CCF Transactions on Pervasive Computing and Interaction
https://doi.org/10.1007/s42486-023-00129-6

SURVEY PAPER

A survey on EEG-based neurophysiological research for emotion recognition

Jenamani Chandrakanta Badajena[1] · Srinivas Sethi[2] · Sanjit Kumar Dash[3] · Ramesh Kumar Sahoo[2]

Received: 14 November 2022 / Accepted: 13 March 2023

Abstract
Emotions play a significant part in a person's social connections, decision-making, and perception of the world. Elicited emotions cause a change in a person's physiological and psychological states. As Electroencephalography (EEG) facilitates a close study of brain activity, it is becoming a standard method among the research community for reliable recognition of human emotions. This work demonstrates various advancements in emotion recognition utilizing EEG signals and points out major changing trends by making a comparison of previously available research in this field. In addition to the survey a detailed explanation of the procedure for refining EEG for emotion recognition has been explained in this work. This aims to help researchers, especially beginners, have a thorough understanding of the developmental research in this field.

Keywords EEG · Human–computer interaction · Human brain · Emotion recognition

1 Introduction

Artificial intelligence is a field of research based on logical and mathematical intelligence. There have been several categories of research work in several fields that have tried to leverage the benefit of artificial intelligence for better results and insights. One such area of focus is emotion recognition which can be associated with various human-centric computational solutions. Machine learning applications can be helpful in improving computation involving human emotions. Understanding emotions is a crucial form of communication that involves several forms of human-gestures and can help make HCI (Human–Computer Interaction) more appreciable for human-centric applications. The sources of human recognition may include facial expressions, voice intonations, body language, and even subjective Self-Assessment Manikin (SAM reports). While these methods can be useful and have been used for years, researchers have concluded that they are unreliable for emotion recognition because subjects can easily hide their original emotion by manipulating various bodily expressions. Thus, the latest standard is to employ physiological signals for the purpose (Homan et al. 1987). The various forms of physiological signals for emotion recognition may include factors like heart rate (HR), electromyogram signal (EMG), respiratory volume (RV), skin conductance (SKC), skin temperature (SKT), blood volume pulse (BVP), etc. Studies have shown EEG signals to be one of the best choices for emotion classification given that they are non-invasive, rapid, and affordable (Wioleta 2013).

In this paper, the survey conducted by Alracao et al. (2017) has been extended and presented in the form of an in-depth analysis of the changing trends of several significant factors while carrying out the experiment to recognize emotion using EEG (Alarcao and Fonseca 2017). For this study, two queries were used with both IEEE Xplore and Google Scholar. The first one being "EEG + Emotions + Recognition" while the latter being "EEG + Emotions + Identification". From the output received from the search query, around 60 research papers were selected between the years 2017 and 2023 to carry out this survey. The work depicted in this paper has been organized in different sections. The Sect. 2 provides a brief background study to provide a more in-depth understanding of EEG and the physiological aspects of emotions in the human brain. It provides information

✉ Jenamani Chandrakanta Badajena
chand.cet@gmail.com

[1] Indira Gandhi Institute of Technology, Sarang,Utkal University, Bhubaneswar, India

[2] Indira Gandhi Institute of Technology, Sarang, India

[3] Odisha University of Technology and Research, Bhubaneswar, India

Published online: 03 April 2023

about the prerequisites for a better understanding of this review. The Sect. 3 presents a detailed comparison of this work along with several previously carried out research in this field and the next section provides concluding remarks about this research work.

2 Background

This section includes the study of Electroencephalography along with its equipment, placements, and models. The detail process of raw EEG signal acquisition and processing has also been explained based on differences in frequencies of signals. The next stage provides an insight into Brouwer's recommendations (Brouwer et al. 2015) for more reliable research in this field.

2.1 Feature extraction from EEG

EEG stands for Electroencephalography, a test done using small metal electrodes with tiny wires that are adhered to the scalp as electrodes (Teplan 2002). It directly records the physiologic signal activity in the brain. Thus, it is a more preferred input for accurate human emotion detection as (unlike non-physiological signals) it cannot be disguised as per the participant's will (Zhao et al. 2016). Zhao et al. (2017) have demonstrated how EEG provides better classification accuracy as compared to other behavioral indicators like speech, facial expressions, intonation, etc. Neurons in the brain interact with each other via sensory neurons. The prefrontal, temporal, parietal, and occipital lobes make up the cortex, which is the biggest region our brain, with each lobe having its own functions to perform. The prefrontal cortex, as the name suggests, is placed in front of the brain and is an important part of the emotion-processing network. Therefore, it acts as a control center, reacting to our activities. The parietal lobe is primarily involved in temperature feeling, the texture of the food, sensation, and mobility, and the occipital lobe is in charge of perception. Before recording an emotion, it is critical to design the experiment around which the emotion will be recorded, with a subject count of around 30 and a balanced gender ratio, and the emotional stimuli being visual, auditory, tactile, or olfactory stimulation. Depending on the emotions to be recorded, electrode placement is done. The most common techniques for designating electrode locations and positions along with the skull were provided by the American Encephalographic Society (1994) and Walsh et al. (2005). These are commonly referred to as the 10–20 and 10–5 systems, respectively. Reference electrodes are put at 10% and 20% sites throughout arcs of coordinates in the 10–20 system (Homan et al. 1987), as shown in

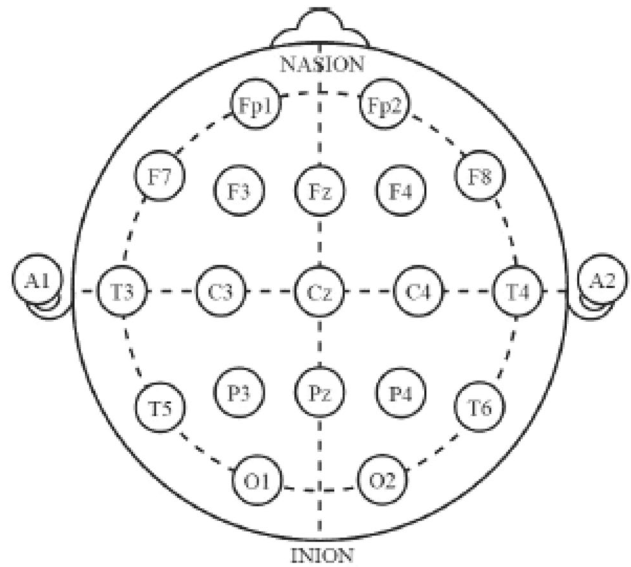

Fig. 1 Placement of electrodes as per 10–20 electrode system

Fig. 1, along the nasion (Nz) and inion (Iz). The standard nomenclature of all electrodes in 10–20 systems has been defined by Clin in a paper published in 1994 (Electrode Position Nomenclature Committee 1994). The raw data recorded is continuous and noisy, with significant artifact deviations like eye movements, heartbeats, muscular movements, and line disturbance and a low signal to noise value (Wang et al. 2019).

Epoching and segmenting work by dividing a continuous signal into segments of interest known as events. The two kinds of potentials- ERP (Goldstein 2010) (related to events known as Event Related Potential) and SEP (Luck and Kappenman 2011) (related to human senses known as Sensory evoked potential) are the most commonly used paradigms. Along with the two kinds of potentials, desynchronizations related to events are also crucial for research of human sentiments (Bekkedal et al. 2011). ERP is best if one wants to measure short responses that are time-locked to an event. This is accomplished with the help of different amplitudes and latencies of both negative (N100) and positive (P100) potentials (Sur and Sinha 2009). ERD is appropriate when you already have existing reactions to compare them to, done with power changes within a specified frequency band. The subject trial average is calculated to obtain a smooth signal across a single subject. Then, the average across all subjects is done. The most important step in extracting distinctive patterns from the raw data for consistent categorization is feature extraction. The next fundamental approach in machine learning for data analysis that categorizes categorical labels is classification done by various classifiers (Nanthini and Santhi 2017).

2.2 Emotion in the brain

Emotions are concentrated on a certain aspect, such as a person or a circumstance, by being more intense. Plutchik (2001) identified eight fundamental emotions that can be crucial for recognition of human emotions. From the emotion-processing network, some of the major regions are the prefrontal cortex, amygdala, basal ganglia, cingulate cortex, and hippocampus. Each area has its own unique function, and they all work together to recognize and control emotion. The amygdala is roughly the size of an almond and it is a very important part of the brain that processes both positive and negative information. The emotional frequency varies from higher to lower emotions based on the hertz frequency rate of vibrational analysis, with neutrality being 250 Hz. Between 20 and 250 Hz, pride, desire, anger, grief, fear, guilt, apathy, and shame; above 250 and 700 Hz, acceptance, will, reason, love, joy, peace, and enlightenment. The human emotions are known to be represented by dimensions of certain parameters like valence, dominance and arousal. Arousal (also known as activation) swings from tired to enthusiastic and eventually to dominance, which corresponds to the intensity of the emotion (Ekman 1999; Lang 1995). The potential elements related to events of intermediate (N200 and P200) to brief (N100 and P100) latencies have already been found to be associated with valence, and elements of moderate to large [P300 and SCP (Slow Cortical Potential)] latencies were shown to correspond with excitement (Posner et al. 2005). This is retrieved while epoching and segmenting events from continuous EEG signals. The convention's naming of N100 or P100 is given by N for the negative rise and P for the positive rise, with 100 representing the EEG graph's time horizon. If the graph has a positive y-axis, the convention is considered negative. When women watch sorrowful films, they feel a high level of arousal. The results also imply that men have fervent emotional encounters while watching videos that provoke an emotional response, but women are more expressive emotionally, typically for negative emotions. Furthermore, from Deng et al. (2016), it is evident that disparities in dimensions of gender may be dependent on the type of emotion, but not on valence of a human emotion.

2.3 Brouwer's Recommendations

Brain–Computer Interface (BCI) and identifying emotions require a lot of knowledge from many different fields, such as advanced machine learning algorithms, system design, mathematical modeling, neurophysiology, etc. Brouwer et al. (2015) proposed six recommendations for avoiding common difficulties associated with the use of neurophysiological data that reflect mental states in order to avoid shortcomings of emotion detection. The recommendations have been explained in detail.

2.3.1 Recommendation-1: Define the ground truth of the state of interest

Different authors sometimes refer to similar concepts using distinct terminology. This is largely dependent on the disciplines from which the researchers come. Thus, to avoid confusion, it is necessary to clearly understand and narrow down the specific mental state that has to be addressed in the study. Key point 1.1 states that it is necessary to understand the operations of ground truth and relevant point of interest. For determining ground truth, key point 1.2 mentions examining multiple measures such as subjective measures (e.g., SAM scale responses, etc.), behavioral measures (e.g., the accuracy of the button press, etc.), and knowledge of the condition the participant is experiencing (e.g., the difficulty level of the task, the response the stimulus aims to elicit, etc.).

2.3.2 Recommendation-2: Neurophysiology as a point of interest

After examining the state of interest, the most important step is to understand and map a specific mental state to corresponding physiological signals (EEG in this case). While emotional experiences are linked to physiological processes, one of the key takeaways of the review is that the physiological reaction is defined not just by the emotions elicited but, more importantly, by the related future reaction. However, there is no such clear-cut explanation of how and to what extent mental states (psychological parameters) connect with neuropsychological parameters. Thus, it becomes even more important that the findings related to the state of interest are explained; all the varying parameters are well formulated and presented as structured hypotheses by the researcher, as stated by key point 2.1.

2.3.3 Recommendation-3: Eliminate confounding factors

Confounding factors are those that vary with the parameters of interest and might interfere with the study at hand. Thus, while designing the system, it is important to check for such factors (key point 3.1). Examples of these can include body movements (because of the small disruptions in sensors and wiring, movement of the eyes, hands, and other body parts while capturing data might introduce artifacts in the data), overloading stimuli (a complete visual stimulus can result in differences in EEG recordings), etc. Key point 3.2 states to examine the occurrence of confounding factors, and a key point 3.3 states to select the data post hoc to reduce confounding factors. Key point 3.4 verifies the consistency of neurophysiological data between hypothesized states and confounding effects and also checks if the confounds in the data were causing an overly optimistic classification result.

2.3.4 Recommendation-4: Stick to good classification practice

A seemingly simple classification practice is to train a classification model using signal data pre-labeled according to the emotional state and then further use this model to classify and label previously unknown neurophysiological data. The classifier's performance can then be evaluated by comparing this estimated label with a known label (if available). However, it is important to remember to have as many independent variables/parameters as feasible. If the classification model's training and testing data are not independent of one another, the model's overall performance will most likely be overstated (Key points 4.1, 4.2). Key point 4.3 states that the evaluation of classification performance should be presented in a proper manner (like a tabular or graphical representation).

2.3.5 Recommendation-5: Insight into the cause of classification success

As stated previously, some confounding factors might contribute to the deceptive success of the classification model. Thus, it's crucial to highlight not only the end output but also how the neurophysiological parameters of the model vary (Key Point 5.1). This confirms that the data is correct and that confounding is not to blame for the overly optimistic classification efficiency. Key point 5.2: check the success of a combination of features to be completely sure of classifier accuracy.

2.3.6 Recommendation-6: Added value of using neurophysiology

It is suggested that researchers provide more information and state under what conditions and in what ways EEG signals are expected to be more useful. This aims to explain the reason why when other techniques are available, researchers should indulge in and invest in EEG (Key Point 6.1). Key Point 6.2 is to mention the further applications for which neurophysiological signals can be beneficial.

3 Analysis

In this section, a detailed analysis of papers on emotion recognition and a meaningful comparison between the different factors to understand more about the changing trends in the research concerned with emotion recognition using EEG have been presented.

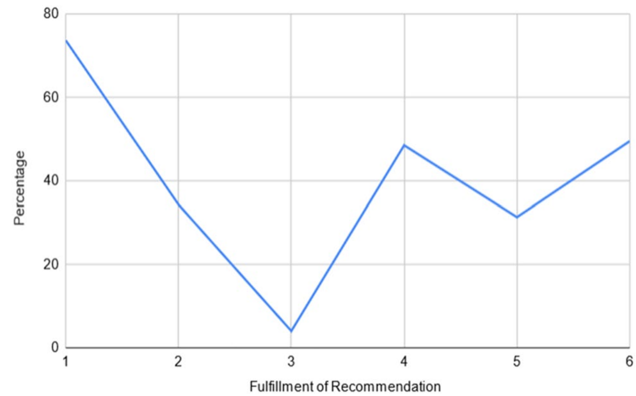

Fig. 2 Percentage fulfilment of the six recommendations in the previous analysis

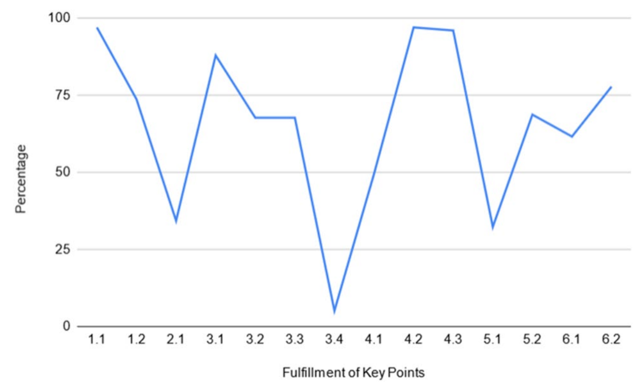

Fig. 3 Percentage fulfilment of the major key points in the previous analysis

3.1 Brouwer's Recommendation

As discussed in Sect. 2.3, Brouwer's recommendations could be used as a checklist for reviewing and evaluating research, as well as to enhance the design and execution of new investigations. Figures 2 and 3 represent the percentage of works that fulfill the six recommendations and key points, respectively, in the previous review conducted by Alarcao and Fonseca (2017). From Figs. 2 and 3, it can be seen that a little more than 70% of the works satisfy recommendation 1, meaning they satisfy both key points (1.1 and 1.2). Considering point 1.1, a huge number of works (97%) comply with it. A lot fewer, 73.7% of works, are in accordance with point 1.2. Only 34.3 percent of the works follow recommendation 2. The percentage fulfillment of recommendation 3 (all key points included) is quite low (4.0%). It can be observed from Fig. 3 that around 90% of works satisfy key point 3.1, but fewer researchers tend to focus on key points

3.2 and 3.3. Nearly 50% of works satisfy recommendation 4. Around 31.3 percent of works are in accordance with recommendation 5. Key point 5.1 is satisfied by 32.3 percent of the works, and key point 5.2 is satisfied by 68% of the works. Approximately 49.5 percent of works are associated with the recommendation 6. Key point 6.1 is satisfied by 61.6 percent of papers and 70.8 percent satisfy key point 6.2.

For this purpose, a survey of papers was conducted on papers from the years 2017–2023 and the findings have been provided in Table 1. The symbol '√' shows that the recommendation's key point has been followed in this particular paper. Figures 4 and 5 represent the percentage of works that fulfill the six recommendations and key points, respectively, in the current analysis. It shows that about 26.9% of the works meet recommendation 1. Taking key point 1.1 into account, a large number of works (around 88.4%) adhere to it. This can be explained by the fact that most researchers try to give detailed presentations of how they intend to operationalize the problem at hand (recognizing emotion). Key point 1.2 is followed by far fewer works (26.9%). The reason could be that researchers consider EEG signals to be sufficient input for experimental emotion recognition. From Table 1, it can be seen that around 73.07% of the works are in accordance with recommendation 2. The percentage

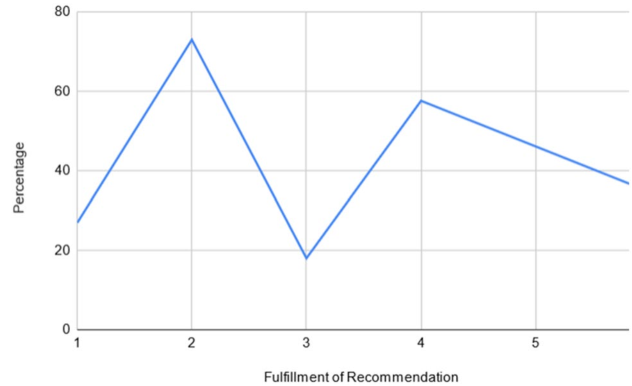

Fig. 4 Percentage fulfilment of the six recommendations in the current analysis

Table 1 Tabular representation of analysis of works from 2017 to 2023 based on Brouwer's Recommendations

References	Year	1.1	1.2	2.1	3.1	3.2	3.3	3.4	4.1	4.2	4.3	5.1	5.2	6.1	6.2	%
Zheng et al. (2021)	2021	√		√	√	√	√			√	√	√	√		√	71.4
Cheng et al. (2020)	2021	√								√	√	√	√		√	42.8
Hasan et al. (2021)	2021			√		√	√			√	√					35.7
Kim et al. (2021)	2021	√		√	√	√	√		√	√	√	√	√	√	√	85.7
Saha et al. (2021)	2021	√			√	√	√		√	√	√				√	57.1
Doma et al. (2020)	2020			√					√	√	√	√	√		√	50.0
He et al. (2020)	2020	√		√	√					√	√		√	√		50.0
Hwang et al. (2020)	2020	√		√	√	√			√	√			√			50.0
Jin et al. (2020)	2020	√		√	√				√	√	√		√	√		57.1
Ozdemir et al. (2021)	2020	√		√	√								√	√		35.7
Mousavi et al. (2020)	2020	√		√	√					√	√			√		42.8
Nawaz et al. (2020)	2020	√		√	√	√	√	√	√	√	√	√	√		√	85.0
Wang et al. (2020)	2020	√		√		√	√		√	√	√		√	√	√	78.5
Qing et al. (2019)	2019	√		√	√	√	√		√	√	√		√		√	71.4
Pandey et al. (2019)	2019				√				√				√	√	√	35.7
Taran et al. (2019)	2019	√		√	√	√	√			√	√	√			√	64.2
Li et al. (2018)	2018	√	√		√	√	√			√	√	√	√		√	71.4
Moon et al. (2018)	2018	√	√	√	√	√	√	√		√	√		√		√	78.5
Alazrai et al. (2018)	2018	√	√		√	√	√		√	√	√	√	√		√	78.5
Becker et al. (2017)	2017	√	√	√	√	√	√	√	√		√	√	√		√	85.0
Pane et al. (2017)	2017	√	√			√	√			√	√	√	√	√	√	71.5
Meng et al. (2017)	2017	√		√	√				√	√	√			√		50.0
Liu et al. (2014)	2021	√	√		√	√	√	√	√	√			√		√	71.4
Samavat et al. (2022)	2022	√		√	√	√	√			√	√	√			√	64.2
Sashi et al. (2022)	2022	√	√	√	√	√	√	√	√		√	√	√		√	85.0
Akhtar et al. (2022)	2022	√		√	√	√	√		√	√	√		√		√	71.4
% Fulfillment of key points		86.3	27.2	68.1	81.8	63.6	59	18.1	54.5	86.3	81.8	45.4	81.8	40.9	68.1	
% Fulfillment of recommendations		27.2		68.1	18.1				54.5			45.4		40.9		

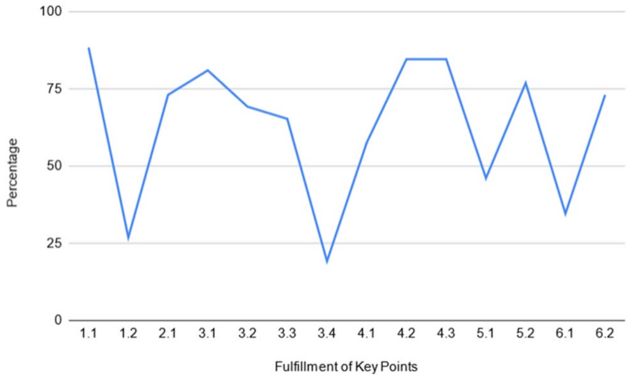

Fig. 5 Percentage fulfilment of the major key points in the current analysis

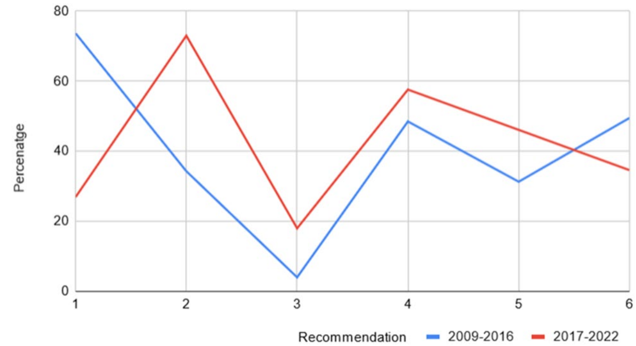

Fig. 6 Comparison of changing trend of recommendation fulfilment between previous and current analysis

of recommendation three that were fulfilled (all key points included) is quite low (19.2%). Using a properly designed experimental setup, 84.61% of the works attempt to remove confounds (Key Point 3.1). One example of this could be to provide habituation time to participants, i.e., they are made to rest after being subjected to particular stimuli to avoid overloading the brain with stimuli. Not as many researchers, however, tend to verify the presence and removal of remaining confounding factors (around 65% of papers satisfy 3.2 and 3.3). Since a lot of experimenters these days use predefined data sets with processed EEG data; they don't find it necessary to manually observe and remove confounds in the input data (very few, 19.2%, satisfy 3.4). Around 57.6% of research work follows recommendation four as a whole (Key Point 4.1 by 57.6%, 4.2 and 4.3 by more than 81%). An explanation for this observation can be found in the fact that sometimes detailed information about the dependency of classification data is not reported. Around 45.4 percent of works are in accordance with recommendation 5. Key point 5.1 is satisfied by 46.1 percent, and key point 5.2 is satisfied by a lot more (76.9 percent). This is because most researchers tend to extract various features from EEG signals without providing deeper insight into the result or why some features show better results than others. Around 34.6% of the works associated with the recommendation satisfy key point 6.1, while 73.07% satisfy key point 6.2.

Figures 6 and 7 show the differing trends between the previous and current analyses. These figures can be used to draw the following conclusions: fulfillment of recommendation one and both of its key points show a similar trend in the current analysis as in the previous analysis. After 2016, more researchers followed recommendation 2, i.e., use and provide a detailed hypothesis of the parameters that vary while carrying out the experiments. Recommendation 3, as a whole, does not show significant deviation; however, it is to be noted that more papers published after 2016 (2017–2023) are in accordance with recommendation 3, i.e., researchers have paid more attention to dealing with confounding factors in the experiment process. Recommendations 4 and 5, as a whole, show a similar trend. For recommendation 6, though it shows a similar trend, it is to be noted that works done between 2017 and 2023 have a lesser percentage of papers in compliance with this recommendation as compared to works ranging from 2009 to 2016. The reason for this could be that EEG is increasingly becoming the BCI standard, so authors tend to assume that EEG and its benefits are already well-known among research peers and do not need to be explained further.

3.2 Emotion Recognition from EEG

In this sub-section, results after performing a detailed study of works on the basis of five major factors have been presented. The factors are: test protocols, EEG recordings, artifact filtering, feature extraction, and the classification process.

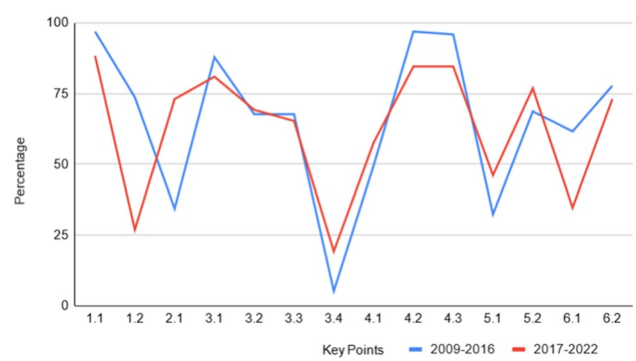

Fig. 7 Comparison of changing trend of key point fulfilment between previous and current analysis

3.2.1 Test Protocol Phase

The analysis of the test protocol phase is presented in Table 2 based on the type of stimulus applied, its duration, the number of subjects categorized according to gender, and the respective emotions displayed. Table 2 shows that a few papers were entered more than once; this is due to a single paper using multiple stimuli methods to record the elicited emotions. From the data presented in Table 2, the median, the range of the number of subjects, the percentage of subjects below the median, other gender-related percentages, and stimuli-related calculations were evaluated and tabulated in Table 3. The number of subjects taken during the study should be at least 30 to get good statistical data, along with the ratio of gender being balanced for the inconsistency between emotional experience and emotional expressivity (Deng et al. 2016). From Table 3, it can be seen that the median of the data increased from 15.0 to 29.5 in the current analysis. More works employ 30 (or so) subjects in their research. Previously, it was found that around 7% of the works only considered male subjects. However, no such papers were used in the current analysis. A noteworthy

Table 2 Tabular representation of analysis of works from 2017 to2023 based on Test Protocol Phase

References	Stimulus in min/s (duration)	#Subjects (female/male)	Emotions
Zheng et al. (2021)	CFAPS (33 ms)	19 (10/9)	Happiness, anger
Cheng et al. (2020)	DEAP (1 min)	32(16/16)	Arousal, valence, dominance
Cheng et al. (2020)	DREAMER (65–393 s)	32(9/14)	Surprise, entertainment, sadness, excitement, fear, happiness, calmness, anger, disgust
Hasan et al. (2021)	Music (60 s)	15(12/3)	Valence, arousal (Joyful, melancholic, neutral)
Kimel et al. (2021)	video clips (3 min)	30(12/18)	Happiness, sadness, rage, disgust, fear
Saha et al. (2021)	DEAP (60 s)	26(–/–)	Arousal, valence, dominance
Doma et al. (2020)	DEAP (1 min)	32(16/16)	Arousal, valence, familiarity, dominance
He et al. (2020)	Music (7 min)	20(10/10)	Positive, negative
He et al. (2020)	DEAP (1 min)	10(–/–)	Valence, arousal
Hwang et al. (2020)	SEED (4 min)	15(8/7)	Positive, negative, neutral
Jin et al. (2020)	DEAP (1 min)	32(16/16)	Arousal, valence
Jin et al. (2020)	SEED (4 min)	15(8/7)	Positive, neutral, and negative
Ozdemir et al. (2021)	DEAP (1 min)	32(16/16)	Dominance, Familiarity, Arousal, Liking, Valence
Mousavi et al. (2020)	Music (1 min)	16 (6/10)	Negative, positive, neutral
Nawaz et al. (2020)	DEAP (1 min)	32(16/16)	Dominance, Familiarity, Arousal, Liking, Valence
Wang et al. (2020)	SEED (4 min)	15(8/7)	Positive, neutral, and negative
Qing et al. (2019)	DEAP (1 min)	32(16/16)	Positive, negative, calm
Qing et al. (2019)	SEED (4 min)	15(8/7)	Positive, negative, neutral
Pandey et al. (2019)	DEAP (60 s)	32 (16/16)	Happy, sad
Taran et al. (2019)	Audio–video	20(–/–)	Happy, fear, sad, relax
Li et al. (2018)	DEAP (60 s)	32(15/17)	Arousal, Valence
Li et al. (2018)	SEED (1 min)	15(8/7)	Positive, negative
Moon et al. (2018)	DEAP (60 s)	32(16/16)	Valence
Alazrai et al. (2018)	DEAP (60 s)	32(16/16)	Valence, arousal
Becker et al. (2017)	Film (10 s)	27(5/22)	Anger, fear, disgust and sadness, amusement and tenderness
Pane et al. (2017)	DEAP (3 s – 12 s.)	32(16/16)	Happy, angry, relaxed, sad
Meng et al. (2017)	DEAP (60 s)	32(16/16)	Happy, fear, sad, relax
Li et al. (2014)	Videos (1–2 min)	29(7/22)	Surprise, anger, sadness, disgust, happiness, fear
Samavat et al. (2022)	DEAP SEED (01 min)	32 (16/16) 15 (8/7)	Valence, arousal, familiarity, like
Mokatren et al. (2021)	SAD DEAP	32	Arousal, valence
Sakalle et al. (2021)	DEAP/SEED (1/2 min)	50 (25/25)	Sadness, disgust, anger, surprise
Huang et al. (2021)	DEAP (01 min)	32 (16/16)	Liking, Arousal, valence, familiarity, dominance
Sashi et al. (2022)	DEAP (01 min)	32(–/–)	Valence, arousal, dominance
Akhtar et al. (2022)	DEAP (01 min)	32 (16/16)	Liking, valence, arousal, dominance

Table 3 Differing trends in current analysis as compared to previous analysis w.r.t the Table 2

Parameters	Previous analysis	Current analysis	Overall analysis
Number of subjects taken for study	1–161	7–32	1–161
Median	15	29.5	19.5
Below median %	47%	39.1	51%
At least 30 subjects	27%	60.80%	47.70%
More males than females	68%	21.70%	66%
Equal no. of male female subjects	9.50%	47.80%	22.50%
Average duration	57.1 s	113.15 s	76.27 s
Range of duration	15–180 s	0.3–420 s	0.3–420
DEAP (60 s)	21%	69.50%	45.50%
SEED (240 s)	0%	21.70%	10.80%

inference to be drawn from Table 3 is that the percentage of works using more male than female subjects decreased from 68% in the previous analysis to 21.7% in the current analysis. It should also be noted that more papers published after 2016 have a diverse range of subjects. Overall, the three papers in the current analysis do not specify the gender-wise distribution of subjects of the experiment. In the predefined dataset DEAP, the age of participants is between 19 and 37, with 26.9 being the mean (Koelstra et al. 2011). This is in accordance with Levenson et al. (1991), who state that autonomic nervous system (ANS) activity is better observable in younger people than in the elderly.

The currently taken stimulus are event-elicited emotions which are usually selected in Table 2 that encompass arousal levels and valence states. Stimuli of visual nature are more difficult to classify efficiently than audio-visual and visual types of stimuli (Bos 2006). This is the reason for consideration studies that mostly involve audio or audio-visual types of stimuli for analysis. Predefined data sets [DEAP and DREAMER] elicit emotions through various film clips with visual and auditory cues (Koelstra et al. 2011; Katsigiannis and Ramzan 2017). A review conducted by Al-Nafjan et al. (2017) shows that most experiments in this field used visual forms of stimulation for their participants, which might result in compromised efficiency. The time used to record the stimulus should be enough to elicit the emotions but not be too long to make the subject conscious of his/her own emotions; for this, Suhami et al. (2020) found that stimulation should last between 15 and 20 s. However, from Table 3, it can be observed that in recent experiments, the average duration has increased from 57.1 to 113.15 s, resulting in an overall increase of 76.2 s. This is the result of an increase in usage of the DEAP and SEED databases, with stimulus durations of 60 s and 240 s, respectively. DEAP usage increased from 21 to 69.5%, and SEED usage increased from 0 to 21.7%. From Table 2, it can be noted that 30.4% of studies have worked on happy or happiness emotions in the years 2017–2023, compared to 48.3% in the previous year's analysis, making an overall percentage of 45. Similarly, sadness (34.7%/62.1%/56.27%), anger (17.3%/44.8%/36%), fear (21.7%/44.8%/34%), surprise (13%/27.6%/18%), and disgust (17.3%/24.1%/20.4%) were also high. A decline in the percentage of the above-stated emotions because of the relatively small sample size of review papers can be noted and the percentage of works done on neutral emotion has increased from 13.8 to 21.7% in recent years, making an increase of 4.2%. This could be because researchers now place a greater emphasis on neutral emotion than in previous years.

3.3 EEG Recording Phase

From Sect. 2.1, it can be deduced that electrode positioning plays a vital role in experiments that use EEG signals for emotion recognition. The corresponding findings have been presented in Table 4 for various papers that have been taken in to consideration for study of this work.

Table 4 shows that the current analysis employs eight different pieces of equipment in various works. Table 5 shows that Biosemi Active 2 is still the most popular piece of equipment in current use. It can also be noted that higher sampling frequencies (i.e., 1000 Hz) have been employed in conjunction with the frequencies used in the previous analysis. However, only a few works (21.7%) mention the use of EEG equipment. This can be explained by the fact that they employ a predefined, publicly available dataset to test their model and hence do not require collecting sample data from scratch. Most of the studies contain details on the usage and positioning of electrodes. Around 8.6% of the works in the current analysis omit this information, compared to 11.6% previously.

3.3.1 Artifact Filtering Phase

As mentioned in Sect. 2.3, artifact filtering is often skipped by researchers. The explanation for this could be the fact that most papers use pre-processed data sets (shown in Sect. 3.2.1) and hence do not need the manual filtering step.

Table 4 Tabular representation of analysis of works from 2017 to 2023 based on EEG equipment used

Ref	Equipment (frequency)	Electrode location (#)
Zheng et al. (2021)	ERP neuroscan system	AF, AF3, C1, HEO, M2, C2, M1, C3, CB1, C5, C4, CZ, C6, CB2, CP1, F2, CP2, F3, CP3, F4, CP4, CP5, CP6, CPZ, F1, F5, F6, F7, F8, FZ, FC1, FC2, FC3, FC4, FC5, FC6, FCZ, FP1FP2, FPZ, FT7, FT8, O1, O2, OZ, P1, P2, P3, P4, P5, P6, P7, P8, PO3, PO4, PO5, PO6, PO8, POZ, TP7, TP8, T8, T7, TP8, TP7
Cheng et al. (2020)	–	AF4, AF3, C3, T7, FP2, C4, CZ, P8, CP2, O2, P3, CP1, CP5, F7, CP6, F3, F4, F8, FC2, T8, FC1, FC6, PO3, FC5, FP1, O1, OZ, P7, P4, PZ, PO4
Hasan et al. (2021)	EMOTIV EPOC (2048) Hz)	AF4, AF3, F3, F4, F7, F8, FC5, FC6, O1, O22, P7, P8, T7, T8
Kim et al. (2021)	EMOTIV EPOC (256 Hz)	AF3, AF4, F3, F4, F7, F8, FC5, FC6, O2, O1, P8, P7, T7, T8
Saha et al. (2021)	–	(FF3-F4) *, (CP1–CP2) *, (P3–P4) *, Oz, Pz*, FC6, PO4, F3, C3, O2
Doma et al. (2020)	Biosemi Active Two (512 Hz)	C3, PO4, F3, CP1, F4, C4, AF3 (7)
He et al. (2020)	EMOTIV EPOC (128 Hz)	–
Hwang et al. (2020)	ESI NeuroScan(1000 Hz)	AF4–CB2, AF3, CB1, F1–P1, F4–P4, F8–P8, F6–P6, F5–P5, F7–P7, F2–P2, F1–P1, F3–P3, FC1–CP1, FC2–CP2, FC3, CP3, FC4–CP4, FC5–CP5, FC6–CP6, FCZ–CPZ, FP2–O2, FP1–O1, FT8–TP8, FT7–TP7, FPZ–OZ, FZ–PZ (23)
Jin et al. (2020)	10–20 system	Fp1–AF3, FP1–FC2, FP1–Oz
Ozdemir et al. (2021)	10–20 electrode placement and using 512 Hz sampling frequency	Via AEP technique
Mousavi et al. (2020)	Encephalan 21-Channel EEG recorder (250 Hz)	F3, F4, F7, F8, FZ, FP2, FP1, T4
Nawaz et al. (2020)	BiosemiActiveTwo(256 Hz)	AF3, AF4, F3, F4, F7, F8, FC6, FC5, O2, O1, P7, P8, T8, T7(14)
Wang et al. (2020)	EEG cap (200 Hz)	C6, C5, CP5, CP6, F5, F6, F7, F8, FC3, FC6, FC4, FC5, FT8, FT7, P7, P8, T8, T7, TP8, TP7(20)
Qing et al. (2019)	ESI NeuroScan System (1000 Hz)	DEAP:AF4, AF3, C3, C4, CZ, CP2, CP1, CP6, CP5, F4, F7, F8, FCS, FC6, FC2, FC1, P3, P4, PZ, P8, P7, PO4, PO3, T7, T8, O2, O1, OZ SEED:AF3, CS, CZ, C1, C3, C4, C6, CB1, CB2, CP1, CP2, CP4, CP6, CPZ, F1, F2, F3, F4, F5, F6, F7, F8, FC1, FC2, FC3, FC4, FC6, FCZ, FZ, FP1, FP2, FPZ, FT7, FT8, FCS, O1, O1, OZ, P1, P2, P3, P4, P5, P6, P7, P8, PO2, PO3, PO4, PO5, PO6, PO8, POZ
Pandey et al. (2019)	EEG Cap and Bio Semi	Fp2, Fp2 + F4, Fp, F4, F3
Taran et al. (2019)	–	F4, F3, F8, F7, FP2, FP1, T6, T4, T3, T5
Li et al. (2018)	–	C5, C6, FC5, FC6, FT8, F17, T7, TP7, TP8, T8, P8
Moon et al. (2018)	–	Fp1
Alazrai et al. (2018)	BiosemiActive Two system	AF3–AF4, F7–F8, F3–F4, FC5–FC6, FC1–FC2, FP1–FP2 (Frontal region) P3–P4, CP5–CP6, P7–P8 (Parietal region) O1–O2 (Occipital region) T7–T8 (Temporal region)
Becker et al. (2017)	Emotiv EPOC with 14 sensors	–
Pane et al. (2017)	Biosemi Active Two system	F7, C3, C4, O2, T7
Meng et al. (2017)	EEG cap (–)	F4, F3, FP2, FP1, PO4, PO3, T8, T7, O2, O1(10)
Liu et al. (2014)	EGI's GTEN 100(1000 Hz)	10–20 System (128)
Mokatren et al. (2021)	Biosemi	FCZ, IZ
Sakalle et al. (2021)	MUSE2 (512 Hz)	AF7, AF8, TP9, TP10, FP2
Huang et al. (2021)	– (512 Hz)	FZ, CZ, PZ, OZ, FP1–FP2, AF3, AF3–AF4, C3–C4

Table 4 (continued)

Ref	Equipment (frequency)	Electrode location (#)
Sashi et al. (2022)	10–20 electrode placement and using 512 Hz sampling frequency	FP1, F4, F3, FP2
Akhtar et al. (2022)	512 Hz	FP1, F8, AF3, C4, F3, D4, F7, FC1, P04, P3, FZ, P03, FP2
Badajena et al. (2023)	NeuroMax32	*P3–O1, T4–T6, O1, T4, P3, T6*
Badajena et al. (2022)	NMX32	*P3–O1, T4–T6, O1, T4, P3, T6*

Table 5 Summarization of inferences w.r.t equipment drawn from Table 4

	Previous analysis	Current analysis	Overall analysis
Equipment used	17	8	25
Most used	Biosemi active two	Biosemi active two	–
Sampling frequencies	512 Hz (21.3%), 256 Hz (19.7%)	1000 Hz (13.6%), 256 Hz (9%)	1000 Hz, 512 Hz, 256 Hz

An analysis of works considering artifact filtering in the current papers is done in Table 6.

Table 7 summarises the conclusions drawn from Table 6 regarding the various methods of artifact filtering used in previous years' studies and experiments, recent years, and the overall trend. Physiological factors such as eye movement, blinks, line noise, muscle, and heartbeat are removed by methods like blind source separation (BSS) and independent component analysis (ICA), which account for 28.1% of previous works, 47.8% of recent works, and 37.9% of overall works. Around 30% of previous analysis and 17% of current analysis have their electrodes re-referenced using methods such as average mean reference (AMR), common average reference (CAR), Laplacian (23.6 percent), and multiple artifact rejection algorithms (MARA) using the EEGLAB toolbox. Because collecting every frequency is pointless, the Bandpass frequencies filter is applied to 84% of previous works, 86.9% of recent works, and 85.4% of the overall analysis. Almost 44% of the works had down-sampled their original EEG signals in the previous analysis, compared to 50.3% in the current analysis. A rather recent study presented by Teplan (2002) provides deeper insight into the various artifacts and their filtration techniques. It concludes that every artifact filtering method has its own shortcomings, and no single method can guarantee the optimal removal of all artifacts at once. Thus, researchers are still working on identifying more techniques that are suited for a wider range of scenarios.

3.3.2 Feature Extraction Phase

This phase is an important phase as it involves find out major features that may be helpful for using a machine learning model. In Table 8, the feature extraction-related data has been summarized and its reference has been provided.

From Table 9, it can be seen that only 17% of respondents provide no details at all. The remaining 78.2% mostly used the δ, θ, α, β, and γ bands. Almost half of these papers (47%) studied above utilize all of the δ, θ, α, β, and γ bands combinedly. The event-related desynchronizations (ERD/ERS), event-related potentials (ERP), and fixed frequency bandwidths were the remaining characteristics utilized (e.g., 0.5–30 Hz, 1–10 Hz, 1–46 Hz, and 2–30 Hz). Due to a minor decline in recent times, the overall band's utilization percentage has reduced somewhat. In recent years, the percentage of people who have no information has risen. Perhaps the authors are unable to relate to it or are uninterested in including it or considering it. There were 42 distinct approaches utilized in the works examined. More than 65% of the works employed many methods, yet only one was chosen as the best in the end. Some of the approaches used in our use the techniques as: HOC (8%), statistics (17.3%), PSD (26%) and the rest use techniques like STFT, TDF, MCMR, PLV, PLI, ERD/ERS, QTFD. Despite the fact that the total number of methods used has remained constant, several new methods have been introduced in recent years, increasing the total number. The use of many methods for feature extraction has become increasingly popular. Perhaps the authors are utilizing many methods to cross-check the results, reducing the error margin to near zero.

3.4 Emotion Classification

There have been several classifiers in use and it is sometimes difficult infer accuracy of each of the classifier in different scenarios. Therefore, this analysis focuses only on the most used classifiers, the two types of classifiers (i.e., offline vs. online), and the data being used to educate and evaluate the classifiers (i.e., whether it is real-time or not). This data is well tabulated in Table 10.

Table 6 Tabular representation of analysis of works from 2017 to 2023 based on Artifact Filtering

References	Artifact filtering
Zheng et al. (2021)	Bandpass filter of 0.05–80 Hz; direct removal methods; translation of raw data to 512 Hz
Cheng et al. (2020)	Down-sampled 128 Hz; using blind source separation techniques like ICA to remove EOG artifacts; linear phase FIR filters to remove eye artifacts; segmentation
Hasan et al. (2021)	Down sampled to 128 Hz; bandpass frequency filter with a cut off frequency (2–42 Hz); independent component analysis; multiple artifact rejection algorithm (MARA); visual inspection to remove the remaining artifacts; EEGLAB toolbox
Kim et al. (2021)	Filtering bandpass frequency (0.5–30 Hz); ICA
Saha et al. (2021)	Down sampling of data to 128 Hz sampling frequency;4–45 Hz cutoff frequency bandpass filter; blind source separation technique to remove eye movement artifacts
Doma et al. (2020)	Hilbert-Huang Transform (HHT)
He et al. (2020)	Bandpass filtering with 4–45 Hz to tackle noise; butterworth bandpass filter of cut-off frequency 0.5–50 Hz; sampled with 128 Hz frequency; notch filter to deal with powerline; sampling with 128 Hz; suppression of electrooculogram artifacts;50 Hz for removal of baseline wandering and high-frequency artifacts
Hwang et al. (2020)	Five band-pass filters with five frequency bands: Alpha (8–13 Hz), Gamma (31–50 Hz), Beta (14–30 Hz), Theta (4–7 Hz), Delta (1–3 Hz); Short Time Fourier Transform (STFT)
Jin et al. (2020)	DEAP: down sampling to 128 Hz; filtering of 4-45 Hz of data to eliminate EOG; Self-assessment to lessen the number of noisy channels SEED: 0-75 Hz bandpass filter; Down sampling to 200 Hz; Categorical labels minimize noisy channels
Ozdemir et al. (2021)	The bandpass filter of the cut-off frequency of 4–45 Hz: down sampling to 128 Hz; elimination of EOG artifacts; segmentation and labeling were performed
Mousavi et al. (2020)	1st order low pass Butterworth filter of frequency 0.5–45 Hz; 50 Hz of power supply removed by Notch filter; Data normalization in between 0.1
Nawaz et al. (2020)	The bandpass filter of 4–45 Hz frequency
Wang et al. (2020)	Visual inspection; filtration of EEG signal in 0.3–49 Hz to eliminate high-frequency noise and power frequency interference (frequency filter)
Qing et al. (2019)	0–75 Hz Bandpass filter; raw data down sampled to 200 Hz
Pandey et al. (2019)	Raw data down sampled to 200 Hz; Independent component analysis (ICA)
Taran et al. (2019)	Correlation-criterion used for removal of noisy intrinsic mode functions (IMF); IF based filtering of VMD mode; Laplacian Filter; STFT and WT
Li et al. (2018)	Hanning window Bandpass filter; raw data down sampled; removing electromyogram (EMG) and electrooculogram (EOG); finite impulse response (FIR)
Moon et al. (2018)	Bandpass frequency filter with a bandwidth of ten frequency band: alpha (8–12 Hz), high-alpha (10.5–12 Hz), low-alpha (8–9.5 Hz), beta (13–29 Hz), high-beta (21–29 Hz), mid-beta (17–20 Hz), low-beta (13–16 Hz), delta (0–3 Hz), gamma (30–50 Hz), theta (4–7 Hz)
Alazrai et al. (2018)	Down sampled to 128 Hz; blind source separation technique, (EMG) to remove electrooculogram (EOG); bandpass frequency filter with a bandwidth of 4–45 Hz to reduce electromyogram (EMG); CAR applied to EEG signals
Becker et al. (2017)	5th order Butterworth filter to filter Bandpass between 2 and 80 Hz; independent component analysis (ICA) with Wavelet Denoising; rejecting ICA components
Pane et al. (2017)	Down sampling to 128 Hz; bandpass frequency filter, blind source separation technique, with a bandwidth of 4–45 Hz, and Bandpass IIR filter with Chebyshev type 2 window
Meng et al. (2017)	De-noising and down sampling
Liu et al. (2014)	Filter between 0.1 to 100 Hz; keeping Net Amps high-pass band frequency filter with a cut-off frequency of 0.1 Hz
Samavat et al. (2022)	Down sampled to 200 Hz using band pass filter of 0–75 Hz
Mokatren et al. (2021)	Down sampling of dataset using 128 Hz frequency band with sampling frequency 0–45 Hz
Sakalle et al. (2021)	Down sampling of data using 256 Hz frequency and sampling frequency 45–64 Hz, Notch filter used
Huang et al. (2021)	Down sampling of data using 128 Hz frequency and sampling frequency 4.0–45.0 Hz
Sashi et al. (2022)	EOG artifact filtering using 128 Hz down sampling frequency with bandpass filter 4–45 Hz
Akter et al. (2022)	Down sampling of data to 128 Hz sampling frequency; 4–45 Hz cutoff frequency bandpass filter; blind source separation technique to remove eye artifacts

Although different classifiers have been used in different research scenarios, it is highly recommended to concentrate on a single type of classifier for research purposes. This work is around analyzing only the most common classifier, i.e., the support vector machine (SVM) classifier, along with its different kernels and other commonly used classifiers.

Table 7 Differing trends in current analysis as compared to previous analysis w.r.t Table 6

	Previous analysis	Current analysis	Overall analysis
Other physiological factors	28.10%	47.80%	37.90%
Re-referencing the electrodes	30%	17%	23.50%
Bandpass frequency filter	84%	86.90%	85.40%
Notch filter	16.5% (50 and 60 Hz)	8.6% (50 Hz)	12.59%
Down sampling their original EEG signals	43.90%	50.30%	47.10%

Table 8 Tabular representation of analysis of works from 2017 to 2023 based on feature extraction

References	EEG features	Feature extraction
Zheng et al. (2021)	Theta, alpha, beta, and gamma	HFD, MPE, DWT
Cheng et al. (2020)	N/A	FFT
Hasan et al. (2021)	Theta, alpha, and beta	MI
Kim et al. (2021)	Delta, theta, alpha, beta, and gamma	WPT, PSD, STFT, DE, FD, SM
Saha et al. (2021)	Theta, alpha, beta, and gamma	PSD, DT-CWPT
Doma et al. (2020)	Delta, theta, alpha, beta, and gamma	FFT, PCA
He et al. (2020)	N/A	FIOA
Hwang et al. (2020)	Delta, theta, alpha, beta, and gamma	STFT, DE
Jin et al. (2020)	Delta (0–4 Hz), theta (4–7.5 Hz), alpha (7.5–12.5 Hz), beta (12.5–30 Hz), and gamma (30–40 Hz)	62 × 62 channel wise feature for SEED, 32 × 32 channel wise feature for DEAP
Ozdemir et al. (2021)	Alpha (1–13 Hz), beta (14–30 Hz), gamma (31–50 Hz)	DNN, CNN, HCNN, SAE, DLN, CNN + LSTM
Mousavi et al. (2020)	N/A	
Nawaz et al. (2020)	Delta, theta, alpha, beta, and gamma	DASM, PE, SE, SVDE, FD, RASM and Statistical
Wang et al. (2020)	Alpha, beta, and gamma	DE, Statistical
Qing et al. (2019)	Delta, theta, alpha, beta, and gamma	Statistical
Pandey et al. (2019)	Delta, theta, alpha, beta, and gamma	EMD
Taran et al. (2019)	SE, TE, HFD, HE	STFT, WT, FAWT, TQWT
Li et al. (2018)	Theta rhythm (4–7 Hz), alpha rhythm (8–15 Hz), beta rhythm (16–31 Hz), and gamma rhythm (> 32 Hz)	ERD/ERS (43.5–94.5 Hz), DASM, PSD, The filter-based methods. The RFE-based method
Moon et al. (2018)	Delta (0–3 Hz), theta (4–7 Hz), low alpha (8–9.5 Hz), high alpha (10.5–12 Hz), alpha (8–12 Hz), low beta (13–16 Hz), mid beta (17–20 Hz), high beta (21–29 Hz), beta (13–29 Hz), and gamma (30–50 Hz)	320 PSD (32 channels × 10 frequency bands) and PLV, PCC, PLI
Alazrai et al. (2018)	N/A	QTFD, statistical, HOC, PSD, DWT, CWD-based TFR
Becker et al. (2017)	θ, α, β, low γ, and high γ frequency bands, and SCF features	HOC features, FD, statistics, spectral moments, SCF
Pane et al. (2017)	Theta (4–8 Hz), alpha (8–12 Hz), beta (12–30 Hz), and gamma (30–45 Hz)	FDF, TDF, PSD
Meng et al. (2017)	Delta, theta, alpha, beta, and gamma	MCMR, MI
Liu et al. (2014)	Delta, theta, alpha, beta, and gamma	STFT, DE
Salankar et al. (2021)	–	EMD, SF
Huang et al. (2021)	–	WT, WE, AE, SE, FE
Subasi et al. (2021)	Gamma, beta, alpha, theta, and delta	TQWT, SF
Topic et al. (2021)	Delta (0–4 Hz), theta (4–8 Hz), alpha (8–16 Hz), beta (16–32 Hz), and gamma (32–64 Hz)	DWT, DE, FD, PSD, HP
Mokatren et al. (2021)	Gamma, beta, alpha, theta, and delta	WPD
Akter et al. (2022)	Theta, low beta, alpha, high beta, gamma	FFT, FE

From Table 11, it can be seen that out of 29 different classifiers, 78% of the time, SVM is being used, with 59% from 26 classifiers (2009–2016) and 98% from 17 different classifiers. The SVM classifier is a supervised learning method based on kernels that divide data into multiple groups. SVM has grown in popularity in recent

Table 9 Differing trends in current analysis as compared to previous analysis w.r.t Table 8

	Previous analysis	Current analysis	Overall analysis
Bands used%	89.40%	77.20%	86.80%
No information%	10%	18%	12.60%
Methods used	42	42	56
More than one%	> 47.6%	> 65%	> 53.3%
PSD	22.20%	26%	24.10%
STFT	25.40%	17.30%	23.10%
Statistical	23.80%	17%	20.90%

years because it works well when there is a clear dividing line between classes. SVM is more effective when the data is in high dimensional space, and the dimensions are greater than the number of samples (Nygaard et al. 2003). It is noteworthy that the larger the sample size, the higher the cost of experimentation, and a larger sample size will necessitate a greater number of subjects. In the previous analysis, the median number of subjects was 19. So, clearly, the authors have not taken a larger sample space to start with; in this case, authors have considered using SVM as the best choice. SVM, using RBF as its kernel, has been used in 29.7% of the work in the years 2009–2016 and 8.6% of the work in 2017–2023. In recent years, most of the works have not specified the name of the kernel used, and therefore, this is the reason for the dramatic decrease in the percentage of RBF kernel used in the studies.kNN, or k-nearest neighbours, is a pattern recognition algorithm that identifies objects in feature space based on the nearest training instances. kNN was used in 14% of the works from 2009 to 2016, increasing to 27.7% from 2017 to 2023, for an overall percentage of 16.5 in the studies from 2009 to 2023. Because kNN is a lazy learner, it has no training period and works best for real-time predictions because it saves the training dataset and only uses it when needed, generating forecasts in real time to train itself. It is faster than those that require training, making it the second most popular classifier over the years. LDA being used in 6.3% of the 2009–2016 works saw a rise up to 8.6% of the 2017–2023 works, for an overall percentage of 7.45.

Another aspect of the classification process is whether the dataset was created online or offline. It is already clear that EEG signals are always changing over time. In order to capture the changes in frequency, a dynamic dataset with frequent updates is highly essential. For this, the data has to become available in real time in a sequential manner. So, if the data is recorded in a real-time manner, then it is called "online learning. Real-time data can be both dynamic and static. Offline learning occurs when data remains static over time and is recorded over an instance.

Table 11 also shows that in previous works (2009–2016), 90% of the work was completed offline, with only 8% completed online. In recent years, from 2017 to 2023, the percentage of offline work has increased to 95.5%, and online work has degraded to nil. The majority of the work has been done offline because most of the classification has been built on the notion of the stationary distribution. Dynamic dataset classification can be done using kNN models (Rani et al. 2006).

The last essential part of the classification procedure is whether it is user-dependent or user-independent. If the classifier has been trained by creating a new model for each of the test subjects and user data in the testing phase, then it is called "user-dependent data. If multiple user data sets are employed for both training and testing of the classifier, then they are called "user-independent data. It can be inferred that creating user-independent data is much easier because the creation of a new model with a new test subject is not needed, and the same model can be used for new users.

Table 11 also shows that in the works from 2009 to 2016, 46.8% used user-independent data, and 43.5% used user-dependent data. While there are 8% of the works have used both independent and dependent data, while working on projects from 2017 to 2023, the percentage of user-independent work has increased to 66%.

4 Conclusion

This work carried out in here makes a critical discussion of some of the earlier work in the time frame of 2017–2023. Several important factors associated with this research work were improved and have been highlighted. In the first case, the gender ration of participating subjects seemed to improve to a more balanced pattern which is shown in Table 2. In the same light, there was substantial improve in the average stimulus duration. In addition to this, the efficiency of filters applied for data retrieval was also improved which is shown in Table 6. Since mere machine learning techniques are not enough to produce genuine throughput approaches like CNN, RNN, DBN, etc. can be taken as another alternative for producing effective results. An analysis of such algorithms has been provided in the Table 10. Further, it can be concluded that while researchers are moving towards making more efficient models, emotion recognition in real-time still poses a big challenge. It can be observed that most of the recent studies use EEG data generated over a very short period, a few seconds to be precise, using an audio-visual stimulus. However, to build a more implementable model, it is expected to recognize emotions for a longer duration of time, preferably using virtual reality simulation. Commonly, researchers rely on just EEG signals for their experiments, but clubbing EEG signals with other

Table 10 Tabular representation of analysis of works from 2017 to 2023 based on classification

References	Off/on	User	Classifier	Results
Zheng et al. (2021)	Offline	Ind	SVM	91.35% all
Cheng et al. (2020)	Offline	Dep	SVM, KNN, DT, MLP	DEAP-97.69% valence and 97.53% arousal; DREAMER-89.03% valence, 90.41% arousal, 89.89% dominance
Hasan et al. (2021)	Offline	Ind	SVM, NN	90%
Kim et al. (2021)	Offline	Dep	SVM RBF, LSTM	Dependent: 91.5%; independent: 89.6%
Saha et al. (2021)	Offline	Ind	SVM RBF	Arousal-54.31% LK and 52.96% RDLK; valence-54.00% LK and 52.91% RDLK
Doma et al. (2020)	Offline	Ind	SVM, kNN, LDA	between 55 and 75%
Hwang et al. (2020)	Offline	Ind	SVM, KPCA, TCA, DNN	75.31%
Jin et al. (2020)	Offline	Dep	SVM	DEAP-98.93% valence, 99.10% arousal; SEED-99.63% over three-class emotion classification (positive, neutral, negative)
Ozdemir et al. (2021)	Offline	Ind	CNN	Negative and positive Valence—90.62%, high and low Arousal—86.13%, high and low Dominance—88.48%, and like/unlike—86.23%
Sheykhivand et al. (2020)	Offline	Dep	SVM	The proposed algorithm's simulation results for three-stage classification (negative, neutral, and positive) and two-stage classification (negative and positive) of emotion for 12 active channels revealed 97.42% and 96.78% accuracy, respectively
Nawaz et al. (2020)	Offline	Ind	SVM, kNN, DT	78.96 percent arousal; 77.62percent valence; 77.60% percent dominance
Liu et al. (2020)	Offline	Ind	DECNN	97.56%
Qing et al. (2019)	Offline	Dep	DT, KNN, RT	DEAP-63.09%; SEED-75%
Pandey et al. (2019)	Offline	Ind	SVM, ELM	58.5%
Taran et al. (2019)	Offline	Ind	SVM, KNN, MLP, ELM	90.63%
Li et al. (2018)	Offline	Ind	SVM and random forest (RF)	59.06% (AUC = 0.605) on the DEAP dataset, 83.33% (AUC = 0.904) on the SEED dataset
Moon et al. (2018)	Offline	Ind	CNN = 10, SVM	55.31% = High valence, 44.69% = low valence, maximum accuracy = 80.86% (CNN = 5, PSD), The best performance = 99.72% (CNN = 5, PLV), 55.42%. = SVM
Alazraiet al. (2018)	Offline	Ind	GELM, SVM, LSTM	Between 73.8 and 86.2%
Becker et al. (2017)	Offline	Ind	SVM, KNN, LDA, LSVO	> 90%
Pane et al. (2017)	Offline	Dep	SVM, k-NN, NN	92.01%, average accuracy of 81.64%
Meng et al. (2017)	Offline	Ind	CB	68.4% happy, 63.8% fear, 61.1% sad, 58.7% relax
Liu et al. (2014)	Offline	Dep	SVM, kNN, DANN	84.51% neutral, 71.71% sadness, 65.37% fear, 86.12% happy, 67.99% anger, 64.42% disgust, 84.07% surprise
Salankar al. (2021)	Offline	Ind	MLPNN	93.8%
Huang et al. (2021)	Offline	Ind	SVM	85.9%
Subasi al. (2021)	Offline	Ind	RFE, SVM	93%
Topic et al. (2021)	–	–	CNN, SVM	DEAP (Arousal: 77.7, Valence: 76.6) DREAMER (Arousal: 90.4, Valence: 88.2) AMIGOS (Arousal: 90.5, Valence: 78.4) SEED (88.5)
Mokatren et al. (2021)	–	Dep/Ind	CNN, SVM	Arousal-93.66, Valence: 94.48 (Dependent)-CNN Arousal-91.06, Valence: 91.85 (Independent)-CNN Arousal-83.04, Valence: 85.31 (Dependent)-SVM Arousal-81.43, Valence: 83.38 (Independent)-SVM
Sakalle et al. (2021)			LSTM	For four class of emotion classification accuracy as 83.12%, 86.94%, 91.67%, and 94.12% for 50–50, 60–40, 70–30, and 10-fold cross-validations For three class of emotions LSTM based deep learning model provides classification Accuracy as 81.33%, 85.41%, 89.44%, and 92.66% for 50–50, 60–40, 70–30, and tenfold cross-validation

Table 10 (continued)

References	Off/on	User	Classifier	Results
Huang et al. (2021)	Offline	Dep/Ind	BiDCNN	4 class: 94.12, 3 class: 92.66 Arousal-94.72, Valence: 94.38 (dependent) Arousal-63.94, Valence: 68.14 (independent)

Table 11 Differing trends in current analysis as compared to previous analysis w.r.t Table 10

	Previous analysis	Current analysis	Overall analysis
Classifiers selected	26	17	29
SVM	59%	98%	78%
RBF (kernel of SVM)	29.70%	8.60%	
kNN	14%	27.70%	16.50%
LDA	6.30%	8.60%	7.45%
Offline%	90%	95.50%	94.20%
Online%	8%	–	4.30%
No. of both	1	1	2
Independent	46.80%	66%	52.20%
Dependent	43.50%	27.70%	38.50%
Both	8%	5.50%	6.20%

extractable physiological signals and behavioural measures can be a turning point in this field and provide extremely reliable output.

It can be believed that these shortcomings will be dealt with in the near future and hope that our study will prove to be helpful, especially for the new researchers in this field, by providing them with a structured starting point.

Author contributions All authors have made substantial contributions to the conception of the work and contributed to the acquisition, analysis, or interpretation of data.

Data availability Data Sharing not applicable as research data has been clearly mentioned in the text of the manuscript.

Declarations

Conflict of interest The authors have no competing interests.

Ethics approval and consent to participate Not Applicable.

Consent for publication Not applicable.

References

Akter, S., Prodhan, R.A., Pias, T.S., Eisenberg, D., Fresneda Fernandez, J.: M1M2: deep-learning-based real-time emotion recognition from neural activity. Sensors **22**(21), 8467 (2022)

Alarcao, S.M., Fonseca, M.J.: Emotions recognition using EEG signals: a survey. IEEE Trans. Affect. Comput. **10**(3), 374–393 (2017)

Alazrai, R., Homoud, R., Alwanni, H., Daoud, M.I.: EEG-based emotion recognition using quadratic time-frequency distribution. Sensors **18**(8), 2739 (2018)

Al-Nafjan, A., Hosny, M., Al-Ohali, Y., Al-Wabil, A.: Review and classification of emotion recognition based on EEG brain-computer interface system research: a systematic review. Appl. Sci. **7**(12), 1239 (2017)

Badajena, J. C., Sethi, S., Dash, A., Rout, P., Sahoo, R. K. Smart wheelchair using brain waves through machine learning. In: Intelligent systems: proceedings of ICMIB 2021, pp. 445–457. Singapore: Springer Nature Singapore (2022)

Badajena, J. C., Sethi, S., & Sahoo, R. K.: Comparative analysis of full training set and cross-validation for machine learning approach to run smart wheelchair. In: Intelligent Systems and Applications: Select Proceedings of ICISA 2022, pp. 121–132. Springer Nature Singapore (2023)

Becker, H., Fleureau, J., Guillotel, P., Wendling, F., Merlet, I., Albera, L.: Emotion recognition based on high-resolution EEG recordings and reconstructed brain sources. IEEE Trans. Affect. Comput. **11**(2), 244–257 (2017)

Bekkedal, M.Y., Rossi, J., III., Panksepp, J.: Human brain EEG indices of emotions: delineating responses to affective vocalizations by measuring frontal theta event-related synchronization. Neurosci. Biobehav. Rev. **35**(9), 1959–1970 (2011)

Bos, D.O.: EEG-based emotion recognition. Influence vis. Audit. Stimuli **56**(3), 1–17 (2006)

Brouwer, A.M., Zander, T.O., Van Erp, J.B., Korteling, J.E., Bronkhorst, A.W.: Using neurophysiological signals that reflect cognitive or affective state: six recommendations to avoid common pitfalls. Front. Neurosci. **9**, 136 (2015)

Cheng, J., Chen, M., Li, C., Liu, Y., Song, R., Liu, A., Chen, X.: Emotion recognition from multi-channel EEG via deep forest. IEEE J. Biomed. Health Inform. **25**(2), 453–464 (2020)

Deng, Y., Chang, L., Yang, M., Huo, M., Zhou, R.: Gender differences in emotional response: inconsistency between experience and expressivity. PLoS ONE **11**(6), e0158666 (2016)

Doma, V., Pirouz, M.: A comparative analysis of machine learning methods for emotion recognition using EEG and peripheral physiological signals. J. Big Data **7**(1), 1–21 (2020)

Ekman, P.: Basic Emotions. Hoboken (1999)

Electrode Position Nomenclature Committee: Guideline thirteen: guidelines for standard electrode position nomenclature. J. Clin. Neurophysiol. **11**, 111–113 (1994)
Goldstein, E. B. (Ed.).: Encyclopedia of Perception. Sage (2010)
GS, S. K., Sampathila, N., & Tanmay, T.: Wavelet based machine learning models for classification of human emotions using EEG signal. Meas. Sens. **24**, 100554 (2022)
Hasanzadeh, F., Annabestani, M., Moghimi, S.: Continuous emotion recognition during music listening using EEG signals: a fuzzy parallel cascades model. Appl. Soft Comput. **101**, 107028 (2021)
He, H., Tan, Y., Ying, J., Zhang, W.: Strengthen EEG-based emotion recognition using firefly integrated optimization algorithm. Appl. Soft Comput. **94**, 106426 (2020)
Homan, R.W., Herman, J., Purdy, P.: Cerebral location of international 10–20 system electrode placement. Electroencephalogr. Clin. Neurophysiol. **66**(4), 376–382 (1987)
Huang, C.: Recognition of psychological emotion by EEG features. Netw. Model. Anal. Health Inform. Bioinform. **10**(1), 1–11 (2021)
Huang, D., Chen, S., Liu, C., Zheng, L., Tian, Z., Jiang, D.: Differences first in asymmetric brain: a bi-hemisphere discrepancy convolutional neural network for EEG emotion recognition. Neurocomputing **448**, 140–151 (2021)
Hwang, S., Ki, M., Hong, K., & Byun, H.: Subject-independent EEG-based emotion recognition using adversarial learning. In: 2020 8th International Winter Conference on Brain–Computer Interface (BCI), pp. 1–4. IEEE (2020)
Jin, L., Kim, E.Y.: Interpretable cross-subject EEG-based emotion recognition using channel-wise features. Sensors **20**(23), 6719 (2020)
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. **22**(1), 98–107 (2017)
Kim, S., Yang, H.J., Nguyen, N.A.T., Prabhakar, S.K., Lee, S.W.: WeDea: a new EEG-based framework for emotion recognition. IEEE J. Biomed. Health Inform. **26**(1), 264–275 (2021)
Koelstra, S., Muhl, C., Soleymani, M., Lee, J.S., Yazdani, A., Ebrahimi, T., Patras, I.: Deap: a database for emotion analysis; using physiological signals. IEEE Trans. Affect. Comput. **3**(1), 18–31 (2011)
Lang, P.J.: The emotion probe: studies of motivation and attention. Am. Psychol. **50**(5), 372 (1995)
Levenson, R.W., Carstensen, L.L., Friesen, W.V., Ekman, P.: Emotion, physiology, and expression in old age. Psychol. Aging **6**(1), 28 (1991)
Li, X., Song, D., Zhang, P., Zhang, Y., Hou, Y., Hu, B.: Exploring EEG features in cross-subject emotion recognition. Front. Neurosci. **12**, 162 (2018)
Liu, Y., Sourina, O. EEG-based subject-dependent emotion recognition algorithm using fractal dimension. In: 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 3166–3171. IEEE (2014)
Liu, S., Wang, X., Zhao, L., Zhao, J., Xin, Q., Wang, S.: Subject-independent emotion recognition of EEG signals based on dynamic empirical convolutional neural network. In: IEEE/ACM Transactions on Computational Biology and Bioinformatics (2020)
Luck, S. J., Kappenman, E. S. (Eds.).: The Oxford Handbook of Event-Related Potential Components. Oxford University Press (2011)
Meng, Q., Yan, J., & Xu, H.: Research on EEG signal recognition based on channel selection. In: 2017 Chinese Automation Congress (CAC), pp. 6413–6417. IEEE (2017)
Mokatren, L.S., Ansari, R., Cetin, A.E., Leow, A.D., Ajilore, O.A., Klumpp, H., Vural, F.T.Y.: EEG classification by factoring in sensor spatial configuration. IEEE Access **9**, 19053–19065 (2021)
Moon, S. E., Jang, S., Lee, J. S.: Convolutional neural network approach for EEG-based emotion recognition using brain connectivity and its spatial information. In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2556–2560. IEEE (2018)
Nanthini, B.S., Santhi, B.: Electroencephalogram signal classification for automated epileptic seizure detection using genetic algorithm. J. Nat. Sci. Biol. Med. **8**(2), 159 (2017)
Naser, D.S., Saha, G.: Influence of music liking on EEG based emotion recognition. Biomed. Signal Process. Control **64**, 102251 (2021)
Nawaz, R., Cheah, K.H., Nisar, H., Yap, V.V.: Comparison of different feature extraction methods for EEG-based emotion recognition. Biocybern. Biomed. Eng. **40**(3), 910–926 (2020)
Nygaard, V., Løland, A., Holden, M., Langaas, M., Rue, H., Liu, F., et al.: Effects of mRNA amplification on gene expression ratios in cDNA experiments estimated by analysis of variance. BMC Genom. **4**(1), 1–13 (2003)
Ozdemir, M.A., Degirmenci, M., Izci, E., Akan, A.: EEG-based emotion recognition with deep convolutional neural networks. Biomed. Eng. **66**(1), 43–57 (2021)
Pandey, P., &Seeja, K. R.: Emotional state recognition with EEG signals using subject independent approach. In: Data Science and Big Data Analytics, pp. 117–124. Springer, Singapore (2019)
Pane, E. S., Hendrawan, M. A., Wibawa, A. D., Purnomo, M. H.: Identifying rules for electroencephalograph (EEG) emotion recognition and classification. In: 2017 5th International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME), pp. 167–172. IEEE (2017)
Plutchik, R.: The nature of emotions. Am. Sci. **89**(4), 344 (2001)
Posner, J., Russell, J.A., Peterson, B.S.: The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. Dev. Psychopathol. **17**(3), 715–734 (2005)
Qing, C., Qiao, R., Xu, X., Cheng, Y.: Interpretable emotion recognition using EEG signals. IEEE Access **7**, 94160–94170 (2019)
Rani, P., Liu, C., Sarkar, N., Vanman, E.: An empirical study of machine learning techniques for affect recognition in human–robot interaction. Pattern Anal. Appl. **9**(1), 58–69 (2006)
Sakalle, A., Tomar, P., Bhardwaj, H., Acharya, D., Bhardwaj, A.: A LSTM based deep learning network for recognizing emotions using wireless brainwave driven system. Expert Syst. Appl. **173**, 114516 (2021)
Salankar, N., Mishra, P., Garg, L.: Emotion recognition from EEG signals using empirical mode decomposition and second-order difference plot. Biomed. Signal Process. Control **65**, 102389 (2021)
Samavat, A., Khalili, E., Ayati, B., Ayati, M.: Deep learning model with adaptive regularization for EEG-based emotion recognition using temporal and frequency features. IEEE Access **10**, 24520–24527 (2022)
Sheykhivand, S., Mousavi, Z., Rezaii, T.Y., Farzamnia, A.: Recognizing emotions evoked by music using CNN-LSTM networks on EEG signals. IEEE Access **8**, 139332–139345 (2020)
Subasi, A., Tuncer, T., Dogan, S., Tanko, D., Sakoglu, U.: EEG-based emotion recognition using tunable Q wavelet transform and rotation forest ensemble classifier. Biomed. Signal Process. Control **68**, 102648 (2021)
Suhaimi, N. S., Mountstephens, J., Teo, J.: EEG-based emotion recognition: a state-of-the-art review of current trends and opportunities. Computat. Intell. Neurosci. (2020)
Sur, S., Sinha, V.K.: Event-related potential: an overview. Ind. Psychiatry J. **18**(1), 70 (2009)
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. **173**, 157–165 (2019)
Teplan, M.: Fundamentals of EEG measurement. Meas. Sci. Rev. **2**(2), 1–11 (2002)

Topic, A., Russo, M.: Emotion recognition based on EEG feature maps through deep learning network. Eng. Sci. Technol. Int. J. **24**(6), 1442–1454 (2021)

Walsh, P., Kane, N., Butler, S.: The clinical role of evoked potentials. J. Neurol. Neurosurg. Psychiatry. **76**(suppl 2), ii16–ii22 (2005)

Wang, Z., Zhang, Z., & Wang, W.: Emotion recognition based on framework of badeba-SVM. Math. Problems Eng. **2019** (2019)

Wioleta, S.: Using physiological signals for emotion recognition. In: 2013 6th international conference on human system interactions (HSI), pp. 556–561. IEEE (2013)

Zhao, G., Song, J., Ge, Y., Liu, Y., Yao, L., Wen, T.: Advances in emotion recognition based on physiological big data. J. Comput. Res. Dev. **53**(1), 80 (2016)

Zhao, G., Ge, Y., Shen, B., Wei, X., Wang, H.: Emotion analysis for personality inference from EEG signals. IEEE Trans. Affect. Comput. **9**(3), 362–371 (2017)

Zheng, X., Liu, X., Zhang, Y., Cui, L., Yu, X.: A portable HCI system-oriented EEG feature extraction and channel selection for emotion recognition. Int. J. Intell. Syst. **36**(1), 152–176 (2021)

Jenamani Chandrakanta Badajena received the M. Tech in Information Technology from College of Engineering and Technology, BPUT, Odisha India and is continuing doctoral work in Computer Science Engineering and Application at Indira Gandhi Institute of Technology, Sarang, Research Nodal center of Utkal University Bhubaneswar. He has published his research work in several national and international conferences and journals. His research interests include emotion recognition, brain computer interaction, human computer interaction, and network security. He has successfully guided several undergraduate and postgraduate students in their projects and final dissertation work.

Srinivas Sethi is currently working as Professor of Computer Science Engineering and Application, IGIT Sarang. He has been actively involved in teaching and research in computer science since 1997. He completed his Ph.D. in the area of Mobile Ad hoc Networks. His research interest includes wireless communication, wireless sensor networks, cognitive radio networks, IoT, BCI, HCI and cloud computing. He has been Volume Editor for Springer LNNS, International Conference Proceedings, board members for different journals, and Program Committee Member for different international conferences/workshops. He has published more than 80 research papers in various international journals and conference proceedings. He has also received grants for research projects from agencies such as DST, AICTE, NPIU, and DRDO.

Sanjit Kumar Dash is currently working as Assistant Professor in the Department of Information Technology in Odisha University of Technology and Research, Bhubaneswar Odisha, India. He obtained his Ph.D. degree in Computer Science Engineering from the Utkal University, Odisha, India. He has more than 15 years of teaching experience to UG and PG students. His research interests include Wireless Mobile Network, Wireless Sensor Network, and Machine Learning. He has published many international journals in the field on Mobile Data Offloading, Wireless Sensor Network, Cloud Computing.

Ramesh K. Sahoo has been Continuing Ph.D. in Computer Science, Department of Utkal University Bhubaneswar, India. He is currently working as Assistant Professor at the Indira Gandhi Institute of Technology, Sarang, India. His area of interest includes cognitive science, crowdsensing, data analytics, wireless sensor networks, IoT, and cognitive networks.