn hierarchical representations of facial features [101][102]. Additionally, Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are commonly applied in sequential data analysis due to their ability to retain memory over sequences, making them particularly useful for tasks involving time-series data, natural language processing, and sequential pattern recognition. RNNs and LSTMs capture temporal dependencies in speech or textual data, extracting sequential features crucial for recognizing nuanced emotions [103][104]. Feature extraction from physiological signals like heart rate variability (HRV), skin conductance, and electroencephalogram (EEG) involves identifying specific signal characteristics linked to different emotional states. For instance, Spectral analysis techniques, such as calculating power in specific frequency bands (e.g., LF for low frequency, HF for high frequency) or employing ratios like the LF/HF ratio, serve to capture sympathetic and parasympathetic activities. These metrics are often associated with distinct emotional states, providing insights into the physiological correlates of emotions [105].

Innovative strides in feature extraction techniques across multiple modalities, from facial analysis to physiological signal processing and sequential data analysis, reflect the dynamic landscape of emotion recognition. These cuttingedge methodologies not only enhance the accuracy and depth of emotion understanding but also pave the way for more nuanced, context-aware systems. As we harness the potential of these diverse extraction techniques, we inch closer to developing emotionally intelligent systems that can perceive, comprehend, and respond to human emotions with unprecedented precision. The ongoing exploration and integration of these techniques are poised to revolutionize not just technology but also human-computer interactions, ushering in an era where machines understand emotions in ways that mirror human comprehension.

6. Conclusion

This article has explored the methods and sensors utilized in emotion recognition, with a focus on computational approaches and different modalities. The discussion highlighted the use of traditional machine learning and deep learning techniques for analyzing and classifying emotions, showcasing the advancements made in the field. Additionally, the various modalities of emotion recognition, including facial expressions, vocal cues, physiological signals, textual data, and brain activity, were examined, considering their respective strengths and limitations.

However, it is important to note that despite the progress achieved, there is no perfect method for emotion recognition. Each approach or modality has its own set of advantages and challenges. Factors such as individual differences, cultural variations, and contextual influences pose ongoing difficulties in accurately and reliably recognizing emotions.

The article underscores the need for continued research and development in emotion recognition to address these challenges. Future advancements may involve refining existing computational methods, exploring novel sensor technologies, and considering multimodal approaches that combine multiple modalities for a more comprehensive understanding of emotions.

Ultimately, a holistic and robust emotion recognition system is yet to be realized. As technology advances and interdisciplinary collaborations flourish, we can look forward to further breakthroughs in emotion recognition, paving the way.

Funding

Algerian Ministry of Higher Education and Scientific Research, Grant/Award Number: C00L07UN310220180008

ACKNOWLEDGEMENT

None

CONFLICTS OF INTEREST

The author declares no conflict of interest

REFERENCES

[1] S. Menon, Brain, self and consciousness: Explaining the conspiracy of experience. New Delhi, India: Springer, 2016.

[2] N. M. Szajnberg, “What the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system (FACS) what the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system (FACS). Edited by Erika L. rosenberg and Paul Ekman. 3rd ed. New York: Oxford university press, 2020, xxii + 627 pp., 80.00 hardcover, 29.99 paperback,” J. Am. Psychoanal. Assoc., vol. 70, no. 3, pp. 591–595, 2022.

[3] Elham S. Salama, Reda A. El-Khoribi, Mahmoud E. Shoman, Mohamed A. Wahby Shalaby, “A 3D-convolutional neural network framework with ensemble learning techniques for multi-modal emotion recognition,” Egyptian Informatics Journal, vol. 2, no. 2, pp. 167–176, 2021.

[4] M.-Z. Poh, N. C. Swenson, and R. W. Picard, “A wearable sensor for unobtrusive, long-term assessment of electrodermal activity,” IEEE Trans. Biomed. Eng., vol. 57, no. 5, pp. 1243–1252, 2010.

[5] X. Liu, Z. Xu, and K. Huang, “Multimodal emotion recognition based on cascaded multichannel and hierarchical fusion,” Comput. Intell. Neurosci., vol. 2023, p. 9645611, 2023.

[6] G. Valenza, L. Citi, A. Lanatà, E. P. Scilingo, and R. Barbieri, “A nonlinear heartbeat dynamics model approach for personalized emotion recognition,” Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., vol. 2013, pp. 2579–2582, 2013.

[7] J. Machajdik and A. Hanbury, “Affective image classification using features inspired by psychology and art theory,” in Proceedings of the 18th ACM international conference on Multimedia, 2010.

[8] H. Wang, X. Li, Z. Ren, M. Wang, and C. Ma, “Multimodal sentiment analysis representations learning via contrastive learning with condense attention fusion,” Sensors (Basel), vol. 23, no. 5, 2023.

[9] R. W. Picard, Affective Computing. London, England: MIT Press, 2000.

[10] D. Liu, Z. Wang, L. Wang, and L. Chen, “Multi-modal fusion emotion recognition method of speech expression based on deep learning,” Front. Neurorobot., vol. 15, p. 697634, 2021.

[11] R. S. Gaikwad and S. C. Gandage, "MCNN: Visual Sentiment Analysis using Various Deep Learning Framework with Deep CNN," InternationalJournalof INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING, vol. 11, no. 2, pp. 265–278, 2023.

[12] K. Feng and T. Chaspari, "A review of generalizable transfer learning in automatic emotion recognition," Front. Comput. Sci., vol. 2, 2020.

[13] S. Liu, P. Gao, Y. Li, W. Fu, and W. Ding, "Multi-modal fusion network with complementarity and importance for emotion recognition," Inf. Sci. (Ny), vol. 619, pp. 679–694, 2023.

[14] K. Scherer, "Vocal communication of emotion: A review of research paradigms," Speech Commun., vol. 40, no. 1–2, pp. 227–256, 2003.

[15] R. A. Calvo and S. D'Mello, "Affect detection: An interdisciplinary review of models, methods, and their applications," IEEE Trans. Affect. Comput., vol. 1, no. 1, pp. 18–37, 2010.

[16] J. M. Salsman et al., "Emotion assessment using the NIH Toolbox," Neurology, vol. 80, no. 11 Suppl 3, pp. S76-86, 2013.

[17] T.-M. Bynion and M. T. Feldner, "Self-Assessment Manikin," in Encyclopedia of Personality and Individual Differences, Cham: Springer International Publishing, 2020, pp. 4654–4656.

[18] S. Katsigiannis and N. Ramzan, "DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices," IEEE J. Biomed. Health Inform., vol. 22, no. 1, pp. 98–107, 2018.

[19] S. Jerritta, M. Murugappan, R. Nagarajan, and K. Wan, "Physiological signals based human emotion Recognition: a review," in 2011 IEEE 7th International Colloquium on Signal Processing and its Applications, 2011.

[20] C. Qing, R. Qiao, X. Xu, and Y. Cheng, "Interpretable emotion recognition using EEG signals," IEEE Access, vol. 7, pp. 94160–94170, 2019.

[21] A. Kushki, J. Fairley, S. Merja, G. King, and T. Chau, "Comparison of blood volume pulse and skin conductance responses to mental and affective stimuli at different anatomical sites," Physiol. Meas., vol. 32, no. 10, pp. 1529–1539, 2011.

[22] J. Z. Lim, J. Mountstephens, and J. Teo, "Emotion recognition using eye-tracking: Taxonomy, review and current challenges," Sensors (Basel), vol. 20, no. 8, p. 2384, 2020.

[23] W.-L. Zheng, B.-N. Dong, and B.-L. Lu, "Multimodal emotion recognition using EEG and eye tracking data," Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., vol. 2014, pp. 5040–5043, 2014.

[24] F. Z. Canal et al., "A survey on facial emotion recognition techniques: A state-of-the-art literature review," Inf. Sci. (Ny), vol. 582, pp. 593–617, 2022.

[25] A. A. Abdelhamid et al., "Robust speech emotion recognition using CNN+LSTM based on stochastic fractal search optimization algorithm," IEEE Access, vol. 10, pp. 49265–49284, 2022.

[26] S. Vaijayanthi, J. Arunnehru, M. Singh, V. Tyagi, P. K. Gupta, and J. Flusser, "Human Emotion Recognition from Body Posture with Machine Learning Techniques," in Advances in Computing and Data Sciences. ICACDS 2022, T. Ören, Ed.

[27] W. Sasaki, J. Nakazawa, and T. Okoshi, "Comparing ESM Timings for Emotional Estimation Model with FineTemporal Granularity," in 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers, 2018.

[28] M. M. Bradley and P. J. Lang, "Measuring emotion: the Self-Assessment Manikin and the Semantic Differential," J. Behav. Ther. Exp. Psychiatry, vol. 25, no. 1, pp. 49–59, 1994.

[29] P. Desmet, K. Overbeeke, and S. Tax, "Designing products with added emotional value: Development and appllcation of an approach for research through design," Des. J., vol. 4, no. 1, pp. 32–47, 2001.

[30] P. Desmet, Designing Emotions. The Netherlands, 2002.

[31] "Development and validation of brief measures of positive and negative affect: the PANAS scales," J. Pers. Soc. Psychol, vol. 54, no. 6, pp. 1063–1070, 1988.

[32] P. Ekman, Emotion in the human face: Guide-lines for research and an integration of findings. Pergamon, 1972.

[33] Q. Gan, C. Wu, S. Wang, and Q. Ji, "Posed and spontaneous facial expression differentiation using deep Boltzmann machines," in 2015 International Conference on Affective Computing and Intelligent Interaction (ACII), 2015.

[34] J. A. Harrigan, R. Rosenthal, and K. R. Scherer, The new handbook of methods in nonverbal behavior research. London, England: Oxford University Press, 2005.

[35] R. Rajoo and C. C. Aun, "Influences of languages in speech emotion recognition: A comparative study using Malay, English and Mandarin languages," in 2016 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE), 2016.

[36] S. Bhosale, R. Chakraborty, and S. K. Kopparapu, "Deep encoded linguistic and acoustic cues for attention-based end to end speech emotion recognition," in ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020.

[37] M. Sakurai and T. Kosaka, "Emotion recognition combining acoustic and linguistic features based on speech recognition results," in 2021 IEEE 10th Global Conference on Consumer Electronics (GCCE), 2021.

[38] Salman , H. A., Kalakech , A., & Steiti, A. (2024). Random Forest Algorithm Overview. Babylonian Journal of Machine Learning, 2024, 69–79. https://doi.org/10.58496/BJML/2024/007
[39] G. Ball and J, Breese, Emotion and personality in a conversational agent. Cambridge: USA: Embodied conversational agents, 2000.
[40] S. Lee, M. Bae, W. Lee, and H. Kim, "CEPP: Perceiving the emotional state of the user based on body posture," Appl. Sci. (Basel), vol. 7, no. 10, p. 978, 2017.
[41] R. Calvo, S. D'Mello, J. Gratch, A. Kappas, M. Lhommet, and S. C. Marsella, "Expressing emotion through posture and gesture," in The Oxford Handbook of Affective Computing, Oxford University Press, 2015.
[42] N. Dael, M. Mortillaro, and K. R. Scherer, "Emotion expression in body action and posture," Emotion, vol. 12, no. 5, pp. 1085–1101, 2012.
[43] B. Kaur, D. Singh, and P. P. Roy, "EEG based emotion classification mechanism in BCI," Procedia Comput. Sci., vol. 132, pp. 752–758, 2018.
[44] P. Ekman, R. W. Levenson, and W. V. Friesen, "Real-Time Recognition of the Affective User State with Physiological Signals," in Proc. Doctoral Consortium Conf. Affective Computing and Intelligent Interaction, 2007, pp. 1–8.
[45] J. Anttonen and V. Surakka, "Emotions and heart rate while sitting on a chair," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2005.
[46] K. Rattanyu, Graduate School of Functional Control Systems Engineering, Shibaura Institute of Technology, 3-7-5 Toyosu, Koto-ku, Tokyo 135-8548, Japan, M. Mizukawa, and Department of Electrical Engineering, Shibaura Institute of Technology, 3-7-5 Toyosu, Koto-ku, Tokyo 135-8548, Japan, "Emotion recognition based on ECG signals for service robots in the intelligent space during daily life," J. Adv. Comput. Intell. Intell. Inform., vol. 15, no. 5, pp. 582–591, 2011.
[47] G. Udovičić, J. Đerek, M. Russo, and M. Sikora, "Wearable emotion recognition system based on GSR and PPG signals," in Proceedings of the 2nd International Workshop on Multimedia for Personal Health and Health Care, 2017.
[48] C. Zong and M. Chetouani, "Hilbert-Huang transform based physiological signals analysis for emotion recognition," in 2009 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 2009.
[49] B. Tag, A. Vargo, A. Gupta, G. Chernyshov, K. Kunze, and T. Dingler, "Continuous Alertness Assessments: Using EOG Glasses to Unobtrusively Monitor Fatigue Levels In-The-Wild," in CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 2019, pp. 1–12.
[50] M. Kolodziej et al., "Fatigue detection caused by office work with the use of EOG signal," IEEE Sens. J., vol. 20, no. 24, pp. 15213–15223, 2020.
[51] P. Mishra, A. Biancolillo, J. M. Roger, F. Marini, and D. N. Rutledge, "New data preprocessing trends based on ensemble of multiple preprocessing techniques," Trends Analyt. Chem., vol. 132, no. 116045, p. 116045, 2020.
[52] Y. Cai, X. Li, and J. Li, "Emotion recognition using different sensors, emotion models, methods and datasets: A comprehensive review," Sensors (Basel), vol. 23, no. 5, 2023.
[53] M. Swain, A. Routray, and P. Kabisatpathy, "Databases, features and classifiers for speech emotion recognition: a review," Int. J. Speech Technol., vol. 21, no. 1, pp. 93–120, 2018.
[54] R. Qiao, C. Qing, T. Zhang, X. Xing, and X. Xu, "A novel deep-learning based framework for multi-subjectemotion recognition," in Proceedings of the International Conference on Information, Cybernetics and Computational Social Systems, Dalian, 2017, pp. 181–185.
[55] J. Huang, X. Xu, and T. Zhang, "Emotion classification using deep neural networks and emotional patches," in 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2017.
[56] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
[57] J. Zhang, M. Chen, S. Hu, Y. Cao, and R. Kozma, "PNN for EEG-based emotion recognition," in 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2016.
[58] T. Song, W. Zheng, P. Song, and Z. Cui, "EEG emotion recognition using dynamical graph convolutional neural networks," IEEE Trans. Affect. Comput., vol. 11, no. 3, pp. 532–541, 2020.
[59] J. Cheng et al., "Emotion recognition from multi-channel EEG via deep forest," IEEE J. Biomed. Health Inform., vol. 25, no. 2, pp. 453–464, 2021.
[60] K. K. Talluri, M.-A. Fiedler, and A. Al-Hamadi, "Deep 3D convolutional neural network for facial micro-expression analysis from video images," Appl. Sci. (Basel), vol. 12, no. 21, p. 11078, 2022.
[61] S. Gannouni, A. Aledaily, K. Belwafi, and H. Aboalsamh, "Emotion detection using electroencephalography signals and a zero-time windowing-based epoch estimation and relevant electrode identification," Sci. Rep., vol. 11, no. 1, 2021
[62] A. Sakalle, P. Tomar, H. Bhardwaj, D. Acharya, and A. Bhardwaj, "A LSTM based deep learning network for recognizing emotions using wireless brainwave driven system," Expert Syst. Appl., vol. 173, no. 114516, p. 114516, 2021.

[63] B. Sun and Z. Lin, “Emotion Recognition using Machine Learning and ECG signals,” arXiv [eess.SP], 2022. Doi: https://doi.org/10.48550/arXiv.2203.08477

[64] Z. Zhang, X. Wang, P. Li, X. Chen, and L. Shao, “Research on emotion recognition based on ECG signal,” J. Phys. Conf. Ser., vol. 1678, no. 1, p. 012091, 2020. Doi: :10.1088/1742-6596/1678/1/012091

[65] L. Wang, J. Hao, and T. H. Zhou, “ECG Multi-Emotion Recognition Based on Heart Rate Variability Signal Features Mining,” Sensors, vol. 23, no. 20, p. 8636, Oct. 2023, doi: 10.3390/s23208636.

[66] D. Fan, M. Liu, X. Zhang, and X. Gong, “Human emotion recognition based on Galvanic Skin Response signal feature selection and SVM,” arXiv [eess.SP], 2023. Doi: https://doi.org/10.48550/arXiv.2307.05383

[67] M. S. Lee, Y. K. Lee, D. S. Pae, M. T. Lim, D. W. Kim, and T. K. Kang, “Fast emotion recognition based on single pulse PPG signal with convolutional neural network,” Appl. Sci. (Basel), vol. 9, no. 16, p. 3355, 2019.

[68] P. Tarnowski, M. Kołodziej, A. Majkowski, and R. J. Rak, “Eye-tracking analysis for emotion recognition,” Comput. Intell. Neurosci., vol. 2020, pp. 1–13, 2020

[69] D. Ayata, Y. Yaslan, and M. E. Kamasak, “Emotion recognition from multimodal physiological signals for emotion aware healthcare systems,” J. Med. Biol. Eng., vol. 40, no. 2, pp. 149–157, 2020.

[70] M. R. Kose, M. K. Ahirwal, and A. Kumar, “A new approach for emotions recognition through EOG and EMG signals,” Signal Image Video Process., vol. 15, no. 8, pp. 1863–1871, 2021, doi: https://doi.org/10.1007/s11760-021-01942-1.

[71] F. Zeng, L. I. N. Yitao, P. Siriaraya, D. Choi, and N. Kuwahara, “Emotion Detection Using EEG and ECG Signals from Wearable Textile Devices for Elderly People,” Journal of Textile Engineering, vol. 66, no. 6, pp. 109–117, doi: https://doi.org/10.4188/jte.66.109

[72] A. Raheel, M. Majid, M. Alnowami, and S. M. Anwar, “Physiological sensors based emotion recognition while experiencing tactile enhanced multimedia,” Sensors (Basel), vol. 20, no. 14, p. 4037, 2020, doi: 10.3390/s20144037

[73] J. Liao, Q. Zhong, Y. Zhu, and D. Cai, “Multimodal physiological signal emotion recognition based on convolutional recurrent neural network,” IOP Conf. Ser. Mater. Sci. Eng., vol. 782, no. 3, p. 032005, 2020, doi: 10.1088/1757-899X/782/3/032005.

[74] A. Dessai and H. Virani, “Emotion Classification Based on CWT of ECG and GSR Signals Using Various CNN Models,” Electronics, vol. 12, no. 13, p. 2795, Jun. 2023, doi: 10.3390/electronics12132795.

[75] S. Tiwari, S. Agarwal, M. Syafrullah, and K. Adiyarta, “Classification of physiological signals for emotion recognition using IoT,” in 2019 6th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), 2019.

[76] M. Mahoor, “AffectNet.”. [Online] Available: http://mohammadmahoor.com/affectnet/.

[77] Ohio-state.edu, “EmotioNet Database.”. [Online] Available: http://cbcsl.ece.ohio state.edu/dbform_emotionet.html.

[78] M. Eu, “MMI facial expression database.”. [Online] Available: https://mmifacedb.eu. [Accessed: 08-Jun-2023].

[79] “Facial Expression Recognition 2013 Database.”. [Online] Available: https://www.kaggle.com/c/challenges-inrepresentation-learning-facial-expressionrecognition-challenge/data.

[80] “The Japanese Female Facial Expression Database.”. [Online] Available: http://www.consortium.ri.cmu.edu/ckagree/.

[81] “The Extended Cohn–Kanade.”. Available: http://www.kasrl.org/jaffe.html

[82] O. Langner, R. Dotsch, G. Bijlstra, D. H. J. Wigboldus, S. T. Hawk, and A. Van Knippenberg, “Presentation and validation of the Radboud faces database, ” Cogn,” Cogn. Emot, vol. 24, no. 8, pp. 1377–1388, 2010.

[83] I. Petrov, V. Shakhuro, and A. Konushin, “Deep probabilistic human pose estimation,” IET Comput. Vis., vol. 12, no. 5, pp. 578–585, 2018.

[84] K. Sun, B. Xiao, D. Liu, and J. Wang, “Deep high-resolution representation learning for human pose estimation,” in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019.

[85] Kaviya and Arumugaprakash, “Group facial emotion analysis system using convolutional neural network,” in 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184), 2020.

[86] R. Ravi, S. V. Yadhukrishna, and R. Prithviraj, “A Face Expression Recognition Using CNN & LBP,” in 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 2020.

[87] A. I. Siam, N. F. Soliman, A. D. Algarni, F. E. Abd El-Samie, and A. Sedik, “Deploying Machine Learning techniques for human emotion detection,” Comput. Intell. Neurosci., vol. 2022, p. 8032673, 2022.

[88] B. Yusra, J. Afshan, N. Nudrat, M. Haroon, Y. Serestina, and V. Sergio, “Facial Expression Recognition of Instructor Using Deep Features and Extreme Learning Machine,” Computational Intelligence and Neuroscience, vol. 2021, pp. 1–17, 2021.

[89] F. Ahmed, A. S. M. H. Bari, and M. L. Gavrilova, “Emotion recognition from body movement,” IEEE Access, vol. 8, pp. 11761–11781, 2020.

[90] T. Sapiński, D. Kamińska, A. Pelikant, and G. Anbarjafari, "Emotion recognition from skeletal movements," Entropy (Basel), vol. 21, no. 7, p. 646, 2019.
[91] T. Keshari and S. Palaniswamy, "Emotion recognition using feature-level fusion of facial expressions and body gestures," in 2019 International Conference on Communication and Electronics Systems (ICCES), 2019.
[92] C. Ilyas, R. Nunes, K. Nasrollahi, M. Rehm, and T. Moeslund, "Deep emotion recognition through upper body movements and facial expression," in Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, 2021.
[93] K. Takahashi, "Remarks on emotion recognition from multi-modal bio-potential signals," in 2004 IEEE International Conference on Industrial Technology, 2004. IEEE ICIT '04, 2005.
[94] Lin, Lin, Wang, and Wu, "Multiple Convolutional Neural Networks fusion using Improved Fuzzy Integral for facial emotion recognition," Appl. Sci. (Basel), vol. 9, no. 13, p. 2593, 2019.
[95] C. G. M. Snoek, M. Worring, and A. W. M. Smeulders, "Early versus late fusion in semantic video analysis," in Proceedings of the 13th annual ACM international conference on Multimedia, 2005
[96] S. Planet and I. Sanz, "Comparison between decision-level and feature-level fusion of acoustic and linguistic features for spontaneous emotion recognition," in Iberian Conference on Information Systems and Technologies, CISTI, 2012, pp. 1–6.
[97] D. Nguyen, K. Nguyen, S. Sridharan, D. Dean, and C. Fookes, "Deep spatio-temporal feature fusion with compact bilinear pooling for multimodal emotion recognition," Comput. Vis. Image Underst., vol. 174, pp. 33–42, 2018.
[98] L. Wen, S. Yang, J. Zeng, X. Liang, and Y. Xu, "The analysis on student' psychologic status of online learning under extraction model from computer face features," J. Phys. Conf. Ser., vol. 1544, no. 1, p. 012198, 2020.
[99] A. Alsarayreh and F. S. Mohamad, "Enhanced constrained local models (CLM) for facial feature detection," International Journal of Engineering Research and Technology, vol. 13, no. 11, p. 3217, 2020.
[100] K. Prabhu, S. SathishKumar, M. Sivachitra, S. Dineshkumar, and P. Sathiyabama, "Facial expression recognition using enhanced convolution neural network with attention mechanism," Comput. Syst. Sci. Eng., vol. 41, no. 1, pp. 415–426, 2022.
[101] E. Tsalera, A. Papadakis, M. Samarakou, and I. Voyiatzis, "Feature extraction with handcrafted methods and convolutional neural networks for facial emotion recognition," Appl. Sci. (Basel), vol. 12, no. 17, p. 8455, 2022.
[102] C. Gautam and K. R. Seeja, "Facial emotion recognition using Handcrafted features and CNN," Procedia Comput. Sci., vol. 218, pp. 1295–1303, 2023.
[103] Y. Ming, H. Qian, and L. Guangyuan, "CNN-LSTM facial expression recognition method fused with two-layer attention mechanism," Comput. Intell. Neurosci., vol. 2022, p. 7450637, 2022.
[104] A. Aggarwal et al., "Two-Way Feature Extraction for Speech Emotion Recognition Using Deep Learning," Sensors, vol. 22, no. 6, p. 2378, Mar. 2022, doi: 10.3390/s22062378.
[105] H. -W. Guo, Y. -S. Huang, C. -H. Lin, J. -C. Chien, K. Haraikawa and J. -S. Shieh, "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), Taichung, Taiwan, 2016, pp. 274-277, doi: 10.1109/BIBE.2016.40