

Systematic Review

Datasets for Automated Affect and Emotion Recognition from Cardiovascular Signals Using Artificial Intelligence—A Systematic Review

Paweł Jemioło ^{1,*} , Dawid Storman ² , Maria Mamica ¹, Mateusz Szymkowski ¹, Wioletta Żabicka ³, Magdalena Wojtaszek-Główka ³ and Antoni Ligeża ^{1,*} 

¹ AGH University of Science and Technology, Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering, al. A. Mickiewicza 30, 30-059 Krakow, Poland; maria.anna.mamica@gmail.com (M.M.); mateuszszykowski97@gmail.com (M.S.)

² Chair of Epidemiology and Preventive Medicine, Department of Hygiene and Dietetics, Jagiellonian University Medical College, ul. M. Kopernika 7, 31-034 Krakow, Poland; dawid.storman@doctoral.uj.edu.pl

³ Students' Scientific Research Group of Systematic Reviews, Jagiellonian University Medical College, ul. M. Kopernika 7, 31-034 Krakow, Poland; wiola.zabicka@gmail.com (W.Ż.); wojtaszekglowka@gmail.com (M.W.-G.)

* Correspondence: pawljmlo@agh.edu.pl (P.J.); ligeza@agh.edu.pl (A.L.)

Simple Summary: We reviewed the literature on the publicly available datasets used to automatically recognise emotion and affect using artificial intelligence (AI) techniques. We were particularly interested in databases with cardiovascular (CV) data. Additionally, we assessed the quality of the included papers. We searched the sources until 31 August 2020. Each step of identification was carried out independently by two reviewers to maintain the credibility of our review. In case of disagreement, we discussed them. Each action was first planned and described in a protocol that we posted on the Open Science Framework (OSF) platform. We selected 18 works focused on providing datasets of CV signals for automated affect and emotion recognition. In total, data for 812 participants aged 17 to 47 were analysed. The most frequently recorded signal was electrocardiography. The authors most often used video stimulation. Noticeably, we did not find much necessary information in many of the works, resulting in mainly low quality among included papers. Researchers in this field should focus more on how they carry out experiments.

Abstract: Our review aimed to assess the current state and quality of publicly available datasets used for automated affect and emotion recognition (AAER) with artificial intelligence (AI), and emphasising cardiovascular (CV) signals. The quality of such datasets is essential to create replicable systems for future work to grow. We investigated nine sources up to 31 August 2020, using a developed search strategy, including studies considering the use of AI in AAER based on CV signals. Two independent reviewers performed the screening of identified records, full-text assessment, data extraction, and credibility. All discrepancies were resolved by discussion. We descriptively synthesised the results and assessed their credibility. The protocol was registered on the Open Science Framework (OSF) platform. Eighteen records out of 195 were selected from 4649 records, focusing on datasets containing CV signals for AAER. Included papers analysed and shared data of 812 participants aged 17 to 47. Electrocardiography was the most explored signal (83.33% of datasets). Authors utilised video stimulation most frequently (52.38% of experiments). Despite these results, much information was not reported by researchers. The quality of the analysed papers was mainly low. Researchers in the field should concentrate more on methodology.

Keywords: systematic review; cardiovascular; artificial intelligence; dataset; automated emotion recognition; automated affect recognition; affective computing



Citation: Jemioło, P.; Storman, D.; Mamica, M.; Szymkowski, M.; Żabicka, W.; Wojtaszek-Główka, M.; Ligeża, A. Datasets for Automated Affect and Emotion Recognition from Cardiovascular Signals Using Artificial Intelligence—A Systematic Review. *Sensors* **2022**, *22*, 2538. <https://doi.org/10.3390/s22072538>

Academic Editor: Raffaele Gravina

Received: 16 February 2022

Accepted: 22 March 2022

Published: 25 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



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

Facilitating access to databases seems to be an essential matter in the field of machine learning (ML). Publicly available, reliable datasets could drive research forward, making it unnecessary to re-run similar yet complicated experiments in order to obtain sufficient data. Credible work relies on proper arrangement, validation, adjustment, and fairness in artificial intelligence (AI) [1,2].

Moreover, sufficient descriptions of the scientific methods in AI are a constant challenge. It seems to be particularly valid in automated affect and emotion recognition (AAER) studies, which fall under the field of human–computer interaction (HCI), linking psychology, computer science, and biomedical engineering. As human emotions affect multiple channels, research on this topic is being conducted based on speech, facial expressions, gestures or physiological signals, which became exceptionally popular in the last decade [3].

Increasing interest in the field, among others, comes from broad application prospects. Recent studies point out the potential usage of emotion recognition techniques in medical fields, public security, traffic safety, housekeeping, and related service fields [4].

The topic is extensive, as it covers both data acquisition and computation. A typical experiment in AAER involves several steps [5]. Firstly, the researchers need to adopt a specific perspective on the field, as many exist that consider the universality [6,7] of emotions or their structure [8]. The theoretical approach imposes an understanding of emotions, selections of material used for stimulation, and interpretations. However, the general structure of elicitation experiments that are carried out to gather the data from human participants remains stable [9]. To evoke emotions, passive (e.g., video, music, or pictures presentation) or active stimulation (e.g., game playing, interaction with virtual reality, or conversation) is used [5]. Eliciting material may have different lengths, types, and quantities. After the stimulation phase, the subjects are asked how they felt. Several validated instruments enable it, e.g., Self-Assessment Manikin (SAM) [10].

During the stimulation phase, subjects are connected to measuring devices. Researchers use dedicated hardware [11–13] and experiment with smartphone [14] or wearable [15,16] technologies, especially with CV signals [17–19]. Among others, gathered data include physiological signals [20] (e.g., heart, skin, brain, respiratory system, and eye work), facial expressions [21], and speech [22,23]. Typically, several signals are collected in order to improve the accuracy of the AI system used for AAER [24].

Next, the recognition phase begins. It involves data preprocessing, classification or regression, and finally, validation [5]. Due to its flexibility resulting in, e.g., reduced data preprocessing time [5], deep learning (DL) techniques are widely adopted [25–27], along with classical approaches in AAER [28–31].

As the data collection process in experiments within this field is complex and multi-stage, the problems may occur on many levels. It is thus crucial to plan the experiment and report upon it in adequate detail [32].

The replicability crisis in both psychology and computer science also affects studies on AAER [5,33,34]. Poor methodological conduct often makes it impossible for existing research to be replicated or reproduced. Even in renowned and well-established research that dictates the social order, the phenomenon is widely present [35,36].

Datasets collected inadequately might contribute to lowering the credibility of emerging research (influencing model development by introducing undesirable biases) and waste of time and resources. This issue has been widely discussed before and is known as the *garbage in, garbage out problem* [37,38]. Avoiding bias and proper validation of experiments are crucial to eliminating it [32].

Promisingly, publishing source codes and data is becoming a desirable standard in computer science [39–42]. Journal initiatives [43,44] on the topic emphasise the importance of computational research reproducibility and promote open research. In turn, preregistration of the research plan, taking into account the hypotheses and defining step-by-step the methodology allows for improving the quality of the research and its reproducibility from a psychological perspective [45,46].

To create a reliable model presenting a high degree of emotion or affect recognition precision, it is relevant to limit external and internal factors potentially confounding the collected measurements [47,48].

The confounding effect of incomplete control may arise from any stage of the study. For instance, subjects with somatic disorders might affect measures of features, mood disorders, or alexithymia, which is estimated to affect 13% of the population [49].

Each stage of an experiment leading to AAER should be repeatable and standardised among subjects. AAER concerns stimuli presentation, assessment of elicited emotions by the subject, collection of physiological parameters, and laboratory environment, including the presence of experimenter and individual factors [5].

While measuring emotional and affective responses in the laboratory environment using objective methods reduces the risk of self-reported bias, the risk of contextual non-integrity remains. This creates the need to document all the contextual environmental aspects that could influence the measurement [50].

Along with the pervasiveness of wearable devices available to register user psychological parameters during daily activities, AAER is reached [51,52]. Wearable devices are proven to measure efficiently CV signals while being offered at low prices [53,54]. However, the challenge remains to design credible ML models able to deal with the broad spectrum of possible emotions and lack of universality in this category among cultures [7].

Studies on ubiquitous computing are growing in number [55–57]. Due to the constraints of time and human resources, all these results could not be read. Therefore, creating summaries along with the analysis of evidence is now necessary [58]. Describing the data together with a critical appraisal helps to determine, for example, the actual accuracy of the methods and to highlight those articles whose results are derived from a high-quality methodological process. The selection of studies answering a similar research question may be chaotic, purposeful, or systematic [59]. The latter method reduces the risk of researchers steering conclusions, as it follows restrictive, transparent criteria [32,60,61].

Because of the above and since previous similar studies on AAER were of weak reliability [5], we decided to present a systematic review on the topic, corresponding to approved standards, to limit the risk of bias (RoB). We review public datasets available for AAER with the use of AI, utilising physiological modalities as an input with the focus on CV signals.

This paper is a part of the project on a systematic review of studies focused on AAER from CV signals with AI methods. For more details, see the protocol [62] and our previous conference paper [63].

Research Questions

1. What are the datasets used for AAER from CV signals with AI techniques?
2. What are the CV signals most often gathered in datasets for AAER?
3. What were other signals are collected in analysed papers?
4. What are the characteristics of the population in included studies?
5. What instruments were used to assess emotion and affect in included papers?
6. What confounders were taken into account in analysed papers?
7. What devices were used to collect the signals in included studies?
8. What stimuli are most often used for preparing datasets for AAER from CV signals?
9. What are the characteristics of investigated stimuli?
10. What is the credibility of included studies?

2. Methods

2.1. Eligibility Criteria, Protocol

Papers in which more than half of the sample constitutes a specific population (e.g., children or people with illness) were excluded. All experiments needed to be carried out in laboratory settings. We considered any type of publication to be eligible in which CV signals and AI methods were used for AAER. The primary focus of our whole project [63]

was the performance of these computer programs (e.g., specificity, sensitivity, accuracy). For this focused systematic review, we imposed additional inclusion criteria, namely public availability of the data.

Due to double referencing, some of the references were overlapping. These were post-conference books and full proceedings. We excluded them as they contained little information about specific chapters. Nevertheless, we did not reject these particular sections. We excluded introductions to Special Issues in a journal or section, letters to editors, reviews, post-conference books, full proceedings (but not qualified papers), and case studies.

The review protocol was published on the Open Science Framework (OSF) [64] and then registered there [62] on 18 March 2021. All additional information about methods can be found in the protocol.

2.2. Search Methods

We searched article databases (MEDLINE, Web of Science, dblp, EMBASE, Scopus, IEEE, Cochrane Library) and preprint databases (medRxiv, arXiv). The complete search was done on 31 August 2020.

To develop the MEDLINE strategy (see protocol on OSF [62]), we combined MeSH (controlled vocabulary) and free-text words related to AAER, CV signals, and AI. Then, these strings were translated for other sources utilised in the search. We adopted no date or language restrictions.

Additionally, we screened full texts of included papers for otherwise not identified studies. We included them in further steps of identification.

2.3. Definitions

We used the following definitions. AAER [65,66] refers to finding patterns with specific signals (e.g., behavioural, physiological) consistent with detected states. AI refers to software able to perform tasks as accurately as intelligent beings (e.g., humans) [67]. DL refers to the architecture of neural networks comprising at least two hidden layers [68]. Performance metrics, which refer to a mathematical evaluation of model predictions with ground truth [69]. CV signals refer to an electrocardiogram (ECG), pulse oximetry (POX), heart rate (HR), intracranial pressure (ICP), pulse pressure variation (PPV), heart rate variability (HRV), photoplethysmogram (PPG), blood volume pressure (BVP), and arterial blood pressure (ABP) [53,70].

2.4. Data Collection

EndNote (Claritive Analytics®) and Rayyan [71] were utilised for deduplication of identified references. P.J., D.S., M.S., and M.M. used the Rayyan [71] application to screen the remaining references independently. Subsequently, full texts were assessed separately by P.J., D.S., M.S., and M.M. for meeting inclusion criteria.

P.J., D.S., M.S., M.M., W.Ż., and M.W.G. collected all necessary data independently using a pre-specified extraction form. We gathered bibliographic data (e.g., year, journal name) and information about authors, funding, and conflicts of interest. We also focused on population, models, and outcomes—AI methods and additional analyses, e.g., interpretability, as specified in the protocol (see OSF [62]).

Pilot exercises were conducted before each phase, namely screening of abstracts and titles, full text evaluation, and extraction of the data. By doing so, we aimed at improving the sense of understanding among the reviewers. When discrepancies occurred (at each step of data identification), they were resolved via discussion.

2.5. Quality Assessment

The methodological credibility of included studies was assessed using a tool developed by our team (see Appendix C). The method was based on well-grounded techniques, namely Quality Assessment of Diagnostic Accuracy Studies (QUADAS) [72], Prediction model Risk Of Bias ASsessment Tool (PROBAST) [73], and an instrument provided by

Benton et al. [74] as it was dedicated to the same study design as included by us. The process of evaluation was preceded by pilot exercises. We rated RoB independently in pairs (P.J., D.S., M.S., M.M., W.Ž., and M.W.G.). Discussion resolved all discrepancies.

The utilised tool constituted of eight questions (items):

1. Was the sample size pre-specified?
2. Were eligibility criteria for the experiment provided?
3. Were all inclusions and exclusions of the study participants appropriate?
4. Was the measurement of the exposition clearly stated?
5. Was the measurement of the outcome clearly stated?
6. Did all participants receive a reference standard?
7. Did participants receive the same reference standard?
8. Were the confounders measured?

Items were assessed using a three-point scale with the following answers: yes/partial yes, no/partial no, and not reported resulting in high, low, or unclear RoB. For more details, see Appendix C.

2.6. Analyses

We concentrate on descriptive synthesis regarding characteristics of populations and collected datasets, i.e., stimuli, signals, devices, emotions, and affect. We also present results regarding credibility.

The quantitative summary with sensitivity, heterogeneity, and subgroup analysis of all papers is not the purpose of this focused review. For more details, please refer to the protocol [62] and other papers from the project [63].

3. Results

From 4649 records, we identified 195 studies that met our eligibility criteria. Then, we selected a sub-sample of 18 papers. Each paper provides one validated, publicly available dataset, including CV signals with labels regarding emotions or affect.

Names of datasets described in included papers are as follows: Database for Emotion Analysis using Physiological signals (DEAP) [75]; Multimodal Analysis of Human Nonverbal Behaviour in real-world settings–Human-Computer Interaction (MAHNOB-HCI) [76]; MEG-based multimodal database for DEcoding Affective physiological responses (DECAF) [77]; a dataset for Affect, personality and Mood research on Individuals and GrOuPs (AMIGOS) [78]; a multimodal databAse for impliCit pERsonaliTy and Affect recognitiON using commercial physiological sensors (ASCERTAIN) [79]; Augsburg Database of Biosignal 4 (AuDB-4) [80]; Emotion Recognition Smartwatch (ERS) [81]; IT Multimodal Dataset for Emotion Recognition (ITMDER) [82]; Database for Affective Gaming (DAG) [83]; Quality Adaptive Multimodal Affect recognition system for user-centric multimedia indexing (QAMAF) [84]; Virtual Reality Affective Dataset (VRAD) [85]; NEME [86]; WEearable Stress and Affect Detection (WESAD) [87]; a Multi-modal Physiological Emotion Database for discrete emotion recognition (MPED) [88]; database of multimodal (Face, Body gesture, Voice and Physiological signals) recordings (emoFBVP) [89]; a database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices DREAMER [90]; Multi-subject Affective Physiological Database (MAPD) [91]; and MazeBall Dataset (MD) [92].

Supplementary File S1 (OSF [64]) and Appendices A and B contain the list of all included studies, the subgroup of datasets analysed in this review, and the excluded studies with reasons, respectively. The remaining included studies are considered in other articles from the project [63]. The flow of our study is presented in Figure 1. Our reporting is consistent with Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines with diagnostic test accuracy (DTA) extension [93].

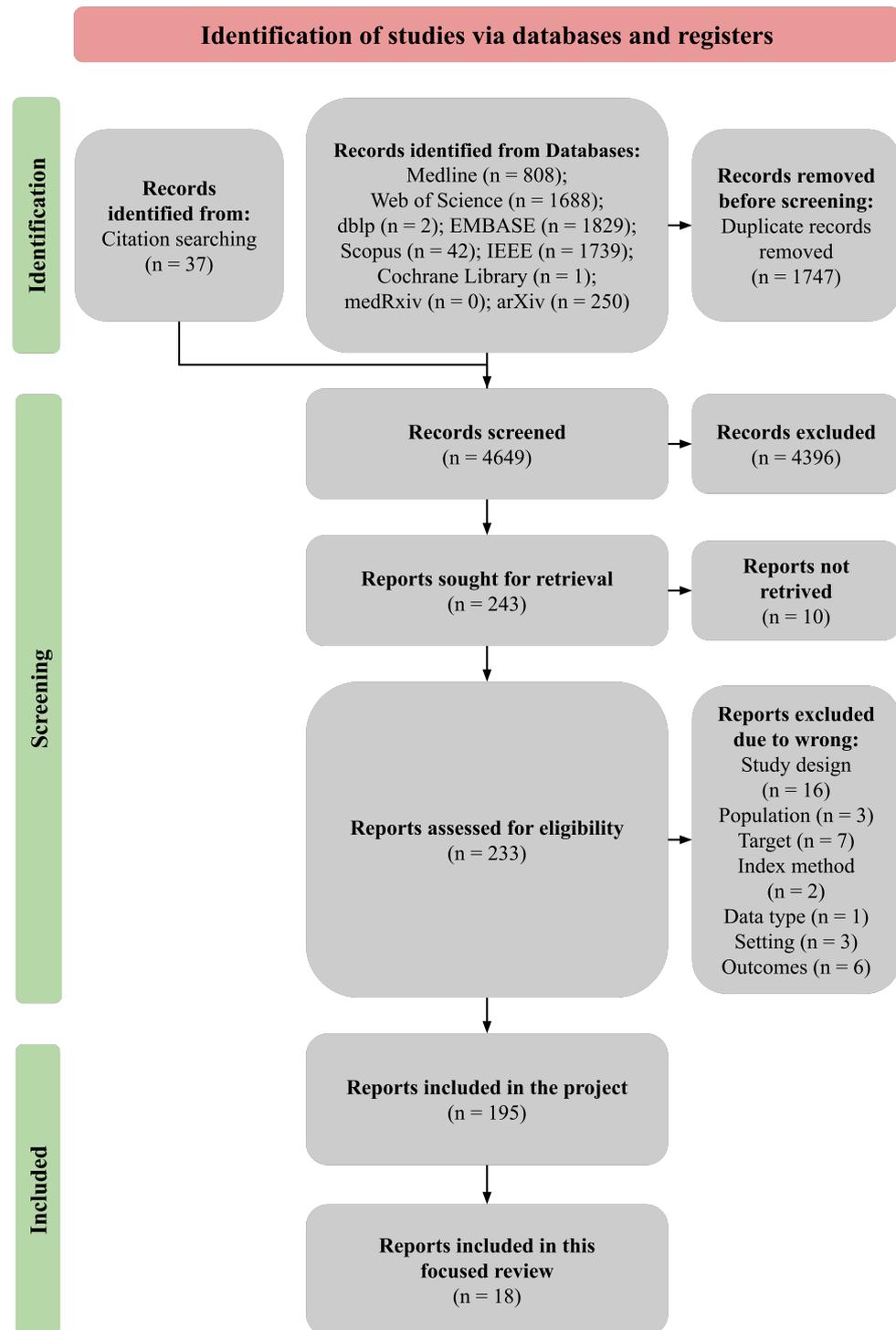


Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) study flow diagram [63].

3.1. Included Studies

Included studies were published mainly in scientific journals (66.67% of papers, mean Impact Factor = 7.01) [75–81,85,86,88,90,92]. The most popular was *IEEE Transactions on*

Affective Computing. Out of all the authors, 27.78% did not report on funding [77,82,87,89,91], while 88.89% did not inform about competing interests [75–80,82–84,86–92].

None of the studies provided source code of executed analyses, while only one study (5.56%) reported registering protocol [92].

3.2. Experiments

The total number of elicitation experiments was 21, presented in 18 papers. Three of the studies (16.67%) carried out two trials each [77,78,81]. It was found that 76.19% of experiments were conducted using passive stimulation solely [75–82,84,86,88,90,91], while 19.05% used only active elicitation (e.g., video games) [83,85,89,92]. One experiment (described in Schmidt et al.'s paper [87]) used both passive (video) and active stimulation (meditation and Trier Social Stress Test (TSST)). The essential characteristics of experiments regarding stimuli are presented in Table 1.

Table 1. Characteristics of stimuli in 18 included studies (21 experiments).

Variable (No. of Experiments Available for Calculations)	No (%) Mean (Range)
Type of stimuli (21)	
Video (music, movie, ads)	11 (52.38)
Audio (music excerpts)	4 (19.05)
Game (FIFA 2016, Maze-Ball)	2 (9.52)
Virtual Reality (videos, scenes)	2 (9.52)
Self elicitation (actors)	1 (4.76)
Mixed (TSST ¹ , video and meditation in one experiment)	1 (4.76)
Length of stimuli [seconds] (17)	304.60 (32–1200)
No. of stimuli in dataset (20)	27.70 (4–144)
No. of elicited emotions [classes] (18)	6.06 (3–23)

¹ Trier Social Stress Test.

Most of the experiments did not use stimuli from validated databases (71.43%, e.g., FIFA 2016, YouTube videos) [75,76,78,80,81,83–89,91,92], whereas public sources (e.g., DEAP, DECAF) accounted for 23.81% [77,79,81,90]. Pinto [82] used both forms. In 47.62% of experiments, the justification for the choice of the database was not reported [77,79,81,83,84,86,87,89,91,92]. Pinto [82] partially reported on it. In 42.85% of experiments, validation was provided by conducting a pilot study or preliminary classification by researchers [75–78,81,85,88,90]. Stimuli were described by authors most frequently in terms of valence (52.38% of experiments) [75–78,80,82,85,86,90,92], arousal (52.38%) [75–78,80,82,85,86,90,92], and discrete emotional tags (38.10%) [76,77,81,88,89,91]. Four experiments [77,79,83,84] did not report on it at all.

The presence of diseases or disorders was the most often controlled factor in participants (61.90% of experiments) [75,76,78,81,82,85–88,90,91]. The mood was controlled using the Positive and Negative Affect Schedule (PANAS) tool in two experiments (9.52%) [78]. In only one experiment, the authors checked if the participants were able to recognise emotions or affective states correctly [85]. In examining factors controlled in the laboratory, the most frequent was found to be brightness (33.33% of experiments) [75,77,78,90,91], followed by volume (28.57%) [78,80,82,84,91], presentation of stimuli (14.29%) [75,80,85], the comfort of participants (9.52%) [77,86], and time of the day (9.52%) [81]. In 23.81% of experiments, it was not reported which factors were controlled [77,80,83,84,92]. Additionally, in four experiments (19.05%), personality was measured in participants using the Big Five Personality Test [78,79] or the Eysenck Personality Questionnaire (EPQ) [91].

In the assessment of emotions and affect by participants, the most prevalent instruments used were: SAM (38.10% of experiments) [75,78,82,85,87,88,90] for valence, arousal and dominance, selecting a discrete emotion from the provided list (23.81%) [76,78,83,89], and PANAS (19.08%) [81,87,91].

3.3. Signals and Devices

Table 2 summarises applied devices and recorded CV signals. Apart from CV signals, electrodermal activity (EDA) [75,76,78–80,82–84,87,88,91,92] is available in 66.67% of datasets. Thus, these are the most prevalent data. The next most common are face video [75–79,83,84,89,91], electroencephalography (EEG) [75,76,78,79,84,85,88,90], respiration [75,76,80,82,83,88], and electromyography (EMG) [75,77,80,83,87] in 50%, 44.44%, 33.33%, and 27.78% of datasets, respectively. The remaining signals include, e.g., magnetoencephalography (MEG), gyroscope, accelerometer, and audio. There is only one dataset that focuses solely on CV signals [86]. The authors recognise devices used for recording CV signals as wearable in 38.89% of studies [78,79,84,85,87,89,90].

Table 2. Characteristics of devices and signals in 18 included studies (21 experiments).

Variable (No. of Datasets Available for Calculations) ¹	No (%) Mean (Range)
Used devices (16)	
Shimmer 2R	3 (18.75)
BIOPAC MP150	3 (18.75)
Biosemi ActiveTwo	2 (12.50)
NeXus-10	1 (6.25)
ProComp Infiniti	1 (6.25)
BIOPAC BioNomadix	1 (6.25)
BIItalino	1 (6.25)
RespiBAN Professional	1 (6.25)
B-Alert ×10	1 (6.25)
Empatica E4	1 (6.25)
Polar H7	1 (6.25)
Zephyr BioHarness	1 (6.25)
IOM Biofeedback	1 (6.25)
CV ² signals recorded (18)	
ECG ³	15 (83.33)
HR ⁴	3 (16.67)
BVP ⁵	3 (16.67)
PPG ⁶	1 (5.56)
Sampling frequency [Hz] (12)	543.31 (32–2048)
Length of baseline recording [seconds] (7)	292.14 (5–1200)

¹ some studies used more than one device or cardiovascular signal; ² cardiovascular; ³ electrocardiogram; ⁴ heart rate; ⁵ blood volume pressure; ⁶ photoplethysmogram.

3.4. Validation

As we included only validated datasets in this analysis, all of the papers explored AAER with AI. Out of all the papers, 55.56% [75–78,82,85,86,90–92] conducted experiments with only one type of ML algorithm, while the rest explored more methods. In total, the data were validated using AI methods 33 times. Support vector machine was used most frequently (33.33%) [76,77,79,82–86,89–91]. Naive Bayes (NB) [75,78,79,84], random forest (RF) [81,83,87,91], and DL [88,89] were the second most explored techniques (12.12% each).

The authors classified 61 different discrete states in total. The most commonly classified one was sadness [81,86,89,91], occurring in 22.22% of papers. The following states were examined in two papers each (11.11%): fear [89,91], anger [89,91], amusement [87,91], anxiety [89,92], boredom [89,92], happiness [81,89], and neutral state [88,89]. Additionally, the authors used affect space in 12 (66.67%) papers [75–80,82–86,90].

All of the datasets were validated in classification experiments. Authors of only two datasets (11.11%) [86,89] compared their results with other publicly available data.

3.5. Population

The total number of analysed people was 916, with a mean number of 43.62 participants and a range from 3 [80] to 250 [91]. However, due to, e.g., missing data, the datasets contain complete information for only 812 of them.

The remaining characteristics of the population are shown in Table 3. five experiments (23.81%) were approved by the ethics committee [81,85,88,90]. Participant consent was obtained in 15 experiments (71.43%) [75–79,82,85–91]. Only one experiment ensured the privacy (by anonymisation) of participants [90].

Table 3. Characteristics of population in 18 included studies (21 experiments).

Variable (No. of Experiments Available for Calculations)	No. (%) Mean (Range)
Participating people (21)	916 43.62 (3–250)
Eligible people (20)	812 40.60 (3–250)
Age (18)	23.8 (17–47)
Percentage of females (16)	45.13 (0–86)
Ethnicity (4)	
Chinese	2 (9.52)
European	2 (9.52)

3.6. Credibility

The general RoB was analysed in two scenarios—with or without the first item of the proposed tool (see Section 2.5). We excluded the first question in the second condition because none of the included studies reported on pre-specification of sample size.

Of all the studies, 77.78% were of low quality in both scenarios, whereas 22.22% [78, 85,86,90] and 11.11% [86,90] were of unclear quality in the first and second conditions, respectively. Two studies [78,85] were of high quality according to the latter scenario. The RoB across all RoB items is presented in Figure 2. The reference standard was provided for all participants in the same way in 16 studies (88.89%) [75–79,81–91]. For participants from 14 studies (77.78%) [75–79,81–86,89–91], the same reference standard was given. Thirteen studies (72.22%) [75,76,78,79,81,82,85–90,92] did not show any flaws in terms of providing eligibility criteria for the experiment. These are the most satisfied questions.

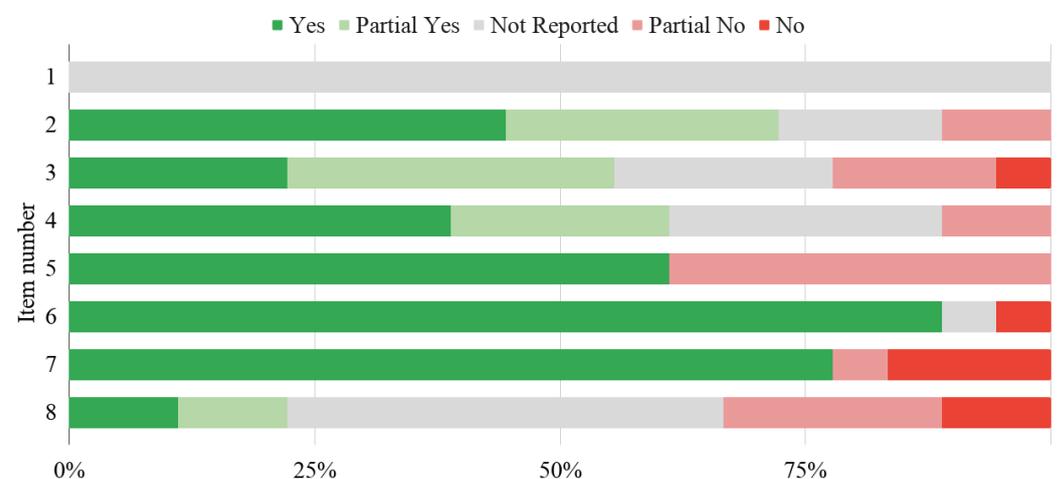


Figure 2. Risk of bias (RoB) in included studies.

However, 14 papers (77.78%) [75–77,79,81–84,86,87,89–92] did not control confounders or did not report it. Measurement of exposition [80,83,84,86,87,89,91] and outcomes [76,77,79,80,83,84,89] was flawed, or authors did not mention it, in seven studies (38.89%).

All ratings are presented in Table 4. Among them, the most frequent was *yes*, marked in 43.06% of cases. However, the second most prevalent was *not reported*, which was assessed 27.08% times.

Table 4. Risk of Bias (RoB) among 18 included studies.

Study ID	RoB Item ¹								Overall Quality ²	
	1	2	3	4	5	6	7	8	Scenario 1	Scenario 2
[75]	NR	PY	PY	Y	Y	Y	Y	PN	Low	Low
[76]	NR	PY	PY	Y	PN	Y	Y	NR	Low	Low
[77]	NR	NR	NR	Y	PN	Y	Y	PN	Low	Low
[78]	NR	Y	PY	Y	Y	Y	Y	PY	Unclear	High
[79]	NR	PY	PN	PY	PN	Y	Y	N	Low	Low
[80]	NR	PN	N	PN	PN	N	N	PY	Low	Low
[81]	NR	Y	PN	PY	Y	Y	Y	PN	Low	Low
[82]	NR	Y	Y	PY	Y	Y	Y	PN	Low	Low
[83]	NR	NR	NR	PN	PN	Y	Y	NR	Low	Low
[84]	NR	NR	NR	NR	PN	Y	Y	N	Low	Low
[85]	NR	Y	Y	Y	Y	Y	Y	Y	Unclear	High
[86]	NR	PY	PY	NR	Y	Y	Y	NR	Unclear	Unclear
[87]	NR	Y	Y	NR	Y	Y	N	NR	Low	Low
[88]	NR	Y	Y	Y	Y	Y	PN	Y	Low	Low
[89]	NR	Y	PN	NR	PN	Y	Y	NR	Low	Low
[90]	NR	PY	PY	PY	Y	Y	Y	NR	Unclear	Unclear
[91]	NR	PN	PY	NR	Y	Y	Y	NR	Low	Low
[92]	NR	Y	NR	Y	Y	NR	N	NR	Low	Low

¹ Y, PY, NR, PN, N stands for yes, partial yes, not reported, partial no, no; ² for more details, see Appendix C.

3.7. Additional Analyses

Please refer to the protocol [62] and our other papers [63] from the project on AAER from CV signals with AI methods for additional analyses.

4. Discussion

The paper search conducted in this study revealed that there are 18 publicly available validated datasets for AAER from CV signals. The methodological credibility assessment showed that only two studies are of high quality, suggesting a significant need for developing good scientific practices.

Furthermore, none of the studies provided a source code used for the validation experiments. It opens a discussion on replicability, which we are witnessing in science nowadays [5]. Experiments in included papers were conducted on small samples. The number of participants exceeded one hundred only in one study.

What is more, the subjects' background information was poorly described. Only four studies established that the participants were either Chinese or European. According to Wierzbicka [94,95], the history behind a person (and the language he or she speaks) may play a crucial role in the emotional states they experience and thus should be controlled. Feldman later disseminated this belief in her approach [7].

Another bothering aspect of the analysis is that an ethical commission approved experiments described by only four papers, and only one study mentioned ensuring the privacy of participants. It lights up red flags in terms of maintaining ethical standards

or suggests negligence of reporting crucial information. Authors of experimental studies should more carefully examine this aspect.

Additionally, the authors either selectively controlled the influence of potential confounders or did not do so at all. Various CV diseases, mental disorders [49], and participants' moods and personalities may affect AAER from physiological signals [78]. Therefore, we believe authors should include such information.

The problem in assessing quality in systematic reviews is about distinguishing how much the authors did not take care of the methodological regime and how much they did not report the details of the research process [60]. Therefore, it is recommended that when submitting an article to the journals' editorial office, the authors fill in a checklist and mark the exact place where they have included the minimum necessary descriptions of the operation process [32,96].

On the other hand, we observed great diversity in the choice of physiological signals, stimuli type and length. What is more, 38.89% of the studies used wearable devices to perform measurements. Considering the increasing popularity and facility of these instruments [78], it gives the excellent potential for future adoption of proposed methods in real-life scenarios. Thanks to recent advances in the field of sensors technology, such devices are well-suited for daily usage. They do not require complicated instalments, are comfortable to wear, and are easy to use [97]. However, one should remember that there are still many limitations standing in the way of the wider use of wearable devices in AAER. First of all, the quality of physiological signals is still noticeably lower than that of medical-level equipment [98]. What is more, the data gathered by such instruments in non-laboratory settings are often flawed, with noise coming from motion or misplacement [99].

Similarly to our study, the CV databases were also explored by Merone et al. [100]. The authors investigated 12 datasets with the inclusion criteria of having an ECG signal. In addition, they analysed included sets in terms of many parameters, e.g., the number of ECG channels and electrodes type. However, they did not primarily focus on emotions or affect. They included only one paper [101] covering this scope, which we did not consider eligible for inclusion as it did not meet the criteria. Since datasets including CV signals are still unexplored, we cannot discuss our results with other authors. Furthermore, Hong et al. [102] analysed ECG data systematically using DL. Still, they identified only one study about AAER [103], but it was not in their primary interest, so they did not describe it in detail.

However, in the literature, there are plenty of reviews (systematic and not) focusing on AAER from multiple signals or focusing on specific ones, e.g., EEG [104–106], or covering multiple modalities [53,107–111]. Still, their quality has been thoroughly criticised in our recently published umbrella review [5].

In line with these results, in the current literature, we found a shortage of highly credible and methodologically reliable publications and thus datasets that could form the basis of further AI research. This review shows a need to create guideline-compliant datasets with a transparent, fully reported methodology and limited RoB.

Models able to accurately recognise emotions using physiological parameters can contribute to the development of many disciplines. They create the possibility of reaching more advanced levels of HCI, where a computer (or system, in general) can modify its behaviour depending on the identified interlocutor's state and choose the reaction closest to natural social schemes [112].

While using wearable devices, users might be supported in maintaining a psychological and healthy life balance, e.g., by identifying sources of stress, anxiety, or tension during their everyday activities and receiving feedback about their organisms reactions and resources [113]. Furthermore, assessments made on the basis of their CV signals can be used to investigate the impact of different emotional and affective states on the risk of developing CV diseases [114].

Well-validated AI models can significantly support research in the field of health and medical sciences and emotion theory by facilitating the simple, quick, and more matter-of-

fact evaluation of emotions and other states and, therefore, reducing the RoB resulting from participants' incorrect reporting.

Among the implications of our study, we should first include the recommendation to incorporate current, reliable guidelines and standards in the methodology development process and use quality assessment and reporting tools, as this translates into more reliable data, which may result in developing better recognition models [32]. For primary studies, we suggest following the proposed checklist for RoB (see Appendix C) or other available tools, e.g., [32].

Strengths and Limitations

The performed review has high standards [32,60,61,115]. The research question was precisely defined. We utilised multiple resources for collecting studies mentioned in Section 2. Inclusion and exclusion criteria were firstly discussed and recorded. Researchers who participated in this review have knowledge in multiple disciplines: computer science, psychology, HCI, medicine, and methodology. To ensure transparency, we provide all necessary information in the Appendices and Supplements with a permanent DOI [64].

On the other hand, we did not search any Chinese databases. Considering the growing amount of evidence in this language, we might not have considered a large amount of evidence and thus weakened our conclusions. Moreover, the use of the search strategy itself and the stages of identifying articles based on titles and abstracts may be a limitation. Due to such action, we may miss an extraordinary piece of work that did not meet our criteria due to its original form.

5. Conclusions

This paper systematically reviewed the datasets that include CV signals for AAER with AI methods and assessed their quality.

Due to poor reporting and not following methodological guidelines, the evidence, however, is limited. Nevertheless, according to our review, the most up-to-standards research was proposed by Correa et al. [78] and Marin et al. [85].

In the future, more attention should be put into controlling bias in research to ensure incremental knowledge gain. The quality of papers and reporting needs to be improved in order to propose and develop models that do not introduce biases. Preferably, authors should focus more on methodology and describe procedures thoroughly. We recommend following standardised guidelines of reporting [116].

Our next steps include the synthesis of gathered evidence with other physiological signals. Furthermore, we want to propose our own unbiased dataset for AAER for public use. Based on these data, we plan to improve our affective games [117–119].

Supplementary Materials: The following supporting information can be downloaded at: <https://osf.io/kzj8y/> (accessed on 15 February 2022) in *Sensors* folder: Protocol; Supplementary File S1—list of all included studies.

Author Contributions: Conceptualization, P.J. and D.S.; methodology, P.J. and D.S.; validation, P.J. and D.S.; formal analysis, P.J. and D.S.; investigation, P.J., D.S., M.S., M.M., W.Ž. and M.W.-G.; resources, P.J. and D.S.; data curation, P.J., D.S., M.S., M.M., W.Ž. and M.W.-G.; writing—original draft preparation, P.J., D.S., M.S., M.M., W.Ž. and M.W.-G.; writing—review and editing, P.J., D.S., M.S., M.M., W.Ž. and M.W.-G.; visualization, P.J.; supervision, P.J., A.L.; project administration, P.J.; funding acquisition, A.L. and P.J. All authors have read and agreed to the published version of the manuscript.

Funding: Research project supported by program *Excellence initiative—research university* for the University of Science and Technology. This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: We would like to thank Weronika Dranka for her help in protocol preparation.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

AAER	automated affect and/or emotion recognition
ABP	arterial blood pressure
AI	artificial intelligence
AMIGOS	a dataset for Affect, personality and Mood research on Individuals and GrOuPS
ASCERTAIN	a multimodal databaASe for impliCit pERsonaliTY and Affect recognitIoN using commercial physiological sensors
AuDB-4	AUGsbuRg Database of Biosignal 4
BVP	blood volume pressure
CV	cardiovascular
DAG	Database for Affective Gaming
DECAF	MEG-based multimodal database for DECoDing Affective physiological responses
DEAP	Database for Emotion Analysis using Physiological signals
DL	deep learning
DTA	diagnostic test accuracy
EKG	electrocardiogram
EDA	electrodermal activity
EEG	electroencephalography
EMG	electromyography
emoFBVP	database of multimodal (Face, Body gesture, Voice and Physiological signals) recordings
EPQ	Eysenck Personality Questionnaire
ERS	Emotion Recognition Smartwatch
HCI	human–computer interaction
HR	heart rate
HRV	heart rate variability
ICP	intracranial pressure
ITMDER	IT Multimodal Dataset for Emotion Recognition
MAHNOB-HCI	Multimodal Analysis of Human NONverbal Behaviour in real-world settings-Human-Computer Interaction
MAPD	Multi-subject Affective Physiological Database
MD	Mazeball Dataset
MEG	magnetoencephalography
ML	machine learning
MPED	Multi-modal Physiological Emotion Database for discrete emotion recognition
NB	naive Bayes
OSF	Open Science Framework
PANAS	Positive and Negative Affect Schedule
POX	pulse oximetry
PPG	photoplethysmogram
PPV	pulse pressure variation
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses
PROBAST	Prediction model Risk Of Bias ASsessment Tool
QUADAS	QUality Assessment of Diagnostic Accuracy Studies
RF	random forest
RoB	Risk of Bias
SAM	Self-Assessment Manikin

TSST	Trier Social Stress Test
VRAD	Virtual Reality Affective Dataset
WESAD	WEarable Stress and Affect Detection

Appendix A. Papers Included in This Focused Review

The following papers were included in this focused review: [75–92].

Appendix B. Papers Excluded

Table A1. Excluded Studies.

Study ID	Reason of Excluding
[120]	Wrong study design
[121]	Wrong study design
[122]	Wrong study design
[123]	Wrong study design
[124]	Wrong study design
[125]	Wrong study design
[126]	Wrong study design
[127]	Wrong study design
[128]	Wrong study design
[129]	Wrong study design
[130]	Wrong study design
[131]	Wrong study design
[132]	Wrong study design
[133]	Wrong study design
[134]	Wrong study design
[135]	Wrong study design
[136]	Wrong population
[137]	Wrong population
[138]	Wrong population
[139]	Wrong target
[140]	Wrong target
[141]	Wrong target
[142]	Wrong target
[143]	Wrong target
[144]	Wrong target
[145]	Wrong target
[146]	Wrong index method
[147]	Wrong index method
[148]	Wrong type of data
[149]	Wrong setting
[150]	Wrong setting
[151]	Wrong setting
[152]	Wrong outcomes
[153]	Wrong outcomes
[154]	Wrong outcomes
[155]	Wrong outcomes
[156]	Wrong outcomes
[157]	Wrong outcomes

Appendix C. Risk of Bias Tool

Table A2. Risk of bias tool.

Domain ¹	Review Authors's Judgement	Criteria for Judgement	
Sample [74]	1. Was the sample size prespecified?	Yes/partial yes	The experiment was preceded by calculating the minimum sample size, and the method used was adequate and well-described.
		No/partial no	It is stated that the minimum sample size has not been calculated, or it has been calculated, but no details of the method used are provided.
		Not reported	No sufficient information is provided in this regard.
Sample [74]	2. Were eligibility criteria for the experiment provided?	Yes/partial yes	The criteria for inclusion in the experiment are specified.
		No/partial no	The criteria for inclusion in the experiment were used, however not specified in the article.
		Not reported	No sufficient information is provided in this regard.
Participants [73]	3. Were all inclusions and exclusions of participants appropriate?	Yes/partial yes	The criteria for inclusion and exclusion are relevant to the aim of the study. Conditions that may affect the participant's state or collected physiological signals and ability to recognise emotions were considered, including cardiovascular and mental disorders.
		No/partial no	The established criteria for inclusion and exclusion are irrelevant to the aim of the study.
		Not reported	No sufficient information is provided in this regard.
Measurement [74]	4. Was the measurement of exposition clearly stated?	Yes/partial yes	The selection of stimuli is adequately justified in the context of eliciting emotions, e.g., selection from a standardised database, pilot studies.
		No/partial no	The selection of stimuli was carried out based on inadequate criteria.
		Not reported	No sufficient information is provided in this regard.
Measurement [74]	5. Was the measurement of outcome clearly stated?	Yes/partial yes	The assessment tool used for emotions measurement is described in detail, adequate, and validated.
		No/partial no	The assessment tool used for emotions measurement is not described, or the measurement method is inadequate, or not validated.
		Not reported	No sufficient information is provided in this regard.
Flow and Timing [72]	6. Did all participants receive a reference standard?	Yes/partial yes	Emotions were measured in all participants, and the measurement was performed after each stimulus.
		No/partial no	Not all participants had their emotions measured.
		Not reported	No sufficient information is provided in this regard.
Flow and Timing [72]	7. Did participants receive the same reference standard?	Yes/partial yes	The same assessment standard was used in all participants who had their emotions measured
		No/partial no	A different assessment standard was used in some of the participants to measure their emotions.
		Not reported	No sufficient information is provided in this regard.
Control of confounders [74]	8. Were the confounders measured?	Yes/partial yes	Adequate confounding factors were measured, and relevant justification is provided.
		No/partial no	The control of confounding factors is not justified, or the measured factors are inadequate.
		Not reported	No sufficient information is provided in regard to confounding factors.
Scenario 1: Overall quality (elicitation)	Scenario 2: Overall quality (without judgement of 1. item)	High	All judgements are yes or partial yes.
		Low	At least one judgement is no or partial no.
		Unclear	All judgements are yes or partial yes with at least one not reported.

¹ the specific domain was based on an instrument provided in the reference.

References

1. Hacker, P. Teaching fairness to artificial intelligence: Existing and novel strategies against algorithmic discrimination under EU law. *Common Mark. Law Rev.* **2018**, *55*, 1143–1185.
2. Butterworth, M. The ICO and artificial intelligence: The role of fairness in the GDPR framework. *Comput. Law Secur. Rev.* **2018**, *34*, 257–268. [\[CrossRef\]](#)
3. Fan, X.; Yan, Y.; Wang, X.; Yan, H.; Li, Y.; Xie, L.; Yin, E. Emotion Recognition Measurement based on Physiological Signals. In Proceedings of the 2020 13th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, China, 12–13 December 2020; pp. 81–86.
4. Xia, H.; Wu, J.; Shen, X.; Yang, F. The Application of Artificial Intelligence in Emotion Recognition. In Proceedings of the 2020 International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI), Sanya, China, 4–6 December 2020; pp. 62–65.
5. Jemioło, P.; Storman, D.; Giżycka, B.; Ligeza, A. Emotion elicitation with stimuli datasets in automatic affect recognition studies—Umbrella review. In Proceedings of the IFIP Conference on Human-Computer Interaction, Bari, Italy, 30 August–3 September 2021; Springer: Berlin/Heidelberg, Germany, 2021; pp. 248–269. [\[CrossRef\]](#)
6. Ekman, P.; Friesen, W.V.; O’Sullivan, M.; Chan, A.; Diacyoyanni-Tarlatzis, I.; Heider, K.; Krause, R.; LeCompte, W.A.; Pitcairn, T.; Ricci-Bitti, P.E.; et al. Universals and cultural differences in the judgments of facial expressions of emotion. *J. Personal. Soc. Psychol.* **1987**, *53*, 712. [\[CrossRef\]](#)
7. Barrett, L.F. *How Emotions Are Made: The Secret Life of the Brain*; Houghton Mifflin Harcourt: Boston, MA, USA, 2017.
8. Plutchik, R. A general psychoevolutionary theory of emotion. In *Theories of Emotion*; Elsevier: Amsterdam, The Netherlands, 1980; pp. 3–33.
9. Sarma, P.; Barma, S. Review on Stimuli Presentation for Affect Analysis Based on EEG. *IEEE Access* **2020**, *8*, 51991–52009. [\[CrossRef\]](#)
10. Bradley, M.M.; Lang, P.J. Measuring emotion: The self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* **1994**, *25*, 49–59. [\[CrossRef\]](#)
11. Bandara, D.; Song, S.; Hirshfield, L.; Velipasalar, S. A more complete picture of emotion using electrocardiogram and electrodermal activity to complement cognitive data. In Proceedings of the International Conference on Augmented Cognition, Toronto, ON, Canada, 17–22 July 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 287–298.
12. Nardelli, M.; Greco, A.; Valenza, G.; Lanata, A.; Bailón, R.; Scilingo, E.P. A novel heart rate variability analysis using lagged poincaré plot: A study on hedonic visual elicitation. 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. 2300–2303.
13. Jang, E.H.; Park, B.J.; Kim, S.H.; Chung, M.A.; Park, M.S.; Sohn, J.H. Emotion classification based on bio-signals emotion recognition using machine learning algorithms. In Proceedings of the 2014 International Conference on Information Science, Electronics and Electrical Engineering, Sapporo, Japan, 26–28 April 2014; Volume 3, pp. 1373–1376.
14. Kofakowska, A.; Szwoch, W.; Szwoch, M. A review of emotion recognition methods based on data acquired via smartphone sensors. *Sensors* **2020**, *20*, 6367. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Zhao, B.; Wang, Z.; Yu, Z.; Guo, B. EmotionSense: Emotion recognition based on wearable wristband. In Proceedings of the 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), Guangzhou, China, 8–12 October 2018; pp. 346–355.
16. Akalin, N.; Köse, H. Emotion recognition in valence-arousal scale by using physiological signals. In Proceedings of the 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, Turkey, 2–5 May 2018; pp. 1–4.
17. Nasoz, F.; Lisetti, C.L.; Vasilakos, A.V. Affectively intelligent and adaptive car interfaces. *Inf. Sci.* **2010**, *180*, 3817–3836. [\[CrossRef\]](#)
18. Hsiao, P.W.; Chen, C.P. Effective attention mechanism in dynamic models for speech emotion recognition. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018; pp. 2526–2530.
19. 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. [\[CrossRef\]](#)
20. Ragot, M.; Martin, N.; Em, S.; Pallamin, N.; Diverrez, J.M. Emotion recognition using physiological signals: Laboratory vs. wearable sensors. In Proceedings of the International Conference on Applied Human Factors and Ergonomics, Los Angeles, CA, USA, 17–21 July 2017; Springer: Berlin/Heidelberg, Germany, 2017; pp. 15–22.
21. Ali, H.; Hariharan, M.; Yaacob, S.; Adom, A.H. Facial emotion recognition using empirical mode decomposition. *Expert Syst. Appl.* **2015**, *42*, 1261–1277. [\[CrossRef\]](#)
22. Liu, Z.T.; Wu, M.; Cao, W.H.; Mao, J.W.; Xu, J.P.; Tan, G.Z. Speech emotion recognition based on feature selection and extreme learning machine decision tree. *Neurocomputing* **2018**, *273*, 271–280. [\[CrossRef\]](#)
23. Gómez-Zaragoza, L.; Marín-Morales, J.; Parra, E.; Guixeres, J.; Alcañiz, M. Speech Emotion Recognition from Social Media Voice Messages Recorded in the Wild. In Proceedings of the International Conference on Human-Computer Interaction, Copenhagen, Denmark, 19–24 July 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 330–336.
24. Abdullah, S.M.S.A.; Ameen, S.Y.A.; Sadeeq, M.A.; Zeebaree, S. Multimodal emotion recognition using deep learning. *J. Appl. Sci. Technol. Trends* **2021**, *2*, 52–58. [\[CrossRef\]](#)

25. Harper, R.; Southern, J. A bayesian deep learning framework for end-to-end prediction of emotion from heartbeat. *IEEE Trans. Affect. Comput.* **2020**. [[CrossRef](#)]
26. Oh, S.; Lee, J.Y.; Kim, D.K. The design of CNN architectures for optimal six basic emotion classification using multiple physiological signals. *Sensors* **2020**, *20*, 866. [[CrossRef](#)] [[PubMed](#)]
27. Ravindran, A.S.; Nakagome, S.; Wickramasuriya, D.S.; Contreras-Vidal, J.L.; Faghih, R.T. Emotion recognition by point process characterization of heartbeat dynamics. In Proceedings of the 2019 IEEE Healthcare Innovations and Point of Care Technologies, (HI-POCT), Bethesda, MD, USA, 20–22 November 2019; pp. 13–16.
28. Gadea, G.H.; Kreuder, A.; Stahlschmidt, C.; Schnieder, S.; Krajewski, J. Brute Force ECG Feature Extraction Applied on Discomfort Detection. In Proceedings of the International Conference on Information Technologies in Biomedicine, Kamień Śląski, Poland, 18–20 June 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 365–376.
29. Moharreri, S.; Dabanloo, N.J.; Maghooli, K. Detection of emotions induced by colors in compare of two nonlinear mapping of heart rate variability signal: Triangle and parabolic phase space (TPSM, PPSM). *J. Med. Biol. Eng.* **2019**, *39*, 665–681. [[CrossRef](#)]
30. Basu, A.; Routray, A.; Shit, S.; Deb, A.K. Human emotion recognition from facial thermal image based on fused statistical feature and multi-class SVM. In Proceedings of the 2015 Annual IEEE India Conference (INDICON), New Delhi, India, 17–20 December 2015; pp. 1–5.
31. Ferdinando, H.; Seppänen, T.; Alasaarela, E. Emotion recognition using neighborhood components analysis and ecg/hrv-based features. In Proceedings of the International Conference on Pattern Recognition Applications and Methods, Porto, Portugal, 24–26 February 2017; Springer: Berlin/Heidelberg, Germany, 2017; pp. 99–113.
32. Higgins, J.P.; Thomas, J.; Chandler, J.; Cumpston, M.; Li, T.; Page, M.J.; Welch, V.A. *Cochrane Handbook for Systematic Reviews of Interventions*; John Wiley & Sons: Hoboken, NJ, USA, 2019.
33. Mamica, M.; Kapłon, P.; Jemioło, P. EEG-Based Emotion Recognition Using Convolutional Neural Networks. In Proceedings of the International Conference on Conceptual Structures, ICCS, Krakow, Poland, 16–18 June 2021; Springer: Berlin/Heidelberg, Germany, 2021. [[CrossRef](#)]
34. Saxena, A.; Khanna, A.; Gupta, D. Emotion recognition and detection methods: A comprehensive survey. *J. Artif. Intell. Syst.* **2020**, *2*, 53–79. [[CrossRef](#)]
35. Resnick, B. More Social Science Studies just Failed to Replicate. Here’s Why This Is Good. Available online: <https://www.vox.com/science-and-health/2018/8/27/17761466/psychology-replication-crisis-nature-social-science> (accessed on 15 February 2022).
36. Maxwell, S.E.; Lau, M.Y.; Howard, G.S. Is psychology suffering from a replication crisis? What does “failure to replicate” really mean? *Am. Psychol.* **2015**, *70*, 487. [[CrossRef](#)]
37. Kilkenny, M.F.; Robinson, K.M. Data Quality: “Garbage In—Garbage Out”. *Health Inf. Manag. J.* **2018**, *47*, 103–105. [[CrossRef](#)]
38. Vidgen, B.; Derczynski, L. Directions in abusive language training data, a systematic review: Garbage in, garbage out. *PLoS ONE* **2020**, *15*, e0243300. [[CrossRef](#)]
39. Stodden, V.; Seiler, J.; Ma, Z. An empirical analysis of journal policy effectiveness for computational reproducibility. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 2584–2589. [[CrossRef](#)]
40. Fehr, J.; Heiland, J.; Himpe, C.; Saak, J. Best practices for replicability, reproducibility and reusability of computer-based experiments exemplified by model reduction software. *arXiv* **2016**, arXiv:1607.01191.
41. Mann, F.; von Walter, B.; Hess, T.; Wigand, R.T. Open access publishing in science. *Commun. ACM* **2009**, *52*, 135–139. [[CrossRef](#)]
42. Kluyver, T.; Ragan-Kelley, B.; Pérez, F.; Granger, B.E.; Bussonnier, M.; Frederic, J.; Kelley, K.; Hamrick, J.B.; Grout, J.; Corlay, S.; et al. *Jupyter Notebooks—A Publishing Format for Reproducible Computational Workflows*; IOS Press: Amsterdam, The Netherlands, 2016.
43. ReScienceX. Available online: <http://rescience.org/x> (accessed on 12 February 2022).
44. ReScience C. Available online: <https://rescience.github.io/> (accessed on 12 February 2022).
45. P Simmons, J.; D Nelson, L.; Simonsohn, U. Pre-registration: Why and how. *J. Consum. Psychol.* **2021**, *31*, 151–162. [[CrossRef](#)]
46. Nosek, B.A.; Ebersole, C.R.; DeHaven, A.C.; Mellor, D.T. The preregistration revolution. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 2600–2606. [[CrossRef](#)] [[PubMed](#)]
47. Joffe, M.M.; Ten Have, T.R.; Feldman, H.I.; Kimmell, S.E. Model selection, confounder control, and marginal structural models: Review and new applications. *Am. Stat.* **2004**, *58*, 272–279. [[CrossRef](#)]
48. Pourhoseingholi, M.A.; Baghestani, A.R.; Vahedi, M. How to control confounding effects by statistical analysis. *Gastroenterol. Hepatol. Bed Bench* **2012**, *5*, 79.
49. Salminen, J.K.; Saarijärvi, S.; Äärelä, E.; Toikka, T.; Kauhanen, J. Prevalence of alexithymia and its association with sociodemographic variables in the general population of Finland. *J. Psychosom. Res.* **1999**, *46*, 75–82. [[CrossRef](#)]
50. Greenaway, K.H.; Kalokerinos, E.K.; Williams, L.A. Context is everything (in emotion research). *Soc. Personal. Psychol. Compass* **2018**, *12*, e12393. [[CrossRef](#)]
51. Saganowski, S.; Dutkowiak, A.; Dziadek, A.; Dziezyc, M.; Komoszyńska, J.; Michalska, W.; Polak, A.; Ujma, M.; Kazienko, P. Emotion recognition using wearables: A systematic literature review-work-in-progress. In Proceedings of the 2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Austin, TX, USA, 23–27 March 2020; pp. 1–6.
52. Peake, J.M.; Kerr, G.; Sullivan, J.P. A critical review of consumer wearables, mobile applications, and equipment for providing biofeedback, monitoring stress, and sleep in physically active populations. *Front. Physiol.* **2018**, *9*, 743. [[CrossRef](#)]

53. Shu, L.; Xie, J.; Yang, M.; Li, Z.; Li, Z.; Liao, D.; Xu, X.; Yang, X. A review of emotion recognition using physiological signals. *Sensors* **2018**, *18*, 2074. [[CrossRef](#)]
54. Kutt, K.; Nalepa, G.J.; Giżycka, B.; Jemioło, P.; Adamczyk, M. Bandreader—a mobile application for data acquisition from wearable devices in affective computing experiments. In Proceedings of the 2018 11th International Conference on Human System Interaction (HSI), Gdansk, Poland, 4–6 July 2018; pp. 42–48. [[CrossRef](#)]
55. Vallejo-Correa, P.; Monsalve-Pulido, J.; Tabares-Betancur, M. A systematic mapping review of context-aware analysis and its approach to mobile learning and ubiquitous learning processes. *Comput. Sci. Rev.* **2021**, *39*, 100335. [[CrossRef](#)]
56. Bardram, J.E.; Matic, A. A decade of ubiquitous computing research in mental health. *IEEE Pervasive Comput.* **2020**, *19*, 62–72. [[CrossRef](#)]
57. Cárdenas-Robledo, L.A.; Peña-Ayala, A. Ubiquitous learning: A systematic review. *Telemat. Inform.* **2018**, *35*, 1097–1132. [[CrossRef](#)]
58. Paré, G.; Kitsiou, S. Methods for literature reviews. In *Handbook of eHealth Evaluation: An Evidence-Based Approach [Internet]*; University of Victoria: Victoria, BC, Canada, 2017.
59. Okoli, C. A guide to conducting a standalone systematic literature review. *Commun. Assoc. Inf. Syst.* **2015**, *37*, 43. [[CrossRef](#)]
60. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G. PRISMA 2009 flow diagram. *PRISMA Statement* **2009**, *6*, 97.
61. Liberati, A.; Altman, D.; Tetzlaff, J.; Mulrow, C.; Gøtzsche, P.; Ioannidis, J.; Clarke, M.; Devereaux, P. The PRISMA statement for reporting systematic and meta-analyses of studies that evaluate interventions. *PLoS Med.* **2009**, *6*, 1–28.
62. Jemioło, P.; Storman, D.; Mamica, M.; Szymkowski, M.; Orzechowski, P.; Dranka, W. Emotion Recognition from Cardiovascular Signals Using Artificial Intelligence—A Systematic Review. Protocol Registration. Available online: <https://osf.io/nj7ut> (accessed on 12 February 2022).
63. Jemioło, P.; Storman, D.; Mamica, M.; Szymkowski, M.; Orzechowski, P. Automated Affect and Emotion Recognition from Cardiovascular Signals—A Systematic Overview of the Field. In Proceedings of the Hawaii International Conference on System Sciences, Maui, HI, USA, 4–7 January 2022; ScholarSpace: Maui, HI, USA, 2022; pp. 4047–4056. [[CrossRef](#)]
64. Jemioło, P.; Storman, D.; Mamica, M.; Szymkowski, M.; Orzechowski, P.; Dranka, W. Emotion Recognition from Cardiovascular Signals Using Artificial Intelligence—A Systematic Review. Supplementary Information. Available online: <https://osf.io/kzj8y/> (accessed on 12 February 2022).
65. Konar, A.; Chakraborty, A. *Emotion Recognition: A Pattern Analysis Approach*; John Wiley & Sons: New York, NY, USA, 2015; pp. 1–45.
66. Kim, K.H.; Bang, S.W.; Kim, S.R. Emotion recognition system using short-term monitoring of physiological signals. *Med. Biol. Eng. Comput.* **2004**, *42*, 419–427. [[CrossRef](#)]
67. Copeland, B. Artificial Intelligence: Definition, Examples, and Applications. Available online: <https://www.britannica.com/technology/artificial-intelligence> (accessed on 10 May 2021).
68. Craik, A.; He, Y.; Contreras-Vidal, J.L. Deep learning for electroencephalogram (EEG) classification tasks: A review. *J. Neural Eng.* **2019**, *16*, 031001.
69. Botchkarev, A. Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology. *arXiv* **2018**, arXiv:1809.03006.
70. McNames, J.; Aboy, M. Statistical modeling of cardiovascular signals and parameter estimation based on the extended Kalman filter. *IEEE Trans. Biomed. Eng.* **2007**, *55*, 119–129. [[CrossRef](#)]
71. Ouzzani, M.; Hammady, H.; Fedorowicz, Z.; Elmagarmid, A. Rayyan—A web and mobile app for systematic reviews. *Syst. Rev.* **2016**, *5*, 210. [[CrossRef](#)]
72. Whiting, P.F.; Rutjes, A.W.; Westwood, M.E.; Mallett, S.; Deeks, J.J.; Reitsma, J.B.; Leeflang, M.M.; Sterne, J.A.; Bossuyt, P.M. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann. Intern. Med.* **2011**, *155*, 529–536.
73. Wolff, R.F.; Moons, K.G.; Riley, R.D.; Whiting, P.F.; Westwood, M.; Collins, G.S.; Reitsma, J.B.; Kleijnen, J.; Mallett, S. PROBAST: A tool to assess the risk of bias and applicability of prediction model studies. *Ann. Intern. Med.* **2019**, *170*, 51–58. [[CrossRef](#)]
74. Benton, M.J.; Hutchins, A.M.; Dawes, J.J. Effect of menstrual cycle on resting metabolism: A systematic review and meta-analysis. *PLoS ONE* **2020**, *15*, e0236025. [[CrossRef](#)] [[PubMed](#)]
75. Koelstra, S.; Muhl, C.; Soleymani, M.; Lee, J.S.; Yazdani, A.; Ebrahimi, T.; Pun, T.; Nijholt, A.; Patras, I. Deap: A database for emotion analysis; using physiological signals. *IEEE Trans. Affect. Comput.* **2011**, *3*, 18–31. [[CrossRef](#)]
76. Soleymani, M.; Lichtenauer, J.; Pun, T.; Pantic, M. A multimodal database for affect recognition and implicit tagging. *IEEE Trans. Affect. Comput.* **2011**, *3*, 42–55. [[CrossRef](#)]
77. Abadi, M.K.; Subramanian, R.; Kia, S.M.; Avesani, P.; Patras, I.; Sebe, N. DECAF: MEG-based multimodal database for decoding affective physiological responses. *IEEE Trans. Affect. Comput.* **2015**, *6*, 209–222. [[CrossRef](#)]
78. Correa, J.A.M.; Abadi, M.K.; Sebe, N.; Patras, I. Amigos: A dataset for affect, personality and mood research on individuals and groups. *IEEE Trans. Affect. Comput.* **2021**, *12*, 479–493. [[CrossRef](#)]
79. Subramanian, R.; Wache, J.; Abadi, M.K.; Vieriu, R.L.; Winkler, S.; Sebe, N. ASCERTAIN: Emotion and personality recognition using commercial sensors. *IEEE Trans. Affect. Comput.* **2016**, *9*, 147–160. [[CrossRef](#)]
80. Kim, J.; André, E. Emotion recognition based on physiological changes in music listening. *IEEE Trans. Pattern Anal. Mach. Intell.* **2008**, *30*, 2067–2083. [[CrossRef](#)]

81. Quiroz, J.C.; Geangu, E.; Yong, M.H. Emotion recognition using smart watch sensor data: Mixed-design study. *JMIR Ment. Health* **2018**, *5*, e10153. [[CrossRef](#)]
82. Pinto, J. *Exploring Physiological Multimodality for Emotional Assessment*; Instituto Superior Técnico (IST): Lisboa, Portugal, 2019.
83. Yang, W.; Rifqi, M.; Marsala, C.; Pinna, A. Physiological-based emotion detection and recognition in a video game context. In Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 8–13 July 2018; pp. 1–8.
84. Gupta, R.; Khomami Abadi, M.; Cárdenes Cabré, J.A.; Morreale, F.; Falk, T.H.; Sebe, N. A quality adaptive multimodal affect recognition system for user-centric multimedia indexing. In Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval, New York, NY, USA, 6–9 June 2016; pp. 317–320.
85. Marín-Morales, J.; Higuera-Trujillo, J.L.; Greco, A.; Guixeres, J.; Llinares, C.; Scilingo, E.P.; Alcañiz, M.; Valenza, G. Affective computing ual reality: Emotion recognition from brain and heartbeat dynamics using wearable sensors. *Sci. Rep.* **2018**, *8*, 13657.
86. Hsu, Y.L.; Wang, J.S.; Chiang, W.C.; Hung, C.H. Automatic ECG-based emotion recognition in music listening. *IEEE Trans. Affect. Comput.* **2017**, *11*, 85–99. [[CrossRef](#)]
87. Schmidt, P.; Reiss, A.; Duerichen, R.; Marberger, C.; Van Laerhoven, K. Introducing wesad, a multimodal dataset for wearable stress and affect detection. In Proceedings of the 20th ACM International Conference on Multimodal Interaction, Boulder, CO, USA, 16–20 October 2018; pp. 400–408.
88. Song, T.; Zheng, W.; Lu, C.; Zong, Y.; Zhang, X.; Cui, Z. MPED: A multi-modal physiological emotion database for discrete emotion recognition. *IEEE Access* **2019**, *7*, 12177–12191. [[CrossRef](#)]
89. Ranganathan, H.; Chakraborty, S.; Panchanathan, S. Multimodal emotion recognition using deep learning architectures. In Proceedings of the 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Placid, NY, USA, 7–10 March 2016; pp. 1–9.
90. Katsigiannis, S.; Ramzan, N. DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices. *IEEE J. Biomed. Health Inform.* **2017**, *22*, 98–107. [[CrossRef](#)] [[PubMed](#)]
91. Huang, W.; Liu, G.; Wen, W. MAPD: A Multi-subject Affective Physiological Database. In Proceedings of the 2014 Seventh International Symposium on Computational Intelligence and Design, Hangzhou, China, 13–14 December 2014; Volume 2; pp. 585–589.
92. Yannakakis, G.N.; Martínez, H.P.; Jhala, A. Towards affective camera control in games. *User Model. User-Adapt. Interact.* **2010**, *20*, 313–340. [[CrossRef](#)]
93. McInnes, M.D.; Moher, D.; Thombs, B.D.; McGrath, T.A.; Bossuyt, P.M.; Clifford, T.; Cohen, J.F.; Deeks, J.J.; Gatsonis, C.; Hooft, L.; et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: The PRISMA-DTA statement. *JAMA* **2018**, *319*, 388–396. [[CrossRef](#)] [[PubMed](#)]
94. Wierzbicka, A. Defining emotion concepts. *Cogn. Sci.* **1992**, *16*, 539–581. [[CrossRef](#)]
95. Wierzbicka, A. Emotion, language, and cultural scripts. In *Emotion and Culture: Empirical Studies of Mutual Influence*; American Psychological Association: Washington, DC, USA, 1994; pp. 133–196.
96. Cook, D.A.; Levinson, A.J.; Garside, S. Method and reporting quality in health professions education research: A systematic review. *Med. Educ.* **2011**, *45*, 227–238. [[CrossRef](#)]
97. Wijasena, H.Z.; Ferdiana, R.; Wibirama, S. A Survey of Emotion Recognition using Physiological Signal in Wearable Devices. In Proceedings of the 2021 International Conference on Artificial Intelligence and Mechatronics Systems (AIMS), Bandung, Indonesia, 28–30 April 2021; pp. 1–6.
98. Saganowski, S.; Kazienko, P.; Dziezyc, M.; Jakimow, P.; Komoszynska, J.; Michalska, W.; Dutkowiak, A.; Polak, A.; Dziadek, A.; Ujma, M. Consumer Wearables and Affective Computing for Wellbeing Support. In Proceedings of the MobiQuitous 2020—17th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, Darmstadt, Germany 7–9 December 2020.
99. Schmidt, P.; Reiss, A.; Dürichen, R.; Laerhoven, K.V. Wearable-Based Affect Recognition—A Review. *Sensors* **2019**, *19*, 4079. [[CrossRef](#)]
100. Merone, M.; Soda, P.; Sansone, M.; Sansone, C. ECG databases for biometric systems: A systematic review. *Expert Syst. Appl.* **2017**, *67*, 189–202. [[CrossRef](#)]
101. Da Silva, H.P.; Lourenço, A.; Fred, A.; Raposo, N.; Aires-de Sousa, M. Check Your Biosignals Here: A new dataset for off-the-person ECG biometrics. *Comput. Methods Programs Biomed.* **2014**, *113*, 503–514. [[CrossRef](#)]
102. Hong, S.; Zhou, Y.; Shang, J.; Xiao, C.; Sun, J. Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review. *Comput. Biol. Med.* **2020**, *122*, 103801. [[CrossRef](#)]
103. Santamaria-Granados, L.; Munoz-Organero, M.; Ramirez-Gonzalez, G.; Abdulhay, E.; Arunkumar, N. Using deep convolutional neural network for emotion detection on a physiological signals dataset (AMIGOS). *IEEE Access* **2018**, *7*, 57–67. [[CrossRef](#)]
104. Soroush, M.Z.; Maghooli, K.; Setarehdan, S.K.; Nasrabadi, A.M. A review on EEG signals based emotion recognition. *Int. Clin. Neurosci. J.* **2017**, *4*, 118.
105. Suhaimi, N.S.; Mountstephens, J.; Teo, J. EEG-based emotion recognition: A state-of-the-art review of current trends and opportunities. *Comput. Intell. Neurosci.* **2020**, *2020*, 8875426. [[CrossRef](#)] [[PubMed](#)]
106. Wagh, K.P.; Vasanth, K. Electroencephalograph (EEG) based emotion recognition system: A review. In *Innovations in Electronics and Communication Engineering*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 37–59.

107. Egger, M.; Ley, M.; Hanke, S. Emotion recognition from physiological signal analysis: A review. *Electron. Notes Theor. Comput. Sci.* **2019**, *343*, 35–55. [[CrossRef](#)]
108. Jerritta, S.; Murugappan, M.; Nagarajan, R.; Wan, K. Physiological signals based human emotion recognition: A review. In Proceedings of the 2011 IEEE 7th International Colloquium on Signal Processing and its Applications, Penang, Malaysia, 4–6 March 2011; pp. 410–415.
109. Ali, M.; Mosa, A.H.; Al Machot, F.; Kyamakya, K. Emotion recognition involving physiological and speech signals: A comprehensive review. In *Recent Advances in Nonlinear Dynamics and Synchronization*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 287–302.
110. Szwoch, W. Using physiological signals for emotion recognition. In Proceedings of the 2013 6th International Conference on Human System Interactions (HSI), Sopot, Poland, 6–8 June 2013; pp. 556–561.
111. Callejas-Cuervo, M.; Martínez-Tejada, L.A.; Alarcón-Aldana, A.C. Emotion recognition techniques using physiological signals and video games-Systematic review. *Rev. Fac. Ing.* **2017**, *26*, 19–28. [[CrossRef](#)]
112. Marechal, C.; Mikolajewski, D.; Tyburek, K.; Prokopowicz, P.; Bougueroua, L.; Ancourt, C.; Wegrzyn-Wolska, K. Survey on AI-Based Multimodal Methods for Emotion Detection. In *High-Performance Modelling and Simulation for Big Data Application*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 307–324.
113. Picard, R.W. *Affective Computing*; MIT Press: Cambridge, MA, USA, 1997.
114. Amira, T.; Dan, I.; Az-eddine, B.; Ngo, H.H.; Said, G.; Katarzyna, W.W. Monitoring chronic disease at home using connected devices. In Proceedings of the 2018 13th Annual Conference on System of Systems Engineering (SoSE), Paris, France 19–22 June 2018; pp. 400–407.
115. Khan, K.S.; Kunz, R.; Kleijnen, J.; Antes, G. Five steps to conducting a systematic review. *J. R. Soc. Med.* **2003**, *96*, 118–121. [[CrossRef](#)]
116. Enhancing the Quality and Transparency of Health Research. Available online: <https://www.equator-network.org/> (accessed on 12 February 2022).
117. Giżycka, B.; Jemioło, P.; Domarecki, S.; Świder, K.; Wiśniewski, M.; Mielczarek, Ł. A Thin Light Blue Line—Towards Balancing Educational and Recreational Values of Serious Games. In Proceedings of the 3rd Workshop on Affective Computing and Context Awareness in Ambient Intelligence, Cartagena, Spain, 11–12 November 2019; Technical University of Aachen: Aachen, Germany, 2019.
118. Jemioło, P.; Giżycka, B.; Nalepa, G.J. Prototypes of arcade games enabling affective interaction. In Proceedings of the International Conference on Artificial Intelligence and Soft Computing, Zakopane, Poland, 16–20 June 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 553–563. [[CrossRef](#)]
119. Nalepa, G.J.; Kutt, K.; Giżycka, B.; Jemioło, P.; Bobek, S. Analysis and use of the emotional context with wearable devices for games and intelligent assistants. *Sensors* **2019**, *19*, 2509. [[CrossRef](#)]
120. Benovoy, M.; Cooperstock, J.R.; Deitcher, J. Biosignals analysis and its application in a performance setting. In Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing, Madeira, Portugal, 3–4 January 2008; pp. 253–258.
121. Mera, K.; Ichimura, T. Emotion analyzing method using physiological state. In Proceedings of the International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, Wellington, New Zealand, 20–25 September 2004; Springer: Berlin/Heidelberg, Germany, 2004; pp. 195–201.
122. Wang, Y.; Mo, J. Emotion feature selection from physiological signals using tabu search. In Proceedings of the 2013 25th Chinese Control and Decision Conference (CCDC), Guiyang, China, 25–27 May 2013; pp. 3148–3150.
123. Guendil, Z.; Lachiri, Z.; Maaoui, C.; Pruski, A. Emotion recognition from physiological signals using fusion of wavelet based features. In Proceedings of the 2015 7th International Conference on Modelling, Identification and Control (ICMIC), Sousse, Tunisia, 18–20 December 2015; pp. 1–6.
124. Zong, C.; Chetouani, M. Hilbert-Huang transform based physiological signals analysis for emotion recognition. In Proceedings of the 2009 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), Ajman, United Arab Emirates, 14–17 December 2009; pp. 334–339.
125. Joesph, C.; Rajeswari, A.; Premalatha, B.; Balapriya, C. Implementation of physiological signal based emotion recognition algorithm. In Proceedings of the 2020 IEEE 36th International Conference on Data Engineering (ICDE), Dallas, TX, USA, 20–24 April 2020; pp. 2075–2079.
126. Leon, E.; Clarke, G.; Sepulveda, F.; Callaghan, V. Neural network-based improvement in class separation of physiological signals for emotion classification. In Proceedings of the IEEE Conference on Cybernetics and Intelligent Systems, Singapore, 1–3 December 2004; Volume 2, pp. 724–728.
127. Siow, S.C.; Loo, C.K.; Tan, A.W.; Liew, W.S. Adaptive Resonance Associative Memory for multi-channel emotion recognition. In Proceedings of the 2010 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES), Kuala Lumpur, Malaysia, 30 November–2 December 2010; pp. 359–363.
128. Perez-Rosero, M.S.; Rezaei, B.; Akcakaya, M.; Ostadabbas, S. Decoding emotional experiences through physiological signal processing. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 5–9 March 2017; pp. 881–885.

129. Sokolova, M.V.; Fernández-Caballero, A.; López, M.T.; Martínez-Rodrigo, A.; Zangróniz, R.; Pastor, J.M. A distributed architecture for multimodal emotion identification. In Proceedings of the 13th International Conference on Practical Applications of Agents and Multi-Agent Systems, Salamanca, Spain, 3–5 June 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 125–132.
130. Shirahama, K.; Grzegorzec, M. Emotion recognition based on physiological sensor data using codebook approach. In Proceedings of the Conference of Information Technologies in Biomedicine, Kamien Slaski, Poland, 20–22 June 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 27–39.
131. Gong, P.; Ma, H.T.; Wang, Y. Emotion recognition based on the multiple physiological signals. In Proceedings of the 2016 IEEE International Conference on Real-time Computing and Robotics (RCAR), Angkor Wat, Cambodia, 6–10 June 2016; pp. 140–143.
132. Jain, M.; Saini, S.; Kant, V. A hybrid approach to emotion recognition system using multi-discriminant analysis & k-nearest neighbour. In Proceedings of the 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Udupi, India, 13–16 September 2017; pp. 2251–2256.
133. Guendil, Z.; Lachiri, Z.; Maaoui, C.; Pruski, A. Multiresolution framework for emotion sensing in physiological signals. In Proceedings of the 2016 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Monastir, Tunisia, 21–23 March 2016; pp. 793–797.
134. Wong, W.M.; Tan, A.W.; Loo, C.K.; Liew, W.S. PSO optimization of synergetic neural classifier for multichannel emotion recognition. In Proceedings of the 2010 Second World Congress on Nature and Biologically Inspired Computing (NaBIC), Kitakyushu, Japan, 15–17 December 2010; pp. 316–321.
135. Guo, X. Study of emotion recognition based on electrocardiogram and RBF neural network. *Procedia Eng.* **2011**, *15*, 2408–2412.
136. Zhu, H.; Han, G.; Shu, L.; Zhao, H. ArvaNet: Deep Recurrent Architecture for PPG-Based Negative Mental-State Monitoring. *IEEE Trans. Comput. Soc. Syst.* **2020**, *8*, 179–190. [[CrossRef](#)]
137. Wu, C.H.; Kuo, B.C.; Tzeng, G.H. Factor analysis as the feature selection method in an Emotion Norm Database. In Proceedings of the Asian Conference on Intelligent Information and Database Systems, Bangkok, Thailand, 7–9 April 2014; Springer: Berlin/Heidelberg, Germany, 2014; pp. 332–341.
138. Akbulut, F.P.; Perros, H.G.; Shahzad, M. Bimodal affect recognition based on autoregressive hidden Markov models from physiological signals. *Comput. Methods Programs Biomed.* **2020**, *195*, 105571.
139. Takahashi, M.; Kubo, O.; Kitamura, M.; Yoshikawa, H. Neural network for human cognitive state estimation. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'94), Munich, Germany, 12–16 September 1994; Volume 3, pp. 2176–2183.
140. Gao, Y.; Barreto, A.; Adjouadi, M. An affective sensing approach through pupil diameter processing and SVM classification. *Biomed. Sci. Instrum.* **2010**, *46*, 326–330. [[PubMed](#)]
141. Mohino-Herranz, I.; Gil-Pita, R.; Rosa-Zurera, M.; Seoane, F. Activity recognition using wearable physiological measurements: Selection of features from a comprehensive literature study. *Sensors* **2019**, *19*, 5524. [[CrossRef](#)] [[PubMed](#)]
142. Bonarini, A.; Costa, F.; Garbarino, M.; Matteucci, M.; Romero, M.; Tognetti, S. Affective videogames: The problem of wearability and comfort. In Proceedings of the International Conference on Human-Computer Interaction, Orlando, FL, USA, 9–14 July 2011; Springer: Berlin/Heidelberg, Germany, 2011; pp. 649–658.
143. Alqahtani, F.; Katsigiannis, S.; Ramzan, N. ECG-based affective computing for difficulty level prediction in intelligent tutoring systems. In Proceedings of the 2019 UK/China Emerging Technologies (UCET), Glasgow, UK, 21–22 August 2019; pp. 1–4.
144. Xu, J.; Hu, Z.; Zou, J.; Bi, A. Intelligent emotion detection method based on deep learning in medical and health data. *IEEE Access* **2019**, *8*, 3802–3811. [[CrossRef](#)]
145. Wendt, C.; Popp, M.; Karg, M.; Kuhnlenz, K. Physiology and HRI: Recognition of over-and underchallenge. In Proceedings of the RO-MAN 2008—The 17th IEEE International Symposium on Robot and Human Interactive Communication, Munich, Germany, 1–3 August 2008; pp. 448–452.
146. Omata, M.; Moriwaki, K.; Mao, X.; Kanuka, D.; Imamiya, A. Affective rendering: Visual effect animations for affecting user arousal. In Proceedings of the 2012 International Conference on Multimedia Computing and Systems, Tangiers, Morocco, 10–12 May 2012; pp. 737–742.
147. van den Broek, E.L.; Schut, M.H.; Westerink, J.H.; Tuinenbreijer, K. Unobtrusive sensing of emotions (USE). *J. Ambient Intell. Smart Environ.* **2009**, *1*, 287–299. [[CrossRef](#)]
148. Quiroz, J.C.; Yong, M.H.; Geangu, E. Emotion-recognition using smart watch accelerometer data: Preliminary findings. In Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Maui, HI, USA, 11–15 September 2017; pp. 805–812.
149. Althobaiti, T.; Katsigiannis, S.; West, D.; Bronte-Stewart, M.; Ramzan, N. Affect detection for human-horse interaction. In Proceedings of the 2018 21st Saudi Computer Society National Computer Conference (NCC), Riyadh, Saudi Arabia, 25–26 April 2018; pp. 1–6.
150. Dobbins, C.; Fairclough, S. Detecting negative emotions during real-life driving via dynamically labelled physiological data. In Proceedings of the 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Athens, Greece, 19–23 March 2018; pp. 830–835.
151. Jo, Y.; Lee, H.; Cho, A.; Whang, M. Emotion Recognition Through Cardiovascular Response in Daily Life Using KNN Classifier. In *Advances in Computer Science and Ubiquitous Computing*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 1451–1456.

152. Hamdi, H.; Richard, P.; Suteau, A.; Allain, P. Emotion assessment for affective computing based on physiological responses. In Proceedings of the 2012 IEEE International Conference on Fuzzy Systems, Brisbane, QLD, Australia, 10–15 June 2012; pp. 1–8.
153. Moghimi, S.; Chau, T.; Guerguerian, A.M. Using prefrontal cortex near-infrared spectroscopy and autonomic nervous system activity for identifying music-induced emotions. In Proceedings of the 2013 6th International IEEE/EMBS Conference on Neural Engineering (NER), San Diego, CA, USA, 6–8 November 2013; pp. 1283–1286.
154. Zhang, Z.; Tanaka, E. Affective computing using clustering method for mapping human's emotion. In Proceedings of the 2017 IEEE International Conference on Advanced Intelligent Mechatronics (AIM), Munich, Germany, 3–7 July 2017; pp. 235–240.
155. Reinerman-Jones, L.; Taylor, G.; Cosenzo, K.; Lackey, S. Analysis of multiple physiological sensor data. In Proceedings of the International Conference on Foundations of Augmented Cognition, Orlando, FL, USA, 9–14 July 2011; Springer: Berlin/Heidelberg, Germany, 2011; pp. 112–119.
156. Roza, V.; Postolache, O.; Groza, V.; Pereira, J.D. Emotions Assessment on Simulated Flights. In Proceedings of the 2019 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Istanbul, Turkey, 26–28 June 2019; pp. 1–6.
157. Savran, A.; Ciftci, K.; Chanel, G.; Mota, J.; Viet, L.; Sankur, B.; Akarun, L.; Caplier, A.; Rombaut, M. Emotion detection in the loop from brain signals and facial images. In Proceedings of the eNTERFACE 2006 Workshop, Dubrovnik, Croatia, 17 July–11 August 2006.