

Review

Assessing the Applicability of Machine Learning Models for Robotic Emotion Monitoring: A Survey

Md Ayshik Rahman Khan ¹, Marat Rostov ², Jessica Sharmin Rahman ^{2,3}, Khandaker Asif Ahmed ^{3,*}
and Md Zakir Hossain ^{2,3,*}

¹ Department of Information and Communication Technology, La Trobe University, Melbourne, VIC 3083, Australia

² School of Computing, College of Engineering and Computer Science, The Australian National University (ANU), Canberra, ACT 2601, Australia

³ The Commonwealth Scientific and Industrial Research Organisation (CSIRO), Canberra, ACT 2601, Australia

* Correspondence: khandakerasif.ahmed@csiro.au (K.A.A.); zakir.hossain@anu.edu.au (M.Z.H.)

Abstract: Emotion monitoring can play a vital role in investigating mental health disorders that contribute to 14% of global diseases. Currently, the mental healthcare system is struggling to cope with the increasing demand. Robot-assisted mental health monitoring tools can take the enormous strain off the system. The current study explored existing state-of-art machine learning (ML) models and signal data from different bio-sensors assessed the suitability of robotic devices for surveilling different physiological and physical traits related to human emotions and discussed their potential applicability for mental health monitoring. Among the selected 80 articles, we subdivided our findings in terms of two different emotional categories, namely—discrete and valence-arousal (VA). By examining two different types of signals (physical and physiological) from 10 different signal sources, we found that RGB images and CNN models outperformed all other data sources and models, respectively, in both categories. Out of the 27 investigated discrete imaging signals, 25 reached higher than 80% accuracy, while the highest accuracy was observed from facial imaging signals (99.90%). Besides imaging signals, brain signals showed better potentiality than other data sources in both emotional categories, with accuracies of 99.40% and 96.88%. For both discrete and valence-arousal categories, neural network-based models illustrated superior performances. The majority of the neural network models achieved accuracies of over 80%, ranging from 80.14% to 99.90% in discrete, 83.79% to 96.88% in arousal, and 83.79% to 99.40% in valence. We also found that the performances of fusion signals (a combination of two or more signals) surpassed that of the individual ones in most cases, showing the importance of combining different signals for future model development. Overall, the potential implications of the survey are discussed, considering both human computing and mental health monitoring. The current study will definitely serve as the base for research in the field of human emotion recognition, with a particular focus on developing different robotic tools for mental health monitoring.

Keywords: emotion monitoring; mental health; machine learning; human-robot interaction; signals; robots



Citation: Khan, M.A.R.; Rostov, M.; Rahman, J.S.; Ahmed, K.A.; Hossain, M.Z. Assessing the Applicability of Machine Learning Models for Robotic Emotion Monitoring: A Survey. *Appl. Sci.* **2023**, *13*, 387. <https://doi.org/10.3390/app13010387>

Academic Editors: Li Liu and Yuxin Peng

Received: 18 November 2022

Revised: 17 December 2022

Accepted: 20 December 2022

Published: 28 December 2022



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

1. Introduction

Mental health plays a vital role in our overall well-being. However, in recent times, mental health issues have significantly escalated. A survey on the societal mental health of 600,000 U.S. people showed that the number of adolescents reporting a depressive episode had doubled between 2009–2017, and many eventually resulted in suicide [1]. This study clearly indicates the constantly growing mental health issues within our society. Numerous mental health issues are linked to social isolation [2–4]. The concern over this issue is even more intensified by the upward trend in single-person households, especially in developed countries, where the number is alarmingly as high as 60% [5]. Furthermore,

loneliness is not limited to adolescents either. Many elderly people are getting less familial support and ending up living alone. Hence, our society is on the verge of brimming with loneliness, and mental health is sure to deteriorate if nothing is done to remedy the situation. Socially assistive robots can be useful in dealing with loneliness as they can also function as companions [6,7].

Assistive robots have already seen widespread success in the healthcare and medicine sectors [8]. Their versatile contributions in these sectors, such as surgeries [9], radiation therapy, cancer treatment, and animal therapies [10], also led us to believe that robots can play a crucial role in coping with the current mental health situation worldwide. One possible use is to monitor patients' mental health and refer them to a professional neuro-therapist. The traditional approach to mental health monitoring is wholly based on patients recounting their days. In the end, professionals have to rely on the patients to give a true and accurate recount of their health. However, people often face difficulties remembering events accurately. Further, sadness is highly correlated with depression among patients and is a prime component of clinical diagnoses [11]. Therefore, robots can prove highly beneficial in monitoring mental health through emotion monitoring. Currently, numerous ML methods and literature reviews have shown the potentiality of different sensors to monitor human emotions, but there is still a research gap between identifying suitable signal sources and ML models for robotic applications. Moreover, our literature search could not find any uniform methodology or analysis to assess available resources. As different studies used different datasets and sources with varying evaluation metrics, it is indeed a challenging task to make a proper comparative analysis.

While conducting the survey, we also came across a few survey and review papers in the same or similar field. Dzedzickis et al. [12] performed a review of the sensor and methods used for mental health monitoring. However, the pivoting factors of their work were the sensors and the engineering view of the emotion recognition process. The survey conducted by Mohammed et al. [13] also did not prioritize the machine learning methods used for emotion recognition. Instead, they focused on the challenges faced by researchers in developing a human–robot interaction system. Saxena et al. [14] performed a separate analysis of the ML methods and feature-based techniques, lacking the robotic applicability of these approaches. Moreover, their survey mostly involved discrete emotion recognition, with a single study of the valence-arousal category. The foremost objective of Yadav et al. [15] was speech emotion recognition and visual systems. Many other signal sources that could potentially contribute to emotion recognition were not considered in their review.

This paper aims to analyze and determine which machine learning methods and signal sources are the most appropriate for emotion monitoring through robots. While there is plenty of research on emotion recognition and mental health, we systematically set up a boundary to capture the latest works in this field. We have considered all the papers relevant to emotion recognition and monitoring through machine learning from June 2015 to August 2022. Moreover, machine learning is one of the core subject matters of this survey. However, all the machine learning algorithms were not contemplated in the survey because the researchers preferred to use the most sought-after methods for their experiments. Further, even though there is a third emotion category (hierarchical model), we only considered the two most widely used emotional categories. Recognizing emotions through robots could provide an accurate account of people's mental states. To make this a reality, we must determine the means to recognize emotions accurately. High accuracy for classifying emotion was prioritized in the decision-making. Ease of implementation, accessibility of signal sources, and highly accurate ML methods are also key factors. The current study can be utilized for future implementations of robotic mental health monitoring.

2. Background

A brief description of different signal sources, machine learning models, and emotional levels is essential to set up baseline knowledge about the types of signal sources, the mode

of acquisition, suitable ML models, and their applicability in humanized robots. It also helps readers to understand the technical and intellectual difficulties of achieving an optimal outcome from ML models or datasets. Below, we first discussed seven human signal sources, followed by ML models, and lastly discussed two different emotional levels.

2.1. Human Signal Sources

Different signals are generated from different parts of the human body. We have found seven signal sources and Figure 1 has illustrated their originating places.

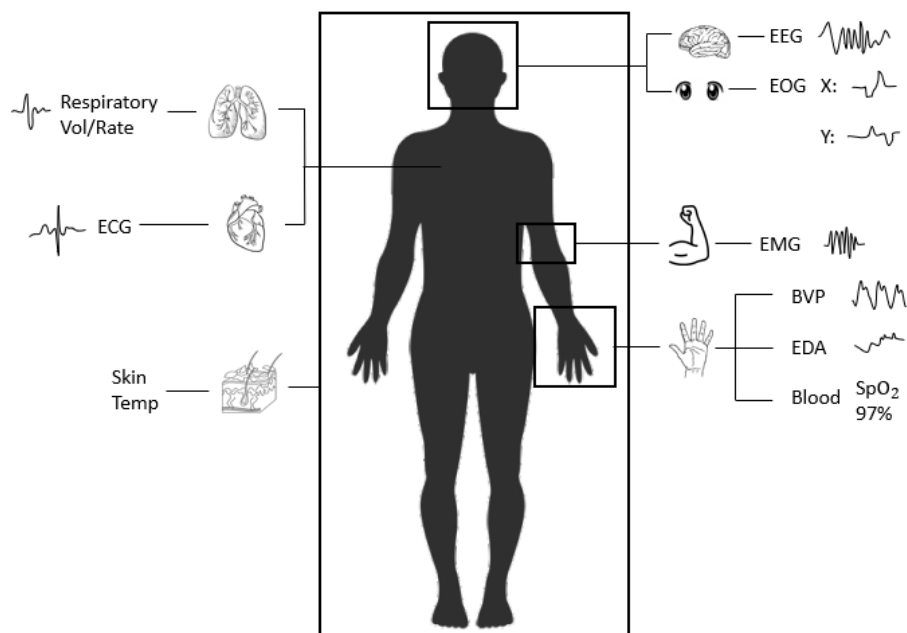


Figure 1. Several physiological signals with sources.

2.1.1. Brain Signals

The signals that are generated from the brain or related to brain activity are considered brain signals. Brain activity can be defined as the neurons sending signals to each other [16]. Two types of signals fall into this category: Electroencephalogram (EEG) and Electrooculography (EOG). Measuring brain activities through EEG is a difficult process. Electroencephalograms (EEG) measure the surface potential of the scalp. Although the brain is insulated from the outside by the skull and other tissue, this electrical activity still presents minuscule changes in the electrical potential across the scalp [17]. These changes can be recorded. However, since these voltages fluctuate in the micro-volt range [18], EEG measurements are very prone to noise. However, as EEG is directly measured by the brain's activity, it has some obvious impacts to our mental state.

Electrooculography (EOG) measures the electrical potential between the front and the back of the eye [19]. Like other electrical potential physiological sensors, this is accomplished through several surface electrodes. This time, the electrodes are placed above and below or left and right of the eye to be measured. Eye position trackers are also often used in this field, although eye position is more voluntarily controlled. Both EOG and eye position have become increasingly relevant in emotion recognition studies [20]. One reason for this is the role eye contact plays in human and primate behavior. Studies have shown that the decoding of facial expressions is significantly impeded without clear eye contact [21].

2.1.2. Heart Signals

The driving force of the cardiovascular system, the heart, plays a key role in our physiology. Even characteristics that can be taken at a non-invasive level, such as blood

pressure [22], contain information that can be used to infer emotional state. The most common example of this is determining whether a person is stressed or not. When stressed, the body releases a hormone that increases heart rate and blood pressure, and affects many other cardiovascular interactions [23]. These signals can be obtained through blood pressure sensors.

Blood volume pulse (BVP) measures how much blood moves to and from a site over time. This signal is usually captured through optical, non-invasive means using a photoplethysmogram (PPG). A PPG works through shining light, usually in the infrared wavelengths, onto the body surface of interest [24]. Since tissues reflect light more than the hemoglobin found in red blood cells [25], measuring the amount of light that returns to the device can give an idea of how much blood is in the site. These devices are so widespread now that they have been integrated into many smartphones [26] and smart watches [27] created in the last decade. Although BVP signals have successfully been used to measure peoples' heart rates in recent years [28], these signals can provide a lot more information than just that. The amplitude of the pulses and how much volume is being transferred can provide key insights into the cardiovascular health of a patient as well as their mental state [25].

Electrocardiograms (ECG) measure the heart's electrical activity. It accomplishes this by sensing the voltage of conductors placed on the chest surface, i.e., electrodes. Historically, this has been used to evaluate how the heart functions in a patient [29]. However, due to the strong connection between heart activity and mental state, it is now a mainstream physiological measurement for effective computing.

2.1.3. Skin Signals

Electrodermal activity (EDA) is the umbrella term for any electrical changes that occur within the skin. Like ECG, this is measured through surface electrodes, often placed on the body's extremities. One signal that is measured under EDA is the galvanic skin response (GSR). GSR, a measure of the conductivity of human skin, can provide an indication of changes in stress levels in the human body [30]. Depending on our emotional state's intensity, our skin produces sweat at different rates [31]. This results in the electrical conductance of our skin changing, which can then be measured.

In addition to conductance, our skin temperature (ST) also changes in response to several internal and external stimuli. External stimuli, such as ambient temperature and any applied heat sources, are of little use to researchers studying physiological signals. However, internal causes of skin temperature change are of great interest. Emotionally charged music causes skin temperatures to rise or fall [32]. Therefore, emotional states are linked to skin temperature, and it may be possible to infer patients' emotions through skin temperature. Skin temperature is measured either through contactless or contact methods, but recently, researchers have taken great interest in contactless methods, such as infrared cameras, that measure temperatures by collecting radiation emitted by the surface [33].

2.1.4. Lungs Signals

Lung signals are physiological data of great interest in numerous studies. Breathing has been proven to change in response to emotional changes [34]. Typically, respiration data comes in the form of respiration volume (RV) and respiration rate (RR) and can be recorded by a single device, such as a wearable strain sensor [35]. Their implementation varies widely, with some new methods even going contactless [36], but often a belt with several sensors is used, such as one by Neulog [37]. Oxygen saturation [38] can also contain information regarding emotional state. SpO₂ and HbO₂ are measures of oxygen saturation.

2.1.5. Imaging Signals

Facial expression recognition from imaging signals is potentially a great technique for emotion recognition. Facial expressions are a great way to detect human emotions, as human faces most often portray their internal emotional states [39]. General facial

expression recognition involves three key phases—preprocessing, feature extraction, and classification [40]. The process of emotion recognition from facial expressions can be performed from either facial images or video extracts that comprise facial expressions. While, in most cases, the images used for the recognition are 2D, depth information can also be incorporated into these 2D images using 3D sensors [41].

Our bodily expressions can also be a significant source of information regarding our emotions. Postural, kinematic, and geometrical features can be conveyors of human emotions [42]. Our emotional states also reflect in our walking and sitting actions [43]. Motion data collected using the RGB sensors are used for detection purposes [42].

2.1.6. Gait Sequences

Human gait refers to the walking style of a person. Alongside uniquely identifying the walker, it can also be used for detecting the walker's emotions [44]. Kinetic or motion data collected by motion capture sensors, in combination with neural network algorithms, can be very effective in recognizing human emotions from gait sequences [45,46]. Traditionally, gait was measured by motion capture systems, force plates, electromyography, etc.; however, the emergence of modern technologies, such as accelerometers, electrogoniometers, gyroscopes, in-shoe pressure sensors, etc., has made gait analysis much easier and more efficient [47].

2.1.7. Speech Signals

Speech is the most commonly used and one of the most important mediums of communication for humans. The signals generated from human voice or audio clips are considered speech signals. Speech signals can contain information regarding the message, speaker, language, and emotion [48]. Therefore, speech signals have been of great interest to researchers for emotion recognition. Similar to imaging signals, the general approach for speech emotion recognition has three stages; signal preprocessing, feature extraction, and feature classification [49]. Besides audio extracts, speech signals are usually collected from smartphones and wearable devices where local interaction of on-body, environmental, and location sensor modalities are merged [50].

2.2. ML Models

In order to develop a reliable mental health monitoring system through emotion recognition, it is important to figure out which machine learning methods perform most efficiently in this sector. To accurately classify emotions, the machine learning methods are fed with input signals (discussed in Section 2.1). The system determines a pattern from this input data and uses this information to correctly classify emotions. Afterwards, when non-classified data are fed to the classifier, it outputs a single discrete emotional category or several numerical values corresponding to a position on an emotional plane.

Among diverse machine learning methods, the most commonly found methods are: Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Convolutional Neural Network (CNN), Deep Neural Network (DNN), Artificial Neural Network (ANN), Decision Tree (DT), Random Forest (RF), Multilayer Perceptron (MLP), etc. We also came across a few hybrid or ensemble methodologies, such as voting, adaboost, a combination of Principal component analysis (PCA) and Linear discriminant analysis (LDA), etc. These techniques are comprehensively included in several studies, and thus interested readers are referred to [51–53]. The majority of machine learning methods evaluate performance in accuracy [54]. However, there are a few other evaluation metrics, e.g., correlation coefficient and RMSE (Root Mean Square Error). The correlation coefficient is the statistical measure of how strongly related two variables are [55]. Moreover, RMSE refers to the numerical error between a prediction and its true value [56]. A lower error corresponds to a closer match between the estimator and the ground truth.

2.3. Emotional Levels

Emotion/ facial expression recognition has been an intriguing topic to explore since last century [57,58]. For emotion recognition, machine learning models are scored against the classification performed prior by psychologists or trial participants. Emotions are classified based on emotional models. Emotional models illustrate how many discrete emotion states there are (e.g., a model just containing happy, neutral, and sad) or how many levels of emotion there are (e.g., anger intensity). Hence, over the last few centuries, there have been numerous emotional models created. We can differentiate between these as models using discrete emotional classes and those using dimensional models. A hierarchical framework has also been introduced by Metallinou et al. [59] as a category of emotion recognition. The hierarchical model is capable of incorporating multimodal information and temporal context from speakers, and their experiments suggest that their multimodal classifiers can proficiently outperform unimodal classifiers. However, the focus of our survey was limited to only discrete and dimensional emotional classes.

2.3.1. Discrete Emotions

The emotions that a person is likely to experience in their day-to-day life are labeled as discrete emotions. For instance, happiness, sadness, boredom, neutrality, and anger are examples of discrete emotions. One type of discrete model that has been used in multiple studies is the “basic 6” or the Ekman model [60]. Ekman et al. identified six basic emotions that covered most human interactions. These are anger, disgust, fear, happiness, sadness, and surprise [61]. Although many other models have introduced more emotions, such as Lazarus’s 15-emotion model [62], many researchers find the basic 6 to work well for their purposes. A potential reason for many to prioritize the basic 6 models over the other models is that having more classes requires higher accuracy to get a similar significance of results.

2.3.2. Continuous Emotions

Some emotional models use dimensions for emotion instead. Amongst them, the most popular is the Valence-Arousal emotion model. This has two dimensions, as the name would suggest: valence and arousal. Valence ranges from feeling pleasant to unpleasant, and arousal ranges from feeling quiet to active [63]. Due to the lack of discretisation of these dimensions, the number of specific emotions that can be located on the valence-arousal plane is infinite. However, our typical everyday emotion models, such as Ekman’s 6-emotion model, can be mapped out onto this plane. Figure 2 shows an example of a correspondence between a discrete emotional model and the V-A (valence-arousal) plane.

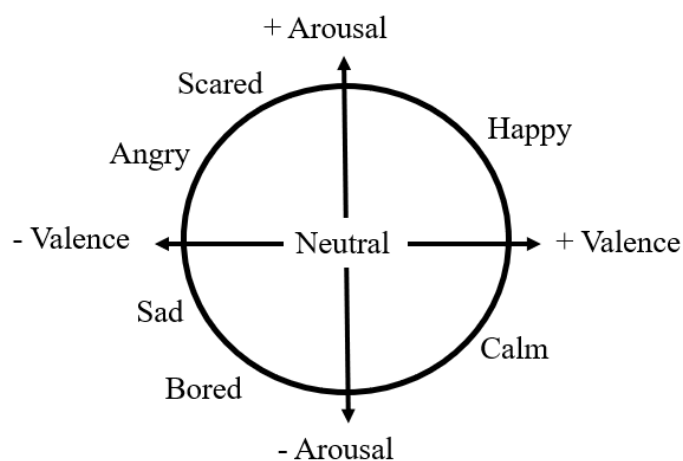


Figure 2. Valence-Arousal emotional model.

Although mapping discrete emotions on continuous models is often performed, classifying emotions on the continuous labels presents an interesting challenge. The classification task, in this case, is exceptionally complicated since there are infinite points along any continuous dimension and hence infinite classes. However, researchers have come up with several ways to accomplish this, and one of them is through the segmentation of the V-A plane into compartments of interest. For example, splitting the plane into four quadrants: high valence + high arousal, high valence + low arousal, low valence + high arousal, and low valence + low arousal.

3. Methods

We explored six academic databases due to their relevance to the topic. These databases were IEEE Xplore, Google Scholar, ANU SuperSearch, Scopus, Pubmed Central, and ResearchGate. In order to search these databases, a set of keywords was derived in consultation with university librarians. These keywords were robot*, emotion recognition, and sensor* (where * denotes wildcard characters). For consistency, the six databases were searched to find papers that contained all three terms in any meta field. As technology is rapidly evolving and to keep our research accurate-to-date, our search results were narrowed down to papers published in the last seven years; from 1 June 2015 to 1 August 2022.

If the databases searched this way had less than 200 results, all papers were added to be screened. In the case where there were more than 200 results (Google Scholar, ANU SuperSearch, and Scopus), the results were sorted by the engine's definition of relevance, and the top 200 results were added to the screening pool. This resulted in a collection of 1141 articles in total, of which 885 were unique.

Following this, the records were screened in order to provide papers relevant to our research interests. Articles were excluded if they were either: (1) publications that were not original peer-reviewed papers, (2) not related to emotion recognition, (3) do not mention the applicability of research to improving robots or machines or agents, or (4) do not state research applicability in a mental health context. For the validation of excluded results, two reviewers acted independently on the 885 unique papers. There were 17% mismatched papers from each reviewer. By discussing with the third reviewer, 80 papers were finally included for detailed analyses. The process and milestones are displayed in Figure 3. Then, both quantitative and qualitative data were extracted from the 80 papers. The aim of each paper, the number of participants, the physiological data used, the methods used for emotion classification, the emotional category type, and the outcome of each paper were recorded.

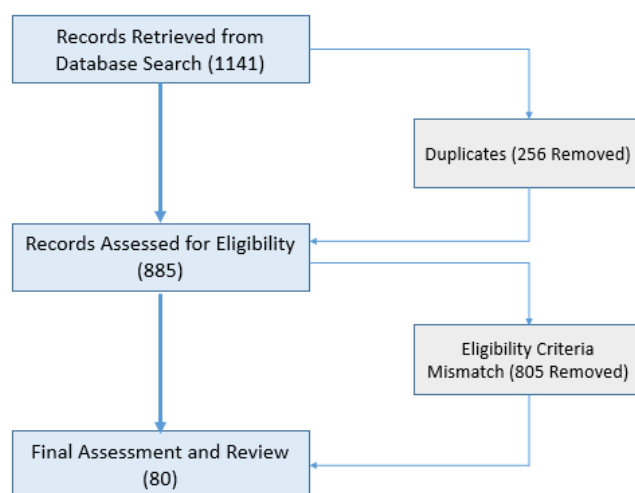


Figure 3. Screening process for the review.

4. Results and Discussion

Following data extraction, we assembled the classification accuracy results from the papers and identified different signal sources. We found a total of 18 signal sources generated from different parts of the human body. Applying different machine learning methods, the studies attempted to identify or monitor human emotions. Among all papers, only one [64] used fully synthesized data, where no participants were involved. Due to many experiments involving multiple sensors, the number of signal sources and sensors is greater than the number of experiments. In total, 112 signals were studied across different physical and physiological sources, namely Brain, Lung, Skin, Heart, Muscle, Imaging, Speech, Tactile, etc. (Details will be found in Supplementary Material). The choice of the classifier plays a key role in accurately classifying emotions. Therefore, in various experiments, we have come across multiple supervised, unsupervised, and hybrid classifiers.

To allow for accurate comparisons, papers are split into two main categories of emotion: discrete and valence-arousal. However, even among discrete emotions studies, there are intensity experiments, e.g., anger intensity [65] and stress level [66]. This adds another dimension to an emotion classification task. Since this is not the same kind of classification as mapping out a user state to the six basic emotions, and all results in a shared category should be comparable to each other, this kind of experiment is categorized as “other”. Gesture recognition tasks that are not validated to emotions are also in the other category. The distribution of emotion classification type is illustrated in Figure 4.

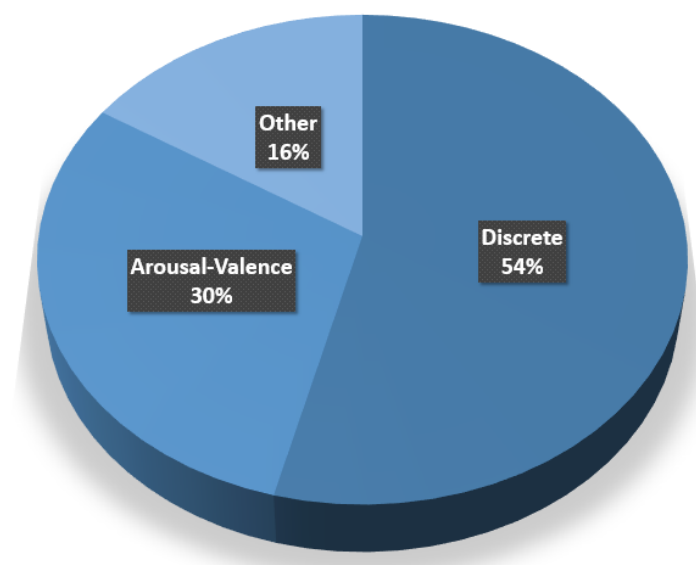


Figure 4. Used emotional model for recognition tasks.

The focus of our study was detecting emotion correctly for better mental health monitoring. Amongst the 80 papers, 70 papers provided single or a range of accuracy percentages as an evaluation matrix. The remaining eight studies used different evaluation metrics for their experiments and, therefore, are excluded from Figures 5 and 6. Carmona et al. [64] calculated their results in terms of sensitivity and specificity, whereas Yu et al. [65] used RMSE to evaluate the accuracy of predicted ratings. Spaulding et al. [67] measured their performance in terms of area under curve where the result varied from 55% to 62%. The experiments performed by Wei et al. [68] and Mencattini et al. [69] were evaluated in terms of the correlation coefficient. On the other hand, Bhatia et al. [45] evaluated their performance on the basis of mean average precision, whereas Hassani et al. [70] and Yun et al. [71] utilized predictive and statistical data analysis rather than classification. Few studies did not include any evaluation matrix at all, as their aims were beyond classification tasks. Al-Qaderi et al. 2018 [72] proposed a perceptual framework for emotion recognition. Miguel et al. 2019 [73] showed that socio-emotional brain areas do not react to effective

touch in infants. These provide conclusions to their research questions but do not yield a percentage accuracy figure.

Although discrete emotion classification experiments take up 54% of the total, it is nearly matched by experiments mapping emotions to the valence-arousal plane (30%). The remaining 16% are emotion models that do not fit any of these emotion-label categories. Most experiments classified emotions using discrete labels, such as happy, neutral, sad, or the continuous valence-arousal plane. How studies classified emotions using the plane varied; however, some split the V-A plane into quadrants to create four emotion labels (Figure 2). Meanwhile, some others measured distance along the valence and arousal axes. To visualize the findings of our survey, we created two separate scatter plots for the discrete (Figure 5) and valence-arousal (Figure 6) categories. Highest accuracies for the discrete and the V-A categories are shown in Tables 1 and 2. The graphs were plotted based on the data we assembled from our study across the 80 papers. Four papers ([50,74–76]) did not provide a separate accuracy for valence and arousal. Instead, they provided an overall accuracy for their whole experiment. The neural network-based methodologies were commonly plotted under the label NN. Similarly, the Bayes variants were commonly denoted as Bayesian and tree-based methods were placed under DT. Hybrid methodologies or the combination of different methods are commonly denoted as fusion.

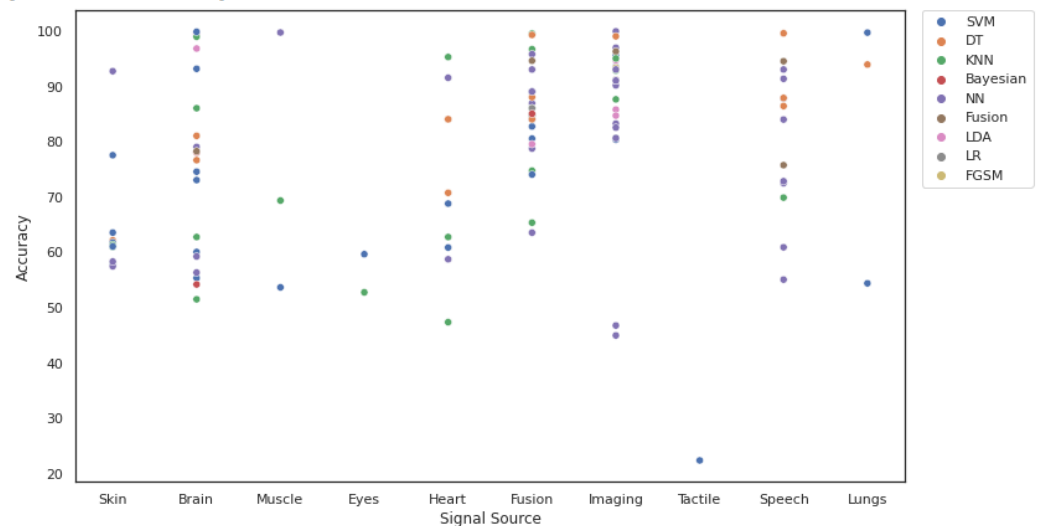


Figure 5. Scatter plot for the discrete emotional category.

Table 1. Summary of the included papers in the discrete emotional category.

Authors	Participant No.	Source	Dataset	Methods	Highest Accuracy
Latif et al. [77]	1	Skin	ST	SVM	63.5
Fan et al. [78]	16	Brain	EEG	KNN	86
Khezri et al. [79]	25	Fusion	EEG, EMG	SVM	82.7
Tivatansakul et al. [80]	8	Heart	ECG	KNN	95.25
Boccanfuso et al. [81]	10	Skin	ST, EDA	SVM	77.5
Ruiz-Garcia et al. [82]	188	Imaging	RGB	NN	96.93
Mehmood et al. [74]	21	Brain	EEG	DT	76.6
Mohammadpour et al. [83]	32	Brain	EEG	NN	59.19
Lowe et al. [84]	64	Tactile	Tactile	SVM	22.3
Noor et al. [85]	44	Fusion	SA, RGB	KNN	96.67

Table 1. Cont.

Authors	Participant No.	Source	Dataset	Methods	Highest Accuracy
Ruiz-Garcia et al. [86]	70	Imaging	RGB	Fusion	96.26
Wei et al. [87]	30	Fusion	EEG, ECG, Resp., EDA	SVM	84.60
Wei et al. [88]	27	Fusion	EEG, ECG, Resp., EDA	SVM	84.62
Goulart et al. [89]	28	Imaging	RGB, ST	LDA	85.75
Gu et al. [90]	14	Fusion	Motion, RGB, Radio	KNN	84.8
Huang et al. [91]	487	Speech	SA	Fusion	94.50
Ilyas et al. [92]	221	Imaging	RGB	Fusion	91
Lopez-Rincon et al. [93]	1192	Imaging	RGB	NN	44.9
Ma et al. [94]	52	Fusion	SA, RGB	NN	86.89
Mithbavkar et al. [95]	1	Muscle	EMG	NN	99.69
Rahim et al. [96]	40	Fusion	ECG, GSR	NN	93
Taran et al. [97]	20	Brain	EEG	SVM	93.13
Balan et al. [98]	8	Fusion	EEG, HR, EDA	KNN	99.5
Chen et al. [99]	4	Speech	SA	DT	87.85
Ding et al. [100]	4	Imaging	3D	KNN	92.8
Melinte et al. [101]	24,336	Imaging	RGB	NN	90.14
Shu et al. [102]	25	Heart	HR	DT	84
Uddin et al. [103]	339	Speech	SA	NN	93
Yang et al. [104]	3	Imaging	RGB	NN	99.9
Zvarevashe et al. [105]	28	Speech	SA	DT	99.55
Ahmed et al. [43]	30	Imaging	RGB	LDA	94.67
Kumar et al. [106]	94	Imaging	RGB	NN	91.02
Hsu et al. [107]	32	Fusion	GSR, SA, ECG	KNN	86
D'Onofrio et al. [108]	27	Imaging	RGB	DT	99
Modi et al. [109]	Sim.	Imaging	RGB	NN	82.5
Chang et al. [110]	Sim.	Brain	EEG	LDA	99.44
Mittal et al. [111]	10	Fusion	RGB, SA	NN	89
Tuncer et al. [112]	Sim.	Brain	EEG	SVM	99.82
Nimmagadda et al. [113]	Sim.	Imaging	RGB	NN	80.6
Zhao et al. [114]	12	Imaging	RGB	FGSM	93.31
Ilyas et al. [115]	10	Speech	SA	NN	91.32
Martínez-Tejada et al. [116]	40	Fusion	EEG	NN	89
Filippini et al. [117]	24	Imaging	RGB	NN	91
Hefter et al. [118]	70,000	Imaging	RGB	NN	93
Shan et al. [119]	84	Lungs	KINECT	SVM	99.67
Gümüslü et al. [120]	15	Fusion	EEG, BVP, ST, SC	Fusion	94.58
Mocanu et al. [121]	24	Speech	SA	NN	83.95

In Figure 5, the accuracies of the methods are plotted against the source of the signals. Different colors denote the methodologies used in the experiments. If any of the experiments were conducted under different experimental settings, the best results from all of those settings were considered. Further, if any of the experiments used multiple sources altogether, they were considered a fusion source. The highest level of accuracy was achieved from imaging signals. Amongst the 27 imaging signals, 25 of the signal studies resulted in above 80% accuracy. While two of the imaging signals showed poor accuracies (44.90% and 46.70%), the rest of the imaging signals showed accuracies ranging from 80.33% to 99.90%. Therefore, facial imaging can potentially be the most prevalent signal for emotion recognition. The brain, heart, and skin signal sources provided good accuracies of above 60–70%. Another signal of interest is speech audio, for which classification accuracies varied a lot, from 55% to 99.55%. However, with an accuracy of up to 90% in some cases, speech is definitely a signal worth considering for mental health robots. On the other hand, the tactile signal did not perform well at all. With an accuracy of 22.30%, tactile signals were the worst performer among the discrete signals. A similar thing can be noticed in Figure 6 as well—tactile signals had very low accuracies. Accuracies procured from the eye signals were unsatisfactory as well (52.70% and 59.60%). However, the lung signals and the muscle signals are particularly worth mentioning in this regard as they had a few data points, and it would not be constructive to reach any conclusion based on the average performances. It is worthwhile to mention that most of the fused source signals had accuracies over 80% and more than 90% in some cases. Therefore, another interesting approach to emotion recognition can be fusing signals from different sources

In Figure 6, the accuracies of valence are plotted against the accuracies of arousal. Different colors, shapes, and sizes represent different methodologies, sources, and the number of participants, respectively. For accuracies provided in a range, the maximum value was used in both figures. The only well-performing signal is the brain signals. None of the other signals provided a good accuracy value. If we consider the 10 best accuracies, 8 of them were from brain signals. The brain signals, including other associated signals, might be useful for diagnosing and maintaining other brain disorders, for example multiple sclerosis [122,123] and autism spectrum disorder [124].

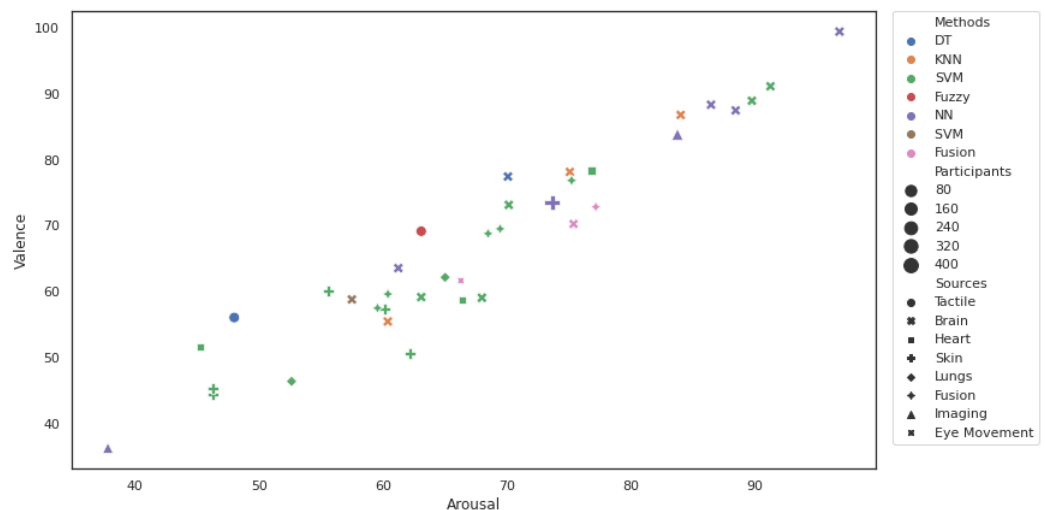


Figure 6. Scatter plot for the Valence-Arousal emotional category.

Table 2. Summary of the included papers in the valence-arousal emotional category.

Authors	Participant No.	Source	Datasets	Methods	Accuracy Valence	Accuracy Arousal
Altun et al. [125]	32	Tactile	Tactile	DT	56	48
Mohammadi et al. [126]	32	Brain	EEG	KNN	86.75	84.05
Wiem et al. [127]	24	Fusion	ECG, RV	SVM	69.47	69.47
Wiem et al. [128]	25	Fusion	ECG, RV	SVM	68.75	68.50
Wiem et al. [129]	24	Fusion	ECG, RV	SVM	56.83	54.73
Yonezawa et al. [130]	18	Tactile	Tactile	Fuzzy	69.1	63.1
Alazrai et al. [131]	32	Brain	EEG	SVM	88.9	89.8
Bazgir et al. [132]	32	Brain	EEG	SVM	91.1	91.3
Henia et al. [133]	27	Fusion	ECG, GSR, ST, RV	SVM	57.44	59.57
Marinoiu et al. [134]	7	Imaging	RGB, 3D	NN	36.2	37.8
henia et al. [135]	24	Fusion	ECG, EDA, ST, Resp.	SVM	59.57	60.41
Salama et al. [136]	32	Imaging	RGB	NN	87.44	88.49
Pandey et al. [137]	32	Imaging	RGB	NN	63.5	61.25
Su et al. [138]	12	Fusion	EEG, RGB	Fusion	72.8	77.2
Ullah et al. [139]	32	Brain	EEG	DT	77.4	70.1
Yin et al. [140]	457	Skin	EDA	NN	73.43	73.65
Algarni et al. [141]	94	Imaging	RGB	NN	99.4	96.88
Panahi et al. [142]	58	Heart	ECG	SVM	78.32	76.83
Kumar et al. [106]	94	Imaging	RGB	NN	83.79	83.79
Martínez-Tejada et al. [116]	40	Brain	EEG	SVM	59	68

For Figures 5 and 6, neural network-based methods outperformed the other methods. K-Nearest Neighbours and Support Vector Machine also performed well in both emotional categories. However, Decision Tree-based methods slightly outperformed KNN and SVM in the discrete emotional category. The highest accuracy was scored by the neural network-based methods, but the most common method found in our studies, with 30% of the total, is SVM. There could be two reasons for the comparatively lower number of papers using NN. First, training NNs or any deep networks require a large number of data. This is hard to come by with physiological signals. Even the largest sample size in all of the papers was 457 for an EDA-based experiment. However, image sets can have potentially thousands of faces, and this does not count video datasets. Another reason for fewer experiments with neural network-based methods could be computational effort, as SVM is a faster process than neural network methods [143].

It is also noticeable that the accuracy of imaging (RGB sensors) appears much higher than other sensors, representing 40% of the papers with the highest accuracies. However, while facial expression accuracy is very high compared to other physiological categories, there is a key difference in emotional validation. At face value, the facial expression is a sort of derived signal where people can counterfeit their smile. Unless we can differentiate between fake and genuine facial expressions, the emotional expressions we get from patients might not always represent their true mental state. However, one could argue that if robots were used in a person's home environment, where they were more likely to be relaxed, they would most likely capture the genuine emotions of the person. Furthermore, there is evidence that facial muscles activate differently depending on whether the person is genuine or acting a smile [144]. However, since none of the papers investigated the

difference between genuine or acted smiles, it remains to be seen how useful standalone imaging (RGB cameras) can be for mental health monitoring.

However, even though emotion recognition using imaging sources scored well, there is much debate on the link between facial recognition and true emotion [145], and for our purpose of mental health monitoring, it is vital that we determine the patients' mental health accurately. Moreover, high computational power is required for deep NN methods to analyse imaging data. Thus it is unclear whether imaging is the most suited sensor for mental health monitoring. In addition, brain signal sources (EEG and ECG), are too invasive and non-consumer friendly to be used in this space [146,147]. However, EDA-based skin signals also performed well. Out of the 80 papers, 17 used skin signals in emotion recognition, and only two used standalone EDA [81,140], while the rest of the studies used a fusion of sensors. Notably, the three EDA emotion recognition experiments that used CNN achieved accuracies ranging from 68.5% to 95%, averaging 79%. It is likely that CNN is a good strategy for the classification of skin data, but incorporating skin-signals in robots remain challenging. The future direction of the study would investigate the feasibility and possibility of skin-based sensors to incorporate physiological signals in robots. Another signal source of interest, at 7.1% of the total, is speech audio. Audio recordings of people talking are used to classify their emotions, which can be easily applied in robots, but its classification accuracy varies a lot through speech audio. With an accuracy of over 90% in some cases [103,105], speech audio is definitely a signal worth considering for mental health robots.

We conducted an extensive survey and found some promising results. However, there are still a few limitations on which we can work in the future. The main limitation of the current study is the lack of common ground for comparison. Each experiment or study is different from another in terms of sources, ML models, and sometimes even in their evaluation metrics. Therefore, our study could not directly compare different models and sources; rather we set up a priority list of sources and ML models to work for robotic emotion monitoring. Moreover, most of the neural network-based methods outperformed other traditional ML methods. However, it was also noticed that, in a lot of cases, the experiments suffered from a lack of data. As per our survey, NN-based LSTM was the highest-performing method for valence-arousal data. However, only a few experiments used NN. Therefore, we still need to explore the applicability of NN models in emotion recognition. Another limitation would be that, even though facial imaging data had the highest level of accuracy among other data sources, most of the works did not consider fake expressions. Humans are capable of faking an expression, which may alter the results. Further, humans are capable of having more than one feeling in a moment, which was not considered in any of the experiments. Therefore, fake emotions and multiple emotions also need to be considered for future experiments in this field.

5. Summary

Our survey assessed 80 latest articles on robotic emotion recognition of two emotional categories—discrete and valence-arousal and discussed the applicability of different sources and ML models in robots. For both categories, our survey found neural network-based methods, especially CNN, performed the best. To be specific, for the discrete category, the highest accuracy of 99.90% was achieved by CNN. Another neural network-based method, LSTM, was the best performer, with accuracies of 99.40% and 96.88% for valence and arousal, respectively. The majority of the experiments that used neural networks had accuracies above 80%. Besides neural network models, SVM can be an alternative model, as this model has been widely used by numerous researchers, with ease of implementation and accuracies of 80% to 99.82%. From the signal sources, Imaging signals were the most proficient and widely used source. Within the VA category, the top eight best-performing models used brain signals as signal sources, showing the great potential of brain signals for this recognition. Despite this, imaging and brain signals also perform well for VA and discrete categories. Among the signal sources, tactile signals performed worst in both

categories, which gives an indication of the cautious usage of tactile signals for human emotion recognition. It is also noticeable that fusion signals performed comparatively better than individual signals. In terms of applicability, brain signals need sophisticated acquisition devices and data processing procedures, while imaging signals can be readily used in ML models. Therefore, to apply within humanized robots for emotion monitoring, we believe imaging sources could be the first choice. Therefore, ML methods, neural networks and SVM, and signal sources of facial imaging should be most promising for further research on emotion monitoring, with some focus on using fusion signals to make the model more robust.

6. Conclusions

We surveyed different ML models and signal sources for emotion monitoring, considering the accessibility, accuracy, and applicability in robots and found imaging can be the most convenient and accurate signal source to consider. Besides imaging, brain and skin signals performed well but were not convenient to implement in robots. However, speech audio has potential applicability in robots, but its varied accuracies raise questions about its application in mental health monitoring. Our survey also showed neural network-based methods, especially CNN, outperformed other machine learning methods. When we compared different emotion categories, we did not find significant differences among the sources and ML models, indicating the potential of using similar sources and models for both discrete and valence-arousal categories. Our study also found fusion signal sources perform better than signal sources. We recommend taking advantage of NN-based models to analyze fusion sources, especially imaging data, as a first choice to solve any tasks related to emotion monitoring. The future direction of our research includes developing a sustainable neural network-based robotic model for monitoring human mental health. Even though we have found that even standalone imaging signals are a convenient way for robotic emotion recognition, using the combination of brain, skin signals, and speech audios can help us attain a very high emotion recognition accuracy, well above standalone CNN-classified data. By conversing with the robot occasionally, the robot would get even more insight into the person's well-being. Thus, this research can help pave the way towards improved social robotics and robot-assisted mental health solutions.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/app13010387/s1>, Table S1: Detailed summary of the included papers.

Author Contributions: Conceptualization, M.Z.H.; methodology, M.R. and M.A.R.K.; validation, M.R., M.A.R.K., J.S.R., K.A.A. and M.Z.H.; formal analysis, M.R. and M.A.R.K.; investigation, M.R. and M.A.R.K.; data curation, M.R., M.A.R.K., K.A.A. and J.S.R.; writing—original draft preparation, M.A.R.K. and K.A.A.; writing—review and editing, M.R., K.A.A. and M.Z.H.; visualization, M.A.R.K.; supervision, K.A.A., J.S.R. and M.Z.H.; project administration, M.Z.H.; funding acquisition, M.Z.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data has been provided in the supplementary section.

Acknowledgments: We thank the Biological Data Science Institute (BDSI) at Australian National University for their extensive support. We also acknowledge Agriculture & Food business unit of the Commonwealth Scientific and Industrial Research Organisation (CSIRO) for their co-operation.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Twenge, J.M.; Joiner, T.E.; Rogers, M.L.; Martin, G.N. Increases in depressive symptoms, suicide-related outcomes, and suicide rates among US adolescents after 2010 and links to increased new media screen time. *Clin. Psychol. Sci.* **2018**, *6*, 3–17. [CrossRef]
2. Novotney, A. The risks of social isolation. *Am. Psychol. Assoc.* **2019**, *50*, 32; Erratum in *Am. Psychol. Assoc.* **2019**, *22*, 2020.
3. Mushtaq, R.; Shoib, S.; Shah, T.; Mushtaq, S. Relationship between loneliness, psychiatric disorders and physical health? A review on the psychological aspects of loneliness. *J. Clin. Diagn. Res. JCDR* **2014**, *8*, WE01. [CrossRef] [PubMed]
4. Loades, M.E.; Chatburn, E.; Higson-Sweeney, N.; Reynolds, S.; Shafran, R.; Brigden, A.; Linney, C.; McManus, M.N.; Borwick, C.; Crawley, E. Rapid systematic review: The impact of social isolation and loneliness on the mental health of children and adolescents in the context of COVID-19. *J. Am. Acad. Child Adolesc. Psychiatry* **2020**, *59*, 1218–1239. [CrossRef]
5. Snell, K. The rise of living alone and loneliness in history. *Soc. Hist.* **2017**, *42*, 2–28. [CrossRef]
6. Bemelmans, R.; Gelderblom, G.J.; Jonker, P.; De Witte, L. Socially assistive robots in elderly care: A systematic review into effects and effectiveness. *J. Am. Med. Dir. Assoc.* **2012**, *13*, 114–120. [CrossRef]
7. Cooper, S.; Di Fava, A.; Vivas, C.; Marchionni, L.; Ferro, F. ARI: The social assistive robot and companion. In Proceedings of the 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), Naples, Italy, 31 August–4 September 2020; pp. 745–751.
8. Castelo, N.; Schmitt, B.; Sarvary, M. Robot Or Human? How Bodies and Minds Shape Consumer Reactions to Human-Like Robots. *ACR N. Am. Adv.* **2019**, *47*, 3.
9. Fox, O.R. Surgeon Completes 100th Knee Replacement Using Pioneering Robot in Bath. *Somersetlive*, 11 December 2020.
10. Case Western Reserve University. 5 Medical Robots Making a Difference in Healthcare at CWRU. Available online: <https://online-engineering.case.edu/blog/medical-robots-making-a-difference> (accessed on 19 December 2022).
11. Mouchet-Mages, S.; Baylé, F.J. Sadness as an integral part of depression. *Dialogues Clin. Neurosci.* **2008**, *10*, 321. [CrossRef]
12. Dzedzickis, A.; Kaklauskas, A.; Bucinskas, V. Human emotion recognition: Review of sensors and methods. *Sensors* **2020**, *20*, 592. [CrossRef]
13. Mohammed, S.N.; Hassan, A.K.A. A survey on emotion recognition for human robot interaction. *J. Comput. Inf. Technol.* **2020**, *28*, 125–146.
14. Saxena, A.; Khanna, A.; Gupta, D. Emotion recognition and detection methods: A comprehensive survey. *J. Artif. Intell. Syst.* **2020**, *2*, 53–79. [CrossRef]
15. Yadav, S.P.; Zaidi, S.; Mishra, A.; Yadav, V. Survey on machine learning in speech emotion recognition and vision systems using a recurrent neural network (RNN). *Arch. Comput. Methods Eng.* **2022**, *29*, 1753–1770. [CrossRef]
16. Brain Basics: The Life and Death of a Neuron. Available online: <https://www.ninds.nih.gov/health-information/public-education/brain-basics/brain-basics-life-and-death-neuron> (accessed on 19 December 2022).
17. Louis, E.K.S.; Frey, L.; Britton, J.; Hopp, J.; Korb, P.; Koubeissi, M.; Lievens, W.; Pestana-Knight, E. Electroencephalography (EEG): An Introductory Text and Atlas of Normal and Abnormal Findings in Adults. *Child. Infants* **2016**. [CrossRef]
18. Malmivuo, J.; Plonsey, R. 13.1 INTRODUCTION. In *Bioelectromagnetism*; Oxford University Press: Oxford, UK, 1995.
19. Soundariya, R.; Renuga, R. Eye movement based emotion recognition using electrooculography. In Proceedings of the 2017 Innovations in Power and Advanced Computing Technologies (i-PACT), Vellore, India, 21–22 April 2017; pp. 1–5.
20. Lim, J.Z.; Mountstephens, J.; Teo, J. Emotion recognition using eye-tracking: Taxonomy, review and current challenges. *Sensors* **2020**, *20*, 2384. [CrossRef] [PubMed]
21. Schurgin, M.; Nelson, J.; Iida, S.; Ohira, H.; Chiao, J.; Franconeri, S. Eye movements during emotion recognition in faces. *J. Vis.* **2014**, *14*, 14. [CrossRef]
22. Nickson, C. Non-Invasive Blood Pressure. 2020. Available online: <https://litfl.com/non-invasive-blood-pressure/> (accessed on 19 December 2022).
23. Porter, M. What Happens to Your Body When You’re Stressed—and How Breathing Can Help. 2021. Available online: <https://theconversation.com/what-happens-to-your-body-when-youre-stressed-and-how-breathing-can-help-97046> (accessed on 19 December 2022).
24. Joseph, G.; Joseph, A.; Titus, G.; Thomas, R.M.; Jose, D. Photoplethysmogram (PPG) signal analysis and wavelet de-noising. In Proceedings of the 2014 Annual International Conference on Emerging Research Areas: Magnetics, Machines and Drives (AICERA/iCMMD), Kottayam, India, 24–26 July 2014; pp. 1–5.
25. Jones, D. The Blood Volume Pulse—Biofeedback Basics. 2018. Available online: <https://www.biofeedback-tech.com/articles/2016/3/24/the-blood-volume-pulse-biofeedback-basics> (accessed on 19 December 2022).
26. Tyapochkin, K.; Smorodnikova, E.; Pravdin, P. Smartphone PPG: Signal processing, quality assessment, and impact on HRV parameters. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; pp. 4237–4240.
27. Pollreisz, D.; TaheriNejad, N. A simple algorithm for emotion recognition, using physiological signals of a smart watch. In Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Korea, 11–15 July 2017; pp. 2353–2356.
28. Tayibnapis, I.R.; Yang, Y.M.; Lim, K.M. Blood Volume Pulse Extraction for Non-Contact Heart Rate Measurement by Digital Camera Using Singular Value Decomposition and Burg Algorithm. *Energies* **2018**, *11*, 1076. [CrossRef]

29. John Hopkins Medicine. Electrocardiogram. Available online: <https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/electrocardiogram> (accessed on 19 December 2022).
30. Shi, Y.; Ruiz, N.; Taib, R.; Choi, E.; Chen, F. Galvanic skin response (GSR) as an index of cognitive load. In Proceedings of the CHI'07 Extended Abstracts on Human Factors in Computing Systems, San Jose, CA, USA, 28 April–3 May 2007; pp. 2651–2656.
31. Farnsworth, B. What is GSR (Galvanic Skin Response) and How Does It Work. 2018. Available online: <https://imotions.com/blog/learning/research-fundamentals/gsr/> (accessed on 19 December 2022).
32. McFarland, R.A. Relationship of skin temperature changes to the emotions accompanying music. *Biofeedback Self-Regul.* **1985**, *10*, 255–267. [[CrossRef](#)]
33. Ghahramani, A.; Castro, G.; Becerik-Gerber, B.; Yu, X. Infrared thermography of human face for monitoring thermoregulation performance and estimating personal thermal comfort. *Build. Environ.* **2016**, *109*, 1–11. [[CrossRef](#)]
34. Homma, I.; Masaoka, Y. Breathing rhythms and emotions. *Exp. Physiol.* **2008**, *93*, 1011–1021. [[CrossRef](#)]
35. Chu, M.; Nguyen, T.; Pandey, V.; Zhou, Y.; Pham, H.N.; Bar-Yoseph, R.; Radom-Aizik, S.; Jain, R.; Cooper, D.M.; Khine, M. Respiration rate and volume measurements using wearable strain sensors. *NPJ Digit. Med.* **2019**, *2*, 1–9. [[CrossRef](#)] [[PubMed](#)]
36. Massaroni, C.; Lopes, D.S.; Lo Presti, D.; Schena, E.; Silvestri, S. Contactless monitoring of breathing patterns and respiratory rate at the pit of the neck: A single camera approach. *J. Sens.* **2018**, *2018*, 4567213. [[CrossRef](#)]
37. Neulog. Respiration Monitor Belt Logger Sensor NUL-236. Available online: <https://neulog.com/respiration-monitor-belt/> (accessed on 19 December 2022).
38. Kwon, K.; Park, S. An optical sensor for the non-invasive measurement of the blood oxygen saturation of an artificial heart according to the variation of hematocrit. *Sens. Actuators A Phys.* **1994**, *43*, 49–54. [[CrossRef](#)]
39. Tian, Y.; Kanade, T.; Cohn, J.F. Facial expression recognition. In *Handbook of Face Recognition*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 487–519.
40. Revina, I.M.; Emmanuel, W.S. A survey on human face expression recognition techniques. *J. King Saud Univ.-Comput. Inf. Sci.* **2021**, *33*, 619–628. [[CrossRef](#)]
41. Li, J.; Mi, Y.; Li, G.; Ju, Z. CNN-based facial expression recognition from annotated rgb-d images for human–robot interaction. *Int. J. Humanoid Robot.* **2019**, *16*, 1941002. [[CrossRef](#)]
42. Piana, S.; Stagliano, A.; Odone, F.; Verri, A.; Camurri, A. Real-time automatic emotion recognition from body gestures. *arXiv* **2014**, arXiv:1402.5047.
43. Ahmed, F.; Bari, A.H.; Gavrilova, M.L. Emotion recognition from body movement. *IEEE Access* **2019**, *8*, 11761–11781. [[CrossRef](#)]
44. Xu, S.; Fang, J.; Hu, X.; Ngai, E.; Guo, Y.; Leung, V.; Cheng, J.; Hu, B. Emotion recognition from gait analyses: Current research and future directions. *arXiv* **2020**, arXiv:2003.11461.
45. Bhatia, Y.; Bari, A.H.; Hsu, G.S.J.; Gavrilova, M. Motion capture sensor-based emotion recognition using a bi-modular sequential neural network. *Sensors* **2022**, *22*, 403. [[CrossRef](#)]
46. Janssen, D.; Schöllhorn, W.I.; Lubienetzki, J.; Fölling, K.; Kokenge, H.; Davids, K. Recognition of emotions in gait patterns by means of artificial neural nets. *J. Nonverbal Behav.* **2008**, *32*, 79–92. [[CrossRef](#)]
47. Higginson, B.K. Methods of running gait analysis. *Curr. Sport. Med. Rep.* **2009**, *8*, 136–141. [[CrossRef](#)] [[PubMed](#)]
48. Koolagudi, S.G.; Rao, K.S. Emotion recognition from speech: A review. *Int. J. Speech Technol.* **2012**, *15*, 99–117. [[CrossRef](#)]
49. Vogt, T.; André, E. Comparing feature sets for acted and spontaneous speech in view of automatic emotion recognition. In Proceedings of the 2005 IEEE International Conference on Multimedia and Expo, Amsterdam, The Netherlands, 6 July 2005; pp. 474–477.
50. Kanjo, E.; Younis, E.M.; Ang, C.S. Deep learning analysis of mobile physiological, environmental and location sensor data for emotion detection. *Inf. Fusion* **2019**, *49*, 46–56. [[CrossRef](#)]
51. Brownlee, J. *Master Machine Learning Algorithms: Discover How They Work and Implement Them from Scratch*. Machine Learning Mastery. 2016. Available online: <https://machinelearningmastery.com/machine-learning-mastery-weka/> (accessed on 9 December 2022).
52. Bonaccorso, G. *Machine Learning Algorithms*; Packt Publishing Ltd.: Birmingham, UK, 2017.
53. Cho, G.; Yim, J.; Choi, Y.; Ko, J.; Lee, S.H. Review of machine learning algorithms for diagnosing mental illness. *Psychiatry Investig.* **2019**, *16*, 262. [[CrossRef](#)] [[PubMed](#)]
54. Juba, B.; Le, H.S. Precision-recall versus accuracy and the role of large data sets. *AAAI Conf. Artif. Intell.* **2019**, *33*, 4039–4048. [[CrossRef](#)]
55. Taylor, R. Interpretation of the correlation coefficient: A basic review. *J. Diagn. Med. Sonogr.* **1990**, *6*, 35–39. [[CrossRef](#)]
56. Willmott, C.J.; Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **2005**, *30*, 79–82. [[CrossRef](#)]
57. Domínguez-Jiménez, J.A.; Campo-Landines, K.C.; Martínez-Santos, J.C.; Delahoz, E.J.; Contreras-Ortiz, S.H. A machine learning model for emotion recognition from physiological signals. *Biomed. Signal Process. Control* **2020**, *55*, 101646. [[CrossRef](#)]
58. Hossain, M.Z.; Gedeon, T.; Sankaranarayana, R. Using temporal features of observers' physiological measures to distinguish between genuine and fake smiles. *IEEE Trans. Affect. Comput.* **2018**, *11*, 163–173. [[CrossRef](#)]
59. Metallinou, A.; Katsamanis, A.; Narayanan, S. A hierarchical framework for modeling multimodality and emotional evolution in affective dialogs. In Proceedings of the 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan, 25–30 March 2012; pp. 2401–2404.

60. Ekman, P.; Friesen, W.V.; Ellsworth, P. *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*; Elsevier: Amsterdam, The Netherlands, 2013; Volume 11.
61. Gu, S.; Wang, F.; Patel, N.P.; Bourgeois, J.A.; Huang, J.H. A model for basic emotions using observations of behavior in *Drosophila*. *Front. Psychol.* **2019**, *10*, 781. [[CrossRef](#)]
62. Krohne, H. Stress and coping theories. *Int. Encycl. Soc. Behav. Sci.* **2002**, *22*, 15163–15170.
63. Kuppens, P.; Tuerlinckx, F.; Russell, J.A.; Barrett, L.F. The relation between valence and arousal in subjective experience. *Psychol. Bull.* **2013**, *139*, 917. [[CrossRef](#)] [[PubMed](#)]
64. Carmona, P.; Nunes, D.; Raposo, D.; Silva, D.; Silva, J.S.; Herrera, C. Happy hour-improving mood with an emotionally aware application. In Proceedings of the 2015 15th International Conference on Innovations for Community Services (I4CS), Nuremberg, Germany, 8–10 July 2015; pp. 1–7.
65. Yu, Y.C. A cloud-based mobile anger prediction model. In Proceedings of the 2015 18th International Conference on Network-Based Information Systems, Taipei, Taiwan, 2–4 September 2015; pp. 199–205.
66. Li, M.; Xie, L.; Wang, Z. A transductive model-based stress recognition method using peripheral physiological signals. *Sensors* **2019**, *19*, 429. [[CrossRef](#)] [[PubMed](#)]
67. Spaulding, S.; Breazeal, C. Frustratingly easy personalization for real-time affect interpretation of facial expression. In Proceedings of the 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII), Cambridge, UK, 3–6 September 2019; pp. 531–537.
68. Wei, J.; Chen, T.; Liu, G.; Yang, J. Higher-order multivariable polynomial regression to estimate human affective states. *Sci. Rep.* **2016**, *6*, 1–13. [[CrossRef](#)]
69. Mencattini, A.; Martinelli, E.; Ringeval, F.; Schuller, B.; Di Natale, C. Continuous estimation of emotions in speech by dynamic cooperative speaker models. *IEEE Trans. Affect. Comput.* **2016**, *8*, 314–327. [[CrossRef](#)]
70. Hassani, S.; Bafadel, I.; Bekhatro, A.; Al Blooshi, E.; Ahmed, S.; Alahmad, M. Physiological signal-based emotion recognition system. In Proceedings of the 2017 4th IEEE International Conference on Engineering Technologies and Applied Sciences (ICETAS), Salmabad, Bahrain, 29 November–1 December 2017; pp. 1–5.
71. de Oliveira, L.M.; Junior, L.C.C.F. Aplicabilidade da inteligência artificial na psiquiatria: Uma revisão de ensaios clínicos. *Debates Psiquiatr.* **2020**, *10*, 14–25. [[CrossRef](#)]
72. Al-Qaderi, M.K.; Rad, A.B. A brain-inspired multi-modal perceptual system for social robots: An experimental realization. *IEEE Access* **2018**, *6*, 35402–35424. [[CrossRef](#)]
73. Miguel, H.O.; Lisboa, I.C.; Gonçalves, Ó.F.; Sampaio, A. Brain mechanisms for processing discriminative and affective touch in 7-month-old infants. *Dev. Cogn. Neurosci.* **2019**, *35*, 20–27. [[CrossRef](#)]
74. Mehmood, R.M.; Du, R.; Lee, H.J. Optimal feature selection and deep learning ensembles method for emotion recognition from human brain EEG sensors. *IEEE Access* **2017**, *5*, 14797–14806. [[CrossRef](#)]
75. Soroush, M.Z.; Maghooli, K.; Setarehdan, S.K.; Nasrabadi, A.M. A novel approach to emotion recognition using local subset feature selection and modified Dempster-Shafer theory. *Behav. Brain Funct.* **2018**, *14*, 1–15.
76. Pan, L.; Yin, Z.; She, S.; Song, A. Emotional State Recognition from Peripheral Physiological Signals Using Fused Nonlinear Features and Team-Collaboration Identification Strategy. *Entropy* **2020**, *22*, 511. [[CrossRef](#)] [[PubMed](#)]
77. Abd Latif, M.; Yusof, H.M.; Sidek, S.; Rusli, N. Thermal imaging based affective state recognition. In Proceedings of the 2015 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS), Langkawi, Malaysia, 18–20 October 2015; pp. 214–219.
78. Fan, J.; Wade, J.W.; Bian, D.; Key, A.P.; Warren, Z.E.; Mion, L.C.; Sarkar, N. A Step towards EEG-based brain computer interface for autism intervention. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; pp. 3767–3770.
79. Khezri, M.; Firoozabadi, M.; Sharafat, A.R. Reliable emotion recognition system based on dynamic adaptive fusion of forehead biopotentials and physiological signals. *Comput. Methods Programs Biomed.* **2015**, *122*, 149–164. [[CrossRef](#)] [[PubMed](#)]
80. Tivatansakul, S.; Ohkura, M. Emotion recognition using ECG signals with local pattern description methods. *Int. J. Affect. Eng.* **2015**, *15*, 51–61. [[CrossRef](#)]
81. Boccanfuso, L.; Wang, Q.; Leite, I.; Li, B.; Torres, C.; Chen, L.; Salomons, N.; Foster, C.; Barney, E.; Ahn, Y.A.; et al. A thermal emotion classifier for improved human-robot interaction. In Proceedings of the 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), New York, NY, USA, 26–31 August 2016; pp. 718–723.
82. Ruiz-Garcia, A.; Elshaw, M.; Altahhan, A.; Palade, V. Deep learning for emotion recognition in faces. In *Proceedings of the International Conference on Artificial Neural Networks*; Springer: Cham, Switzerland, 2016; pp. 38–46.
83. Mohammadpour, M.; Hashemi, S.M.R.; Houshmand, N. Classification of EEG-based emotion for BCI applications. In Proceedings of the 2017 Artificial Intelligence and Robotics (IRANOPEN), Qazvin, Iran, 9 April 2017; pp. 127–131.
84. Lowe, R.; Andreasson, R.; Alenljung, B.; Lund, A.; Billing, E. Designing for a wearable affective interface for the NAO Robot: A study of emotion conveyance by touch. *Multimodal Technol. Interact.* **2018**, *2*, 2. [[CrossRef](#)]
85. Noor, S.; Dhruvo, E.A.; Minhaz, A.T.; Shahnaz, C.; Fattah, S.A. Audio visual emotion recognition using cross correlation and wavelet packet domain features. In Proceedings of the 2017 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), Dehradun, India, 18–19 December 2017; pp. 233–236.
86. Ruiz-Garcia, A.; Elshaw, M.; Altahhan, A.; Palade, V. A hybrid deep learning neural approach for emotion recognition from facial expressions for socially assistive robots. *Neural Comput. Appl.* **2018**, *29*, 359–373. [[CrossRef](#)]

87. Wei, W.; Jia, Q.; Feng, Y.; Chen, G. Emotion recognition based on weighted fusion strategy of multichannel physiological signals. *Comput. Intell. Neurosci.* **2018**, *2018*, 5296523. [[CrossRef](#)] [[PubMed](#)]
88. Wei, W.J. Development and evaluation of an emotional lexicon system for young children. *Microsyst. Technol.* **2019**, *27*, 1535–1544. [[CrossRef](#)]
89. Goulart, C.; Valadão, C.; Delisle-Rodriguez, D.; Funayama, D.; Favarato, A.; Baldo, G.; Binotte, V.; Caldeira, E.; Bastos-Filho, T. Visual and thermal image processing for facial specific landmark detection to infer emotions in a child-robot interaction. *Sensors* **2019**, *19*, 2844. [[CrossRef](#)]
90. Gu, Y.; Wang, Y.; Liu, T.; Ji, Y.; Liu, Z.; Li, P.; Wang, X.; An, X.; Ren, F. EmoSense: Computational intelligence driven emotion sensing via wireless channel data. *IEEE Trans. Emerg. Top. Comput. Intell.* **2019**, *4*, 216–226. [[CrossRef](#)]
91. Huang, Y.; Tian, K.; Wu, A.; Zhang, G. Feature fusion methods research based on deep belief networks for speech emotion recognition under noise condition. *J. Ambient. Intell. Humaniz. Comput.* **2019**, *10*, 1787–1798. [[CrossRef](#)]
92. Ilyas, C.M.A.; Schmuck, V.; Haque, M.A.; Nasrollahi, K.; Rehm, M.; Moeslund, T.B. Teaching pepper robot to recognize emotions of traumatic brain injured patients using deep neural networks. In Proceedings of the 2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), New Delhi, India, 14–18 October 2019; pp. 1–7.
93. Lopez-Rincon, A. Emotion recognition using facial expressions in children using the NAO Robot. In Proceedings of the 2019 International Conference on Electronics, Communications and Computers (CONIELECOMP), Cholula, Mexico, 27 February–1 March 2019; pp. 146–153.
94. Ma, F.; Zhang, W.; Li, Y.; Huang, S.L.; Zhang, L. An end-to-end learning approach for multimodal emotion recognition: Extracting common and private information. In Proceedings of the 2019 IEEE International Conference on Multimedia and Expo (ICME), Shanghai, China, 8–12 July 2019; pp. 1144–1149.
95. Mithbavkar, S.A.; Shah, M.S. Recognition of emotion through facial expressions using EMG signal. In Proceedings of the 2019 International Conference on Nascent Technologies in Engineering (ICNTE), Navi Mumbai, India, 4–5 January 2019; pp. 1–6.
96. Rahim, A.; Sagheer, A.; Nadeem, K.; Dar, M.N.; Rahim, A.; Akram, U. Emotion Charting Using Real-time Monitoring of Physiological Signals. In Proceedings of the 2019 International Conference on Robotics and Automation in Industry (ICRAI), Rawalpindi, Pakistan, 21–22 October 2019; pp. 1–5.
97. Taran, S.; Bajaj, V. Emotion recognition from single-channel EEG signals using a two-stage correlation and instantaneous frequency-based filtering method. *Comput. Methods Programs Biomed.* **2019**, *173*, 157–165. [[CrossRef](#)] [[PubMed](#)]
98. Bălan, O.; Moise, G.; Moldoveanu, A.; Leordeanu, M.; Moldoveanu, F. An investigation of various machine and deep learning techniques applied in automatic fear level detection and acrophobia virtual therapy. *Sensors* **2020**, *20*, 496. [[CrossRef](#)]
99. Chen, L.; Su, W.; Feng, Y.; Wu, M.; She, J.; Hirota, K. Two-layer fuzzy multiple random forest for speech emotion recognition in human-robot interaction. *Inf. Sci.* **2020**, *509*, 150–163. [[CrossRef](#)]
100. Ding, I.J.; Hsieh, M.C. A hand gesture action-based emotion recognition system by 3D image sensor information derived from Leap Motion sensors for the specific group with restlessness emotion problems. *Microsyst. Technol.* **2020**, *28*, 403–415. [[CrossRef](#)]
101. Melinte, D.O.; Vladareanu, L. Facial expressions recognition for human–robot interaction using deep convolutional neural networks with rectified adam optimizer. *Sensors* **2020**, *20*, 2393. [[CrossRef](#)] [[PubMed](#)]
102. Shu, L.; Yu, Y.; Chen, W.; Hua, H.; Li, Q.; Jin, J.; Xu, X. Wearable emotion recognition using heart rate data from a smart bracelet. *Sensors* **2020**, *20*, 718. [[CrossRef](#)] [[PubMed](#)]
103. Uddin, M.Z.; Nilsson, E.G. Emotion recognition using speech and neural structured learning to facilitate edge intelligence. *Eng. Appl. Artif. Intell.* **2020**, *94*, 103775. [[CrossRef](#)]
104. Yang, J.; Wang, R.; Guan, X.; Hassan, M.M.; Almogren, A.; Alsanad, A. AI-enabled emotion-aware robot: The fusion of smart clothing, edge clouds and robotics. *Future Gener. Comput. Syst.* **2020**, *102*, 701–709. [[CrossRef](#)]
105. Zvarevashe, K.; Olugbara, O. Ensemble learning of hybrid acoustic features for speech emotion recognition. *Algorithms* **2020**, *13*, 70. [[CrossRef](#)]
106. Kumar, A.; Sharma, K.; Sharma, A. MEMoR: A multimodal emotion recognition using affective biomarkers for smart prediction of emotional health for people analytics in smart industries. *Image Vis. Comput.* **2022**, *123*, 104483. [[CrossRef](#)]
107. Hsu, S.M.; Chen, S.H.; Huang, T.R. Personal Resilience Can Be Well Estimated from Heart Rate Variability and Paralinguistic Features during Human–Robot Conversations. *Sensors* **2021**, *21*, 5844. [[CrossRef](#)] [[PubMed](#)]
108. D’Onofrio, G.; Fiorini, L.; Sorrentino, A.; Russo, S.; Ciccone, F.; Giuliani, F.; Sancarlo, D.; Cavallo, F. Emotion Recognizing by a Robotic Solution Initiative (EMOTIVE Project). *Sensors* **2022**, *22*, 2861. [[CrossRef](#)] [[PubMed](#)]
109. Modi, S.; Bohara, M.H. Facial emotion recognition using convolution neural network. In Proceedings of the 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 6–8 May 2021; pp. 1339–1344.
110. Chang, Y.; Sun, L. EEG-Based Emotion Recognition for Modulating Social-Aware Robot Navigation. In Proceedings of the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Virtual, 1–5 November 2021; pp. 5709–5712.
111. Mittal, T.; Bhattacharya, U.; Chandra, R.; Bera, A.; Manocha, D. M3er: Multiplicative multimodal emotion recognition using facial, textual, and speech cues. In Proceedings of the AAAI Conference on Artificial Intelligence, Virtual, 25–29 October 2021; Volume 34, pp. 1359–1367.
112. Tuncer, T.; Dogan, S.; Subasi, A. A new fractal pattern feature generation function based emotion recognition method using EEG. *Chaos Solitons Fractals* **2021**, *144*, 110671. [[CrossRef](#)]

113. Nimmagadda, R.; Arora, K.; Martin, M.V. Emotion recognition models for companion robots. *J. Supercomput.* **2022**, *78*, 13710–13727. [[CrossRef](#)]
114. Zhao, Y.; Xu, K.; Wang, H.; Li, B.; Qiao, M.; Shi, H. MEC-enabled hierarchical emotion recognition and perturbation-aware defense in smart cities. *IEEE Internet Things J.* **2021**, *8*, 16933–16945. [[CrossRef](#)]
115. Ilyas, O. Pseudo-colored rate map representation for speech emotion recognition. *Biomed. Signal Process. Control.* **2021**, *66*, 102502.
116. Martínez-Tejada, L.A.; Maruyama, Y.; Yoshimura, N.; Koike, Y. Analysis of personality and EEG features in emotion recognition using machine learning techniques to classify arousal and valence labels. *Mach. Learn. Knowl. Extr.* **2020**, *2*, 7. [[CrossRef](#)]
117. Filippini, C.; Perpetuini, D.; Cardone, D.; Merla, A. Improving Human–Robot Interaction by Enhancing NAO Robot Awareness of Human Facial Expression. *Sensors* **2021**, *21*, 6438. [[CrossRef](#)]
118. Hefter, E.; Perry, C.; Coiro, N.; Parsons, H.; Zhu, S.; Li, C. Development of a Multi-sensor Emotional Response System for Social Robots. In *Interactive Collaborative Robotics*; Springer: Cham, Switzerland, 2021; pp. 88–99.
119. Shan, Y.; Li, S.; Chen, T. Respiratory signal and human stress: Non-contact detection of stress with a low-cost depth sensing camera. *Int. J. Mach. Learn. Cybern.* **2020**, *11*, 1825–1837. [[CrossRef](#)]
120. Gümüslü, E.; Erol Barkana, D.; Köse, H. Emotion recognition using EEG and physiological data for robot-assisted rehabilitation systems. In Proceedings of the Companion Publication of the 2020 International Conference on Multimodal Interaction, Virtual, 25–29 October 2021; pp. 379–387.
121. Mocanu, B.; Tapu, R. Speech Emotion Recognition using GhostVLAD and Sentiment Metric Learning. In Proceedings of the 2021 12th International Symposium on Image and Signal Processing and Analysis (ISPA), Zagreb, Croatia, 13–15 September 2021; pp. 126–130.
122. Hossain, M.Z.; Daskalaki, E.; Brüstle, A.; Desborough, J.; Lueck, C.J.; Suominen, H. The role of machine learning in developing non-magnetic resonance imaging based biomarkers for multiple sclerosis: A systematic review. *BMC Med. Inform. Decis. Mak.* **2022**, *22*, 242. [[CrossRef](#)]
123. Lam, J.S.; Hasan, M.R.; Ahmed, K.A.; Hossain, M.Z. Machine Learning to Diagnose Neurodegenerative Multiple Sclerosis Disease. In *Asian Conference on Intelligent Information and Database Systems*; Springer: Singapore, 2022; pp. 251–262.
124. Deng, J.; Hasan, M.R.; Mahmud, M.; Hasan, M.M.; Ahmed, K.A.; Hossain, M.Z. Diagnosing Autism Spectrum Disorder Using Ensemble 3D-CNN: A Preliminary Study. In Proceedings of the 2022 IEEE International Conference on Image Processing (ICIP), Bordeaux, France, 16–19 October 2022; pp. 3480–3484.
125. Altun, K.; MacLean, K.E. Recognizing affect in human touch of a robot. *Pattern Recognit. Lett.* **2015**, *66*, 31–40. [[CrossRef](#)]
126. Mohammadi, Z.; Frounchi, J.; Amiri, M. Wavelet-based emotion recognition system using EEG signal. *Neural Comput. Appl.* **2017**, *28*, 1985–1990. [[CrossRef](#)]
127. Wiem, M.B.H.; Lachiri, Z. Emotion assessing using valence-arousal evaluation based on peripheral physiological signals and support vector machine. In Proceedings of the 2016 4th International Conference on Control Engineering & Information Technology (CEIT), Hammamet, Tunisia, 16–18 December 2016; pp. 1–5.
128. Wiem, M.B.H.; Lachiri, Z. Emotion recognition system based on physiological signals with Raspberry Pi III implementation. In Proceedings of the 2017 3rd International Conference on Frontiers of Signal Processing (ICFSP), Paris, France, 6–8 September 2017; pp. 20–24.
129. Wiem, M.B.H.; Lachiri, Z. Emotion sensing from physiological signals using three defined areas in arousal-valence model. In Proceedings of the 2017 International Conference on Control, Automation and Diagnosis (ICCAD), Hammamet, Tunisia, 19–21 January 2017; pp. 219–223.
130. Yonezawa, T.; Mase, H.; Yamazoe, H.; Joe, K. Estimating emotion of user via communicative stuffed-toy device with pressure sensors using fuzzy reasoning. In Proceedings of the 2017 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), Jeju, Korea, 28 June–1 July 2017; pp. 916–921.
131. Alazrai, R.; Homoud, R.; Alwanni, H.; Daoud, M.I. EEG-based emotion recognition using quadratic time-frequency distribution. *Sensors* **2018**, *18*, 2739. [[CrossRef](#)] [[PubMed](#)]
132. Bazgir, O.; Mohammadi, Z.; Habibi, S.A.H. Emotion recognition with machine learning using EEG signals. In Proceedings of the 2018 25th National and 3rd International Iranian Conference on Biomedical Engineering (ICBME), Qom, Iran, 29–30 November 2018; pp. 1–5.
133. Henia, W.M.B.; Lachiri, Z. Emotion classification in arousal-valence dimension using discrete affective keywords tagging. In Proceedings of the 2017 International Conference on Engineering & MIS (ICEMIS), Monastir, Tunisia, 8–10 May 2017; pp. 1–6.
134. Marinoiu, E.; Zafir, M.; Olaru, V.; Sminchiescu, C. 3d human sensing, action and emotion recognition in robot assisted therapy of children with autism. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 2158–2167.
135. Henia, W.M.B.; Lachiri, Z. Multiclass SVM for affect recognition with hardware implementation. In Proceedings of the 2018 15th International Multi-Conference on Systems, Signals & Devices (SSD), Yasmine Hammamet, Tunisia, 19–22 March 2018; pp. 480–485.
136. Salama, E.S.; El-Khoribi, R.A.; Shoman, M.E.; Shalaby, M.A.W. EEG-based emotion recognition using 3D convolutional neural networks. *Int. J. Adv. Comput. Sci. Appl.* **2018**, *9*, 329–337. [[CrossRef](#)]
137. Pandey, P.; Seeja, K. Subject independent emotion recognition from EEG using VMD and deep learning. *J. King Saud Univ.-Comput. Inf. Sci.* **2019**, *34*, 1730–1738. [[CrossRef](#)]

138. Su, Y.; Li, W.; Bi, N.; Lv, Z. Adolescents environmental emotion perception by integrating EEG and eye movements. *Front. Neurobotics* **2019**, *13*, 46. [[CrossRef](#)]
139. Ullah, H.; Uzair, M.; Mahmood, A.; Ullah, M.; Khan, S.D.; Cheikh, F.A. Internal emotion classification using EEG signal with sparse discriminative ensemble. *IEEE Access* **2019**, *7*, 40144–40153. [[CrossRef](#)]
140. Yin, G.; Sun, S.; Zhang, H.; Yu, D.; Li, C.; Zhang, K.; Zou, N. User Independent Emotion Recognition with Residual Signal-Image Network. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 22–25 September 2019; pp. 3277–3281.
141. Algarni, M.; Saeed, F.; Al-Hadhrani, T.; Ghabban, F.; Al-Sarem, M. Deep Learning-Based Approach for Emotion Recognition Using Electroencephalography (EEG) Signals Using Bi-Directional Long Short-Term Memory (Bi-LSTM). *Sensors* **2022**, *22*, 2976. [[CrossRef](#)]
142. Panahi, F.; Rashidi, S.; Sheikhan, A. Application of fractional Fourier transform in feature extraction from ELECTROCARDIOGRAM and GALVANIC SKIN RESPONSE for emotion recognition. *Biomed. Signal Process. Control.* **2021**, *69*, 102863. [[CrossRef](#)]
143. Raschka, S. When Does Deep Learning Work Better Than SVMs or Random Forests®? 2016. Available online: <https://www.kdnuggets.com/2016/04/deep-learning-vs-svm-random-forest.html> (accessed on 19 December 2022).
144. Ugail, H.; Al-dahoud, A. A genuine smile is indeed in the eyes—The computer aided non-invasive analysis of the exact weight distribution of human smiles across the face. *Adv. Eng. Inform.* **2019**, *42*, 100967. [[CrossRef](#)]
145. Heaven, D. Why faces don't always tell the truth about feelings. *Nature* **2020**, *578*, 502–505. [[CrossRef](#)] [[PubMed](#)]
146. Ball, T.; Kern, M.; Mutschler, I.; Aertsen, A.; Schulze-Bonhage, A. Signal quality of simultaneously recorded invasive and non-invasive EEG. *Neuroimage* **2009**, *46*, 708–716. [[CrossRef](#)] [[PubMed](#)]
147. Sokolov, S. Neural Network Based Multimodal Emotion Estimation. *ICASSP 2018* **2018**, *12*, 4–7.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.