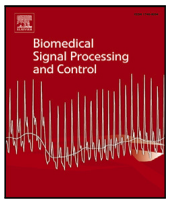
Biomedical Signal Processing and Control 106 (2025) 107674

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc



# Emotion recognition in virtual and non-virtual environments using EEG signals: Dataset and evaluation

Naseem Babu [a],*, Udit Satija [b], Jimson Mathew [a], A.P. Vinod [c]

[a] Computer Science & Engineering, Indian Institute of Technology Patna, Bihta, Patna, 801103, Bihar, India
[b] Electrical Engineering, Indian Institute of Technology Patna, Bihta, Patna, 801103, Bihar, India
[c] Infocomm Technology Cluster, Singapore Institute of Technology, 10 Dover Drive, 138683, Singapore

ARTICLE INFO

*Keywords:*
Electroencephalography (EEG)
Emotions recognition
Dataset
Virtual reality
Virtual environment
Virtual reality EEG (VREEG)
Frequency bands

ABSTRACT

Emotion recognition is essential in human interaction and understanding, significantly impacting cognition, perception, learning, communication, and decision-making. The classification of emotions from EEG signals has garnered interest due to its wide-ranging applications, from psychology to human–computer interaction. Although several publicly available emotion datasets based on EEG signals exist, these datasets often fail to introduce strong emotional stimuli in their recordings due to subject-specific behavior in response to class-distinct emotion video stimuli. This leads to inaccurate labeling and reduces applicability across different datasets due to inadequate model training. In this paper, we present a new EEG emotions dataset collected in a virtual reality environment (VREEG). The virtual reality setup employs a VR headset with an EEG cap to enhance the response to emotional stimuli during data collection. This data was collected from 70 subjects (50 males and 20 females), categorized into a 3-class subset (negative, neutral, and positive) named VREEG03 and a 4-class subset (happy, sad, neutral, and fear) named VREEG04. Differential Entropy (DE) and Power Spectral Density (PSD) features were extracted and analyzed across required frequency bands (delta, theta, alpha, beta, and gamma). The results demonstrated the highest classification accuracy of 77.75% with DE for VREEG03, and 62.67% with DE for VREEG04. The code and dataset are both publicly available at https://github.com/vreegemotions/VREEG_Datasets.

## Introduction

Recognizing and understanding emotions are fundamental to human interaction, influencing social connections, psychological health, and quality of life. Emotions shape individual experiences, guide decision-making, and impact interpersonal relationships, making emotion recognition a crucial field of study across various disciplines [1,2]. The capability to accurately detect and interpret emotions holds significant value in psychology, healthcare, education, and human–computer interaction fields. In this context, Brain–Computer Interfaces (BCIs) have become a powerful technology that enables direct interaction between the human brain and external devices. BCIs enable users to control machines or communicate without physical movement by translating neural activity into commands [3,4]. This technology has profound possibilities for individuals with disabilities, offering new ways for interaction and control. Advancements in machine learning and signal processing techniques have significantly propelled the advancements in BCI technology. These developments have enabled more accurate interpretation of complex brain signals, particularly those captured through electroencephalography (EEG) [5,6]. EEG-based BCIs are especially promising due to their non-invasive nature and ability to provide real-time insights into brain activities, and it has a range of applications, from medical diagnostics and neurorehabilitation to gaming and cognitive enhancement. The ability to decode emotional states from EEG signals is interesting, as it opens up possibilities for personalized user experiences, mental health monitoring, and even improved human–computer interaction. Emotion recognition has been an active area of research for several years. EEG-based emotion recognition differs from speech or face image-based studies as it directly captures neural activity, providing an internal measure of emotional states rather than relying on external expressions, unlike speech [7–9] or facial expressions [10], which can be consciously controlled or influenced by social context, EEG signals reflect more genuine responses, potentially offering a deeper and less biased view of an individual's emotions. This internal perspective makes it a unique and valuable modality for studying complex emotional states. There are so many studies exploring traditional approaches that often rely on handcrafted features

* Corresponding author.
*E-mail address:* naseem_2021cs22@iitp.ac.in (N. Babu).

https://doi.org/10.1016/j.bspc.2025.107674
Received 7 October 2024; Received in revised form 20 January 2025; Accepted 7 February 2025
Available online 25 February 2025

such as power spectral density, differential entropy [11–13], and other statistical measures extracted from these signals, and various classifiers are used to classify these different emotional states like happy, sad, neutral, and fear [14,15]. Recent advancements in deep learning have significantly enhanced emotion recognition through the integration of various neural networks such as graph neural networks [16], convolutional neural networks, and recurrent neural networks. Studies such as [17] introduced multi-head residual graph networks that used to enhance the model's ability to capture short and long-range dependencies. EEG signals are crucial in analyzing, present challenges for traditional feature extraction methods due to their complex nature, to address this, the tunable Q-factor wavelet transform with a constant-Q transform known for its invertibility and moderate oversampling has shown success in extracting meaningful information from signals [18]. Additionally, an optimal wavelet filter bank has been developed for automated diagnosis of Alzheimer's disease and emotions classification using EEG signals, demonstrating its effectiveness for capturing relevant spectral information [19,20]. The study [21] proposed an input variable selection method for nonlinear modeling using principal component analysis (PCA) and cluster analysis, highlighting the potential of combining wavelet decomposition with PCA for dimensionality reduction. This approach can be adapted to EEG-based emotion classification, where wavelet-transformed EEG signals benefit from PCA-driven feature reduction to enhance model performance. Advanced signal decomposition techniques have been adopted in recent studies on human emotion detection to enhance the analysis of non-stationary signals like EEG. Complete ensemble empirical mode decomposition with adaptive noise and local mean decomposition are commonly used to break down signals into intrinsic modes, effectively isolating patterns associated with various emotional states [22,23], multivariate variational mode decomposition further extends this approach to multivariate time-series data, enabling the decomposition of signals into multiple modes, which can help in separating emotion-specific components [24]. Time–frequency analysis techniques play an essential role in signal processing, offering insights into the time-varying characteristics of signals. Traditional methods such as the short-time Fourier transform, wavelet transform, and quadratic time–frequency transforms provide foundational approaches for analyzing these dynamics. Additionally, advancements in the field have led to the development of more refined methods, including adaptive time–frequency transforms and enhanced wavelet-based techniques. Recent methods, such as the Fourier-Bessel series expansion [25], wavelet transform, tunable-Q wavelet transform [26], iterative eigenvalue decomposition of the Hankel matrix, variational mode decomposition, and Fourier decomposition method, demonstrate substantial progress in achieving higher performance and adaptability in time–frequency analysis [27]. Some studies utilizing EEG signals in different tasks are discussed in Table 1.

Most of the emotion classification methods have been evaluated on publicly available datasets, each varying in recording frequency, elicitation methods, and the types of emotions captured. The DEAP dataset [11], for instance, records at a sampling frequency of 512 Hz and includes 32 channels. It utilizes music videos as stimuli to elicit emotional responses, targeting emotional states such as valence, arousal, liking, and dominance across 32 subjects. The SEED dataset [12], recorded at a sampling frequency of 1000 Hz with 62 channels, focuses on emotions elicited by movie clips, includes 15 subjects, and categorizes emotions into positive, neutral, and negative. SEED-IV [13] also uses movie clips at a sampling frequency of 1000 Hz with the same 62 channels, including happy, fear, neutral, and sad. None of these datasets were collected in virtual reality (VR) environments. The challenges with current datasets lie in accurately capturing intense emotions, especially when emotional responses are subjective, for instance, one person may experience fear while watching a horror video clip, whereas another may respond neutrally. This variability in subject responses can lead to false annotations, resulting in poor detection performance. To address these challenges and capture intense emotions effectively, we used an HP virtual reality headset to evoke strong emotional responses from participants within a controlled virtual environment. This immersive setup helps isolate individuals from external distractions, allowing for a more engaged experience with the presented scenarios. The controlled virtual setting facilitated more genuine and natural emotional responses than non-virtual ones, for example, when participants watch a joyful movie in the virtual environment, they immerse themselves entirely in the situation, genuinely laughing and forming a deep emotional connection with the present situation. Table 2 shows that public datasets vary in sampling rates, number of channels, and EEG recording devices.

The research contributions (RC) of this paper : are summarized as follows:

- ***RC01:*** This study introduces a new EEG emotion dataset (VREEG) collected in a virtual environment using a VR headset and an EEG cap to elicit intense emotional responses under four emotion classes: happy, sad, neutral, and fear, enhancing signal quality.
- ***RC02:*** Conducted a comprehensive evaluation of emotion classification performance using our dataset, giving detailed insights into its effectiveness.
- ***RC03:*** The study conducts a thorough experimental analysis; it includes a comparative analysis to evaluate the effectiveness of the dataset collected in virtual environments for classifying negative, neutral, and positive emotions in virtual and non-virtual settings.

The rest of this paper is structured as follows: Section 1 provides a workflow summary, including data collection and participants' self-assessments. Section 2 details the preprocessing steps, feature extraction, and classification models. Section 3 discusses the evaluation metrics, performance analysis, comparative analysis between virtual and non-virtual environments, and a comparison with existing datasets. Finally, Section 4 presents the conclusion.

## 1. Workflow summary

The paper workflow summary is depicted in Fig. 1, which includes key steps, *Step 1: Input Test Protocol*, *Step 2: EEG Recording*, *Step 3: Preprocessing & Feature Extraction*, *Step 4: Model Training*, and *Step 5: Performance Evaluation*. In the input test protocol stage, subjects are presented with audio and visual stimuli to elicit emotional responses, which are then recorded. The EEG recording involves devices such as the Emotiv EPOC Flex and HP VR headset, followed by preprocessing and feature extraction steps, and models are trained for classification. Finally, the performance evaluation stage utilizes metrics such as accuracy, precision, recall, and F1 score to assess the model's effectiveness.

### 1.1. Setup for experiment and data collection

This section discussed experimental setup and data collection protocols. Before recording, participants (subjects) were informed about the safety measures related to the equipment and asked to provide written consent to participate in the experiment. Importantly, all individuals involved in this study had no previous mental health issues or emotional problems. The experimental summary is given in Table 3.

#### 1.1.1. Experimental devices

The experimental setup includes an Emotiv Epoc Flex cap with 32 EEG electrodes, an HP virtual reality (VR) headset, and two computational display units. The Emotiv Epoc Flex cap is a wireless device designed for brain signal recording, with 32 electrodes arranged according to the 10-20 EEG montage, covering locations such as Fp1, F3, F7, FC1, FC5, C3, Cz, FC2, C4, FC6, T7, CP5, CP6, CP1, CP2, P3, P7, Pz, PO9, PO10, O1, Oz, O2, P4, P8, T8, FT9, FT10, F4, F8,

**Table 1**
Studies utilizing EEG signals: Highlighting methods, applications, and key findings across diverse domains.

| Citation | Application | Methods/Techniques | Findings |
|---|---|---|---|
| [28] | Detection of Major Depressive Disorder using EEG signals | Graph convolutional transformer network | Leveraged spatial dependencies between brain regions to improve model performance. |
| [29] | Emotion recognition | Channel relationships-based graph network | High accuracy in subject-dependent and subject-independent experiments. |
| [30] | Sleep apnea detection | Sparse spectrum-based swarm decomposition | Proven effective for robust nonstationary signal processing. |
| [31] | Emotion recognition | Hybrid convolutional and recurrent neural networks | Captured spatial and temporal features for improved emotion classification. |
| [32] | Upper limb movement recognition | Clustering sparse swarm decomposition | Demonstrated potential in motor function analysis. |
| [33] | Cognitive visual object classification | Sparse swarm decomposition with multichannel eeg-meg signals | Highlighted versatility across modalities. |
| [34] | Emotion recognition | Graph neural networks | Used RGNet, a model designed to learn regional-level EEG representations. |
| [35] | Atrial fibrillation detection | Variational mode decomposition | Framework applicable for EEG-based emotion recognition. |
| [36] | Imagined speech detection | Multivariate swarm sparse decomposition with joint time–frequency analysis | Providing an intuitive communication tool for physically impaired patients. |
| [37] | Emotions recognition | Tunable Q-factor wavelet transform and rotation forest ensemble classifier | Used to extract meaningful information from complex EEG signals. |
| [38] | Human emotion recognition | Tunable Q-factor wavelet transform with eeg sub-bands | Effective for extracting meaningful information from complex time–frequency characteristics. |
| [39] | Motor imagery-based EEG-BCI | Enhanced sparse swarm decomposition | Improved classification accuracy and robustness in complex BCI scenarios. |
| [40] | Human emotion recognition using EEG signals | Framework using EEG bands extracted with wavelet filter banks | Extracts wavelet energy and Shannon entropy features, fuses them, and classifies emotions. |
| [41] | EEG-based automatic emotion recognition | Proposed the Connectivity Uncertainty GCN | Effectively capture long-path dependencies, mitigate over-smoothing, and address noisy labels. |
| [42] | EEG-based emotion recognition | Attention-Based Temporal Graph Representation Network | Integrates attention mechanism, graph convolution to capture spatial, frequency, and temporal features. |
| [43] | Epilepsy seizure control | Real-time wireless wearable EEG devices | Enabled efficient real-time monitoring and intervention. |

**Table 2**
This table presents a summary of public EEG datasets commonly used in emotion recognition alongside the newly created dataset. It outlines key attributes such as participant numbers, recording channels, sampling frequency, emotion labels, and stimuli types.

| Datasets (frequency) | Emotions elicitation | Number of channels | Total subjects | Emotions types | VR |
|---|---|---|---|---|---|
| DEAP (512 Hz) [11] | Music videos | 32 | 32 | Valence, arousal, liking, and dominance | No |
| SEED (1000 Hz) [12] | Movie clips | 62 | 15 | Positive, neutral, and negative | No |
| SEED-IV (1000 Hz) [13] | Movie clips | 62 | 15 | Happy, fear, neutral, and sad | No |
| NVEEG (**Ours**) (128 Hz) | Movie clips | 32 | 15 | Positive, neutral, and negative | No |
| VREEG03 (**Ours**) (128 Hz) | Movie clips | 32 | 70 | Positive, neutral, and negative | Yes |
| VREEG04 (**Ours**) (128 Hz) | Movie clips | 32 | 70 | Happy, fear, neutral, and sad | Yes |

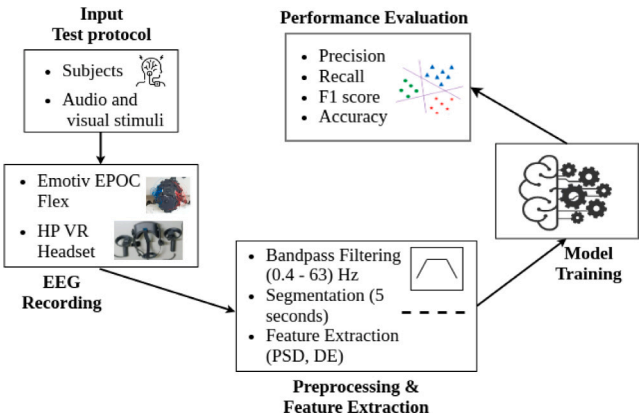
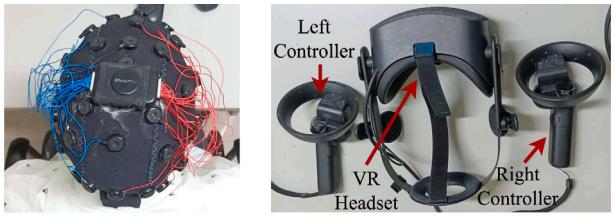



**Fig. 1.** Workflow summary illustrating EEG signal recording using Emotiv EPOC Flex, preprocessing and feature extraction, model training, and performance evaluation with accuracy, precision, recall, and F1-score.

**Table 3**
Overview of experimental details.

| | |
|---|---|
| **Audio and visual stimuli** | |
| Number of videos | 100 |
| Content types: | Audio-video |
| Duration of videos: | 5–10 min |
| **Details about the experiment** | |
| Subjects (Participants) | 70 (50 male, 20 female) |
| Age range of male Subjects | 19–36 years (median = 26.5 years) |
| Age range of female subjects | 18–30 years (median = 27 years) |
| EEG recording device | Emotiv Epoc Flex (32 channels) |
| Virtual environment device | HP virtual reality (VR) headset |
| Recorded signals | 32 channels, sampled at 128 Hz |
| EEG signal duration | Approximately 5 min per emotion class |




**(a): Emotiv cap** **(b): Virtual Reality (VR) Headset**

**Fig. 2.** Devices used in the experimental setup include the Emotiv Epoc Flex cap with 32 channels and an HP virtual reality headset with left and right controllers.

Fz, and Fp2. This HP virtual reality headset is equipped with left and right controllers for interaction within the virtual environment, Fig. 2 shows the EEG cap and headset used during data collection. The VR headset operated on a computational display unit with a 64-bit Windows 10 operating system, 32 GB of RAM, and an Intel Xeon(R) W-1350 processor running at 3.30–3.70 GHz. A second computational unit with a 64-bit Windows 11 operating system, 8 GB RAM, and an Intel(R) Core(TM) i5 processor clocked at 1.60–2.11 GHz is used for EEG signal collection and visualization.

### 1.1.2. Stimulus

We selected audio and video clips as stimuli to evoke emotions in our experimental setup. More than 100 videos were gathered under the supervision of expert neurologists for different emotions, i.e., happy, sad, neutral, and fearful (or positive, neutral, and negative) from YouTube in the Hindi language due to participants comfortability in their native language. Selected videos, including stand-up comedy and funny movie scenes, aim to evoke feelings of happiness (positive).

**Fig. 3.** Experimental setup for EEG signal recording during emotion elicitation. The subject is seated while wearing an EEG recording cap to capture brain activity, and a virtual reality headset presents a video clip that evokes emotions.

Similarly, videos featuring emotional movie scenes or heartfelt scenes can elicit a sense of sadness (negative). For a neutral or relaxed experience, instrumental music like Jazz, Piano, and Guitar have been considered, and to evoke fear emotion, we have selected horror video clips (negative). To annotate the recorded files and assign a proper label for them, we consulted with the doctor. Video and audio stimuli tend to evoke stronger emotional responses in participants compared to static images, as subjects experience more intense emotions while watching these clips. Therefore, these visual and auditory stimuli were presented to participants to elicit specific emotional responses in the brain.

### 1.1.3. EEG signals recording in virtual environment

In this virtual environment, data was collected from a mixed group of 70 participants, including 50 males and 20 females. Participants experienced emotional stimuli shown through videos. Concurrently, EEG signals were recorded in response to these stimuli using an EEG recording cap. Based on the number of emotion classes, the recorded datasets from this virtual environment are categorized into two sets, i.e., VREEG04 and VREEG03. Here, VREEG04 includes 4 classes: happy, sad, neutral, and fear and VREEG03 consists of 3 classes: negative, neutral, and positive, where 'happy' emotion classes are kept in the positive subset, and fear and sad classes are merged into a negative subset. The complete emotions recording setup in the virtual environment is shown in Fig. 3, and the recorded EEG signals for the 32 channels are illustrated in Fig. 4.

### 1.1.4. Recording time

The recording time refers to the total duration for data collection from one subject; during the recording of these emotional responses, the relaxation time is given to the subject, which is around 3–5 min before recording another emotional response. Hence, the total recording time required for collecting signals from a single subject, including the relaxation period following each emotion recording for four classes, typically takes 40–50 min, as Fig. 5 illustrates the required time for recording one subject.

Let $t_{\text{setup}}$ represent the setup time, $t_{\text{relax}}$ the time taken during the relaxation period, $t_{\text{happy}}$ the time for happy emotion, $t_{\text{fear}}$ for fear emotion, $t_{\text{neutral}}$ for neutral emotion, and $t_{\text{sad}}$ for sad emotion. The total recording time for one subject is denoted by $t_{\text{total}}$.

$$t_{total} = t_{setup} + t_{happy} + t_{relax} + t_{fear} + t_{relax} + t_{neutral} + t_{relax} + t_{sad}$$
$$t_{total} = t_{setup} + 3 \times t_{relax} + t_{happy} + t_{fear} + t_{neutral} + t_{sad} \quad (1)$$

### 1.2. Participant's self-assessment

At the end of each recording, participants were asked for their arousal and valence scores on a scale of 1 to 9. To ensure the validity of the self-assessment, we performed these evaluations in the presence of a neurologist confirmation, which verified that the recorded signals aligned with the predefined emotional labels of the stimuli. When discrepancies were observed, these were addressed case-by-case, and the final label was decided based on the neurologist's confirmation.

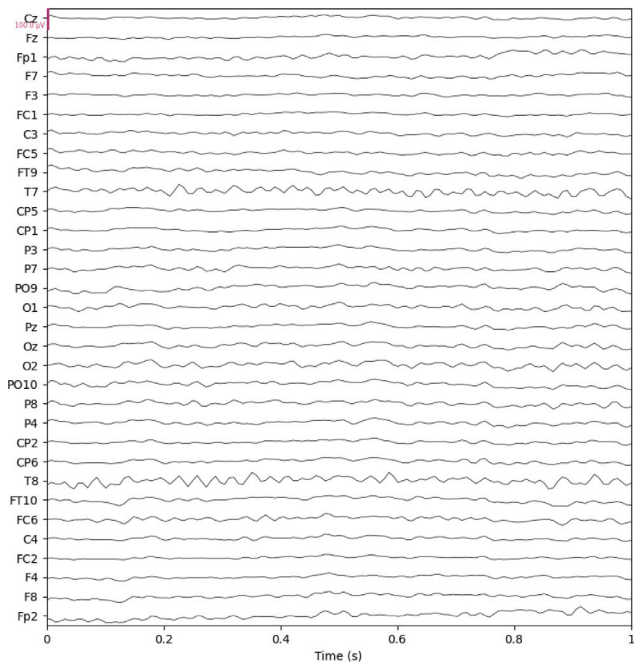



**Fig. 4.** This figure illustrates the recorded EEG signals across 32 channels over a 1-second interval where the $x$-axis represents time (in seconds), while the $y$-axis corresponds to the individual EEG channel.

#### 1.2.1. Valence

Valence refers to a situation's intrinsic attractiveness or aversiveness, measuring how pleasant (positive valence) or unpleasant (negative valence) an emotion is. Positive valence includes emotions like happiness and joy, while negative valence includes sadness, anger, and fear, and it scaled from 1 (very unpleasant) to 9 (very pleasant) [11].

#### 1.2.2. Arousal

Arousal refers to an emotion's intensity or energy level, measuring how activated or calm a state is. High arousal includes intense emotions like excitement and fear, while low arousal includes calm or subdued emotions like relaxation and boredom, and it scaled from 1 (very calm) to 9 (very excited or activated) [11].

#### 1.2.3. Valence-Arousal (VA) space for emotion categorization

The combination of valence and arousal creates a two-dimensional space [11], often referred to as the valence-arousal space, which is used to categorize and visualize emotions. Different emotional states are positioned within this space as follows: In this space, happiness is classified as a high-valence, high-arousal (HVHA) emotion, as it is both pleasant and energizing. Fear is a low-valence, high-arousal (LVHA) emotion, reflecting its unpleasant but highly stimulating nature. Neutral falls in the mid-range for valence and arousal, representing a balanced emotional state. Finally, sad is a low-valence, low-arousal (LVLA) emotion, as it is both unpleasant and calming or de-energizing, as shown in Fig. 6. The valence and arousal scores obtained from the subjects during recording with mean and standard deviation across all subjects discussed in Table 4 and Fig. 7 show the box plots of these scores.

## 2. Methodology

The methodology for EEG-based emotion classification starts with data preprocessing using a bandpass filter to ensure signal quality. After filtering, the signals are segmented into non-overlapping windows, as described in Algorithm 1. Following preprocessing, features are extracted from the segmented signals, which are then used to train the classifier model. To evaluate the performance of the classifier, metrics such as precision, recall, F1 score, and accuracy are employed, providing a comprehensive analysis of the model's effectiveness in recognizing emotional states. Finally, a detailed performance analysis is conducted to assess performance.

**Table 4**
This table represents the mean (Mean) and standard deviation (Std) of valence and arousal scores with ratings ranging from 1 to 9 for happy, sad, neutral, and fear emotions.

| Emotion | Valence (Mean) | Valence (Std) | Arousal (Mean) | Arousal (Std) |
|---|---|---|---|---|
| Happy | 7.13 | 1.12 | 7.20 | 1.08 |
| Sad | 2.09 | 1.05 | 2.10 | 1.04 |
| Neutral | 4.20 | 0.66 | 4.40 | 0.70 |
| Fear | 3.07 | 1.07 | 6.08 | 1.28 |

**Algorithm 1** EEG Recording and Emotion Recognition

1: **Input and Test Protocol**
Subjects (S) = $\{S_1, S_2, \ldots, S_{70}\}$, (50 Male, 20 Female)
Stimuli = $\{V_1, V_2, \ldots, V_{100}\}$ (Audio and visual)
Devices = Emotiv Epoc Flex Cap, HP VR Headset
2: **Output**
$\mathbf{Y}_{\text{pred}} = \{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_{(3 \text{ or } 4)}\}$ (Predicted Emotion Classes)
3: **EEG Recording**
$\mathbf{X}(t) = [x_c(t)]_{c=1}^{C},\ t \in \{1, 2, \ldots, T\},$ and $C \in \{1, 2, \ldots, 32\}$
$\mathbf{X}(t) \in \mathbb{R}^{C \times T},\quad T = f_s \cdot D$ (EEG dimensions)
$C$ (Number of channels), $f_s = 128\,\text{Hz}$ (Sampling frequency)
$D = 5$ minutes (Recording duration)
4: **Preprocessing and Feature Extraction**
$\mathrm{X}_{\text{filtered}}(t) = \mathrm{X}(t) * H(t)$
$H(t) = \text{Bandpass}(f_{\text{low}}, f_{\text{high}}),\ f_{\text{low}} = 0.4\,\text{Hz},\ f_{\text{high}} = 63\,\text{Hz}$
$\mathrm{X}_{\text{segmented}} = \{\mathrm{X}_{\text{filtered}}[t_0 : t_0 + T] \mid t_0 = n \cdot T,\ n \in \mathbb{Z}^+,\ T = 5\,\text{seconds}\}$
**for** $k \leftarrow 1$ to $N$ (total segments)
$\text{PSD}[k] = \frac{1}{T} \int_0^T \left| \mathcal{F}(\mathrm{X}_{\text{segmented},k}(t)) \right|^2 dt$
$\sigma_k^2 = \frac{1}{T} \int_0^T \left( \mathrm{X}_{\text{segmented},k}(t) - \mu_k \right)^2 dt$
$\text{DE}[k] = \frac{1}{2} \log 2\pi e \sigma_k^2$
**end for**
5: **Model Training and Prediction**
$\mathbf{X}_{\text{features}} = \{\text{PSD}, \text{DE}\}$ (Feature set)
$C(\mathbf{X}; \theta) = \arg\max_k\ \mathcal{P}(y = k \mid \mathbf{X}; \theta)$
$\mathcal{L}_{\text{train}} = \frac{1}{N} \sum_{i=1}^{N} \ell(C(\mathbf{X}_i; \theta), y_i)$
$\theta^* = \arg\min_\theta\ \mathcal{L}_{\text{train}}$
$\mathbf{Y}_{\text{pred}} = C(\mathbf{X}_{\text{test}}; \theta^*)$
6: **Performance Evaluation**
$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+TN+FN}}$
$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}},\quad \text{Recall} = \frac{\text{TP}}{\text{TP+FN}}$
$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision+Recall}}$

### 2.1. Preprocessing

Based on the subject's responses, only the experiment segments where the target emotions were elicited were selected for further analysis. The EEG signals were visually inspected and accurately annotated. A bandpass filter was applied to isolate the frequencies in the range of 0.4 to 63 Hz and to remove the noise and artifacts. After preprocessing, we extracted 5-second non-overlapping segments from each EEG signal, and features were subsequently computed for each segment.

### 2.2. Feature extraction

Two types of features are extracted in this analysis on the preprocessed data segments: differential entropy (DE) [44], which performs

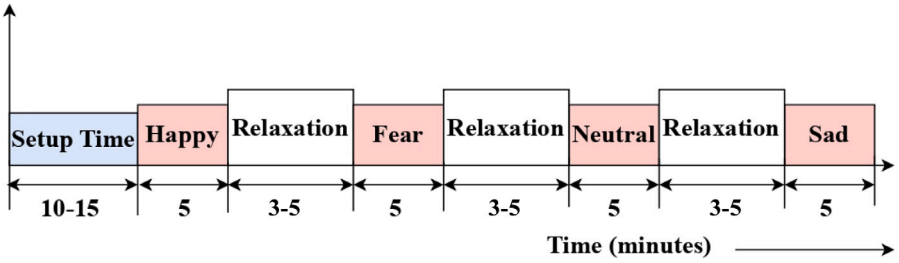



**Fig. 5.** This figure shows a timeline of EEG signal recording for an individual subject.

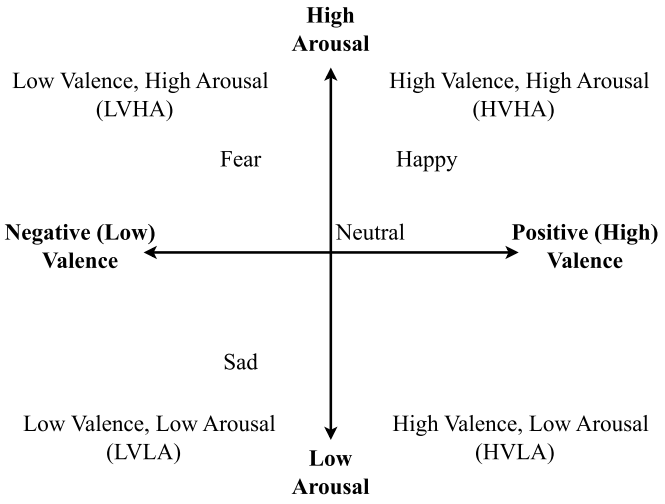



**Fig. 6.** Valence-Arousal (VA) model of emotions.

better compared to power spectral density (PSD) [45] using a short-term fourier transform with a 1-second [12] and 5-second [13] window length without overlapping. The DE feature is defined with the original formula for calculating the feature $DE(X)$ for a segment $X$:

$$DE(X) = -\int_x p(x)\log(p(x))\,dx \tag{2}$$

where $p(x)$ represents the probability density, if the EEG signals follow a Gaussian distribution $N(\mu, \sigma^2)$, the DE feature $DE(X)$ can be simplified and calculated using the following equation:

$$DE(X) = -\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\frac{(x-\mu)^2}{2\sigma^2} \log\frac{1}{\sqrt{2\pi\sigma^2}} \exp\frac{(x-\mu)^2}{2\sigma^2} dx$$

$$DE(X) = \frac{1}{2}\log 2\pi e\sigma^2$$

Differential entropy is similar to the logarithmic power spectral density when calculated for a fixed-length EEG sequence within a specific frequency band. Unlike PSD, DE features provide a more balanced ability to differentiate EEG patterns based on high-frequency energy. We extracted both features across five frequency bands: (1) delta ($\delta$): 0.4–4 Hz; (2) theta ($\theta$): 4–8 Hz; (3) alpha ($\alpha$): 8–12 Hz; (4) beta ($\beta$): 12–25 Hz; and (5) gamma ($\gamma$): 25–45 Hz [12,13,46] for each channel.

### 2.3. Classification with Support Vector Machine (SVM)

Support Vector Machine was chosen for its effectiveness in handling high-dimensional data, as well as its proven performance in EEG-based emotion classification tasks. It has consistently delivered reliable results in various studies [12,13]. This analysis used it with default settings for multi-class classification, utilizing a Radial Basis Function (RBF) kernel, a regularization parameter $C$ of 1.0, and the 'one-vs-one' strategy to handle multiple classes [47–49]. Given a training dataset $\{(x_i, y_i)\}_{i=1}^{N}$ where $x_i \in \mathbb{R}^d$ represents the feature vectors and $y_i \in \{1, 2, \ldots, K\}$ denotes the class labels for $K$ classes (3 or 4), the SVM aims to find a decision boundary that maximizes the margin between classes. For a binary classification problem, the decision function is defined as:

$$f(\mathbf{x}) = \mathbf{w}^\top \mathbf{x} + b \tag{3}$$

where $\mathbf{w} \in \mathbb{R}^d$ is the weight vector and $b \in \mathbb{R}$ is the bias term. The optimization problem to find $\mathbf{w}$ and $b$ can be formulated as:

$$\min_{\mathbf{w},b,\xi} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}\xi_i$$

$$y_i(\mathbf{w}^\top \mathbf{x}_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \quad i = 1, \ldots, N$$

Here, $\xi_i$ are the slack variables that allow for some misclassifications, and $C$ is used to control the trade-off between maximizing the margin and minimizing the classification error. For multi-class classification using the "one-vs-one" approach, $K$ binary classifiers are trained, and each classifier distinguishes between a single class $k$ and the rest of the classes. The decision function for the classifier distinguishing class $k$ from all other classes is given by:

$$f_k(\mathbf{x}) = \mathbf{w}_k^\top \mathbf{x} + b_k$$

During prediction, the decision function outputs are combined, and the class with the highest score is assigned as the predicted label:

$$\hat{y} = \arg\max_k f_k(\mathbf{x})$$

### 2.4. Classification with Lightweight EEGNet (LEEGNet)

In this work, EEGNet was employed due to its efficiency and effectiveness in handling EEG-based emotion recognition tasks [50–53]. The EEGNet architecture is specifically designed to extract both temporal and spatial features from EEG signals while maintaining computational efficiency. The model begins with an initial convolutional block that captures temporal patterns across EEG data, followed by batch normalization to standardize activations and improve training stability. A depthwise convolution layer is then applied to extract spatial features across EEG channels, accompanied by activation functions, pooling, and dropout for regularization. The subsequent block refines the spatial–temporal features further through separable convolution, followed again by normalization, activation, pooling, and dropout to ensure robust feature extraction. The final features are flattened and passed through a fully connected dense layer with regularization constraints to maintain stability, followed by a softmax activation function for class probability estimation, as shown in Fig. 8. The well-structured design of EEGNet provides a balance between high classification performance and low computational overhead, making it suitable for real-time EEG emotion recognition applications. EEGNet efficiently extracts spatial and temporal information through depthwise and separable convolutions. It employs a larger pooling size ((1, 4) in the first block and (1, 8) in the second block), which helps in reducing the temporal dimensions but may lead to some loss of detailed temporal

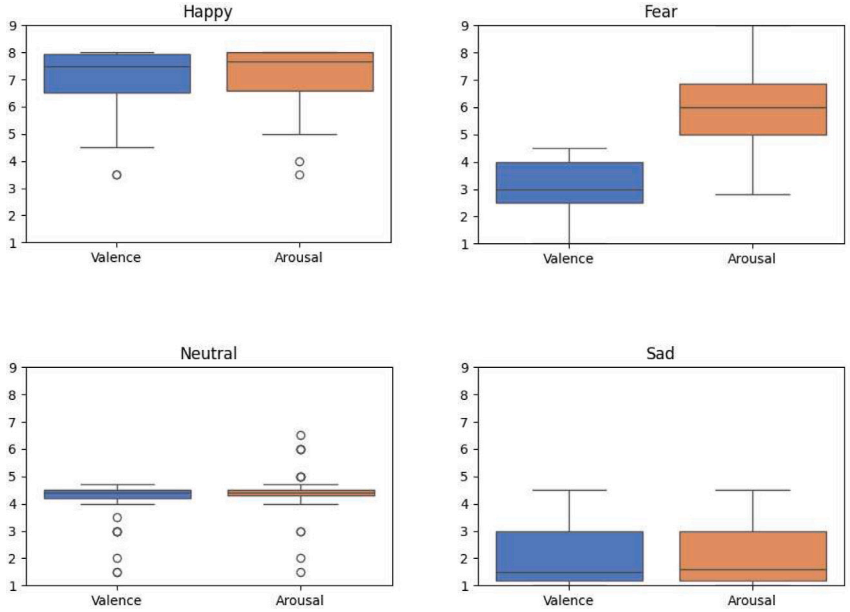



**Fig. 7.** This figure represents box plots of valence and arousal scores for four distinct emotions—happy, fear, neutral, and sad obtained from participants during EEG signal recording. Each box plot illustrates the distribution of valence and arousal ratings.

**Table 5**
LEEGNet model architecture and parameters.

| Layer | Type | Filters/Units | Kernel size | Additional parameters |
|---|---|---|---|---|
| Input | Input Layer | – | – | Channels x features |
| Conv2D | 2D Convolution | 4 | (1, 32) | Stride = 1 |
| | Batch Normalization | – | – | – |
| DepthwiseConv2D | Depthwise Convolution | Multiplier = 2 | (Channels, 1) | Activation = ELU |
| | Batch Normalization | – | – | – |
| | Average Pooling | – | (1, 2) | Stride = 2 |
| | Dropout | – | – | Rate = 0.5 |
| SeparableConv2D | Separable Convolution | 16 | (1, 16) | Activation = ELU |
| | Batch Normalization | – | – | – |
| | Average Pooling | – | (1, 2) | Stride = 2 |
| | Dropout | – | – | Rate = 0.5 |
| Flatten | Flatten | – | – | – |
| Dense | Fully Connected | Classes | – | Activation = Softmax |

information. LEEGNet is a modified version of EEGNet that is designed to enhance classification performance while maintaining computational efficiency with smaller pooling sizes ((1, 2) in both blocks); it retains more temporal information during down-sampling, enabling it to capture the fine-grained features essential for emotion recognition, as discussed in Table 5. Experiments were conducted on a server equipped with NVIDIA-SMI 440.44, Driver Version 440.44, and CUDA Version 10.2, ensuring consistent and reproducible results.

## 3. Evaluation metrics and performance analysis

### 3.1. Evaluation metrics

The classification performance of the classifiers was evaluated using four metrics: Accuracy, Precision, Recall, and F1-Score [11,54]. Accuracy measures the overall correctness of the model across all classes, providing a general sense of how well it distinguishes between them. Precision reflects the ability of the model to avoid false positives by calculating the proportion of true positive predictions out of all positive predictions made. Recall assesses the model sensitivity by determining how well it captures all relevant instances (true positives) in each class. Finally, the F1-Score balances Precision and Recall, offering a single metric that is particularly useful when dealing with imbalanced classes, as it provides a harmonic mean of the two. These metrics collectively offer a comprehensive view of the model performance across various aspects of classification.

### 3.2. Performance analysis

This analysis involved 70 subjects, including 50 males and 20 females, categorized into three and four emotion classes: (negative, neutral, positive) and (happy, sad, neutral, and fear). Raw EEG signals, further power spectral density (PSD), and differential entropy (DE) features extracted were all separately used for performance evaluation. In performance analysis, 60 subjects are used for model training, and 10 are used for validation and testing with multiple sets of 60 and 10, and the final average numbers are reported in this work across these sets. The results are reported in Table 6, which represent the testing scores for 3-classes and 4-classes in different cases. The support vector machine shows lower classification scores than LEEGNet, as shown in the table.

The Figs. 9 and 10 represent t-SNE plots both visually represent the feature separability across different emotion classes; these visualizations offer insights into how each feature set captures distinct emotional

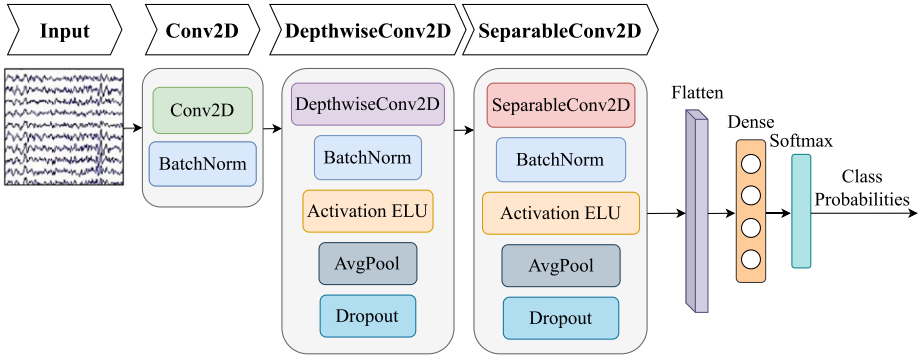



**Fig. 8.** This figure shows an EEG network that begins with a convolution layer to extract features from the EEG signals. Following this, a depthwise convolution layer is applied to each feature map, enabling the model to learn spatial information. Then, it passes to a separable convolution layer, where the depthwise convolution captures temporal information for each feature map.

**Table 6**
The table represents the emotions classification performance for (positive, neutral, and negative), and (fear, happy, neutral, and sad) emotions.

| VREEG03 (Positive, negative, and neutral emotions) | | | | |
|---|---|---|---|---|
| Method | Precision | Recall | F1-Score | Accuracy |
| SVM+PSD | 74.37 | 72.97 | 73.20 | 72.97 |
| LEEGNet+raw EEG | 76.04 | 73.56 | 73.35 | 73.56 |
| LEEGNet+PSD | 78.51 | 75.44 | 75.13 | 75.44 |
| LEEGNet+DE | 79.71 | 77.75 | 77.20 | **77.75** |
| VREEG04 (Happy, sad, fear, and neutral) | | | | |
| Method | Precision | Recall | F1-Score | Accuracy |
| SVM+PSD | 57.71 | 56.31 | 56.82 | 56.31 |
| LEEGNet+raw EEG | 61.18 | 55.94 | 56.01 | 55.94 |
| LEEGNet+PSD | 60.77 | 59.10 | 57.79 | 59.10 |
| LEEGNet+DE | 66.12 | 62.67 | 60.62 | **62.67** |

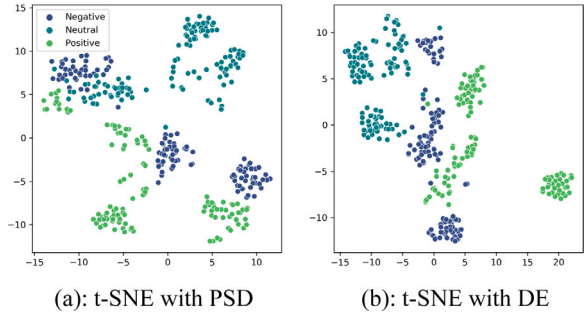



**Fig. 9.** These figures represent t-SNE plots for positive, neutral, and negative emotions with power spectral density and differential entropy features.

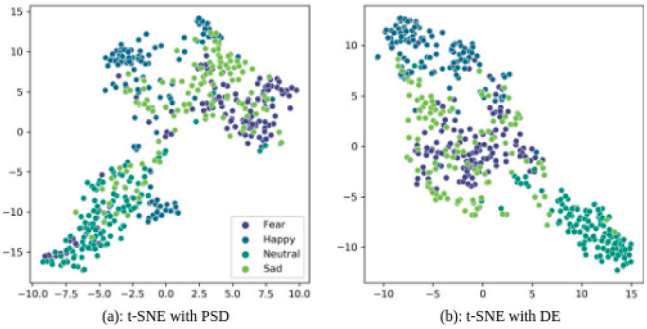



**Fig. 10.** These figures show the t-SNE plots for fear, happy, neutral, and sad emotions.

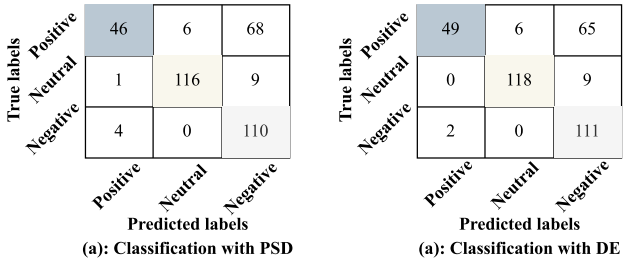



**Fig. 11.** These figures represent the confusion matrices for positive, neutral, and negative emotions. The entries along the diagonal represent accurate predictions, while the off-diagonal entries indicate misclassifications.

states and organizes them in a low-dimensional space. Fig. 9(a) shows clusters of the three emotion classes (negative, neutral, and positive) based on PSD features, and Fig. 9(b), shows a different distribution of the clusters; some separation is visible with DE features. Fig. 10(a), with PSD features, shows that the classes have some degree of separability, but there is a noticeable overlap among the clusters. Fig. 10(b), when using DE features, demonstrates more defined clusters for each emotion class (fear, happy, neutral, and sad). The clusters are separable, indicating that DE may be capable of capturing small variations in EEG signals. These visualizations show that DE features provide better class separability in both three and four classes compared to PSD.

### 3.3. Class-wise performance evaluation

To assess the performance of the classifier model, confusion matrices are used to represent the accuracy of predictions across different emotion classes. Figs. 11 and 12 show these matrices for three-class and four-class classifications, respectively, offering a detailed insight into model performance. These matrices show the model ability to correctly classify each emotion class, highlighting correct predictions and areas where misclassifications occurred. Until this analysis, it was observed that LEEGNet outperformed the support vector machine, that is why the rest of the analysis focused on LEEGNet in this study.

### 3.4. Subject-wise performance analysis

Subject-wise performance analysis was done to evaluate the robustness and generalizability of the classification model. The performance of the classification model was checked with individual subjects, and this analysis shows variations in performance that may be attributed to inter-subject variability in EEG patterns. The obtained results are visualized with bar charts 13 and 14, both compare accuracy, precision, recall, and f1 scores across these subjects.

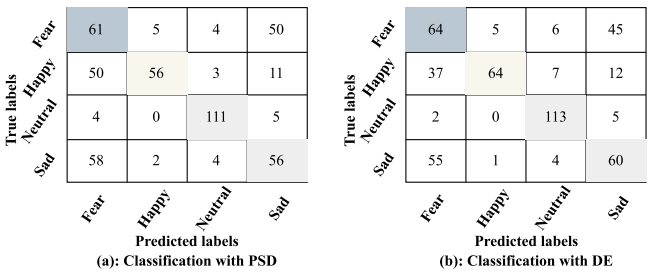



**Fig. 12.** Figure shows the confusion matrices for happy, sad, fear, and neutral emotions.

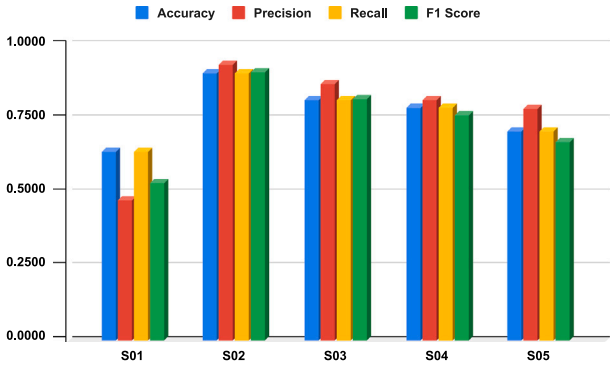



**Fig. 13.** The chart shows the individual performance, showcasing the variability across different subjects (S01, S02, S03, S04, and S05) and highlighting the model's effectiveness in capturing subject-specific differences for positive, neutral, and negative emotions.

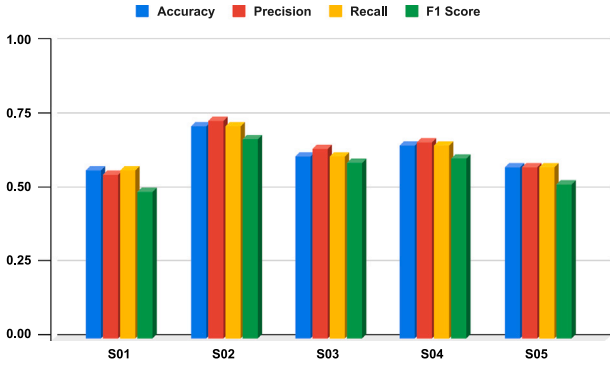



**Fig. 14.** This chart plot shows the subject-wise classification performance for happy, sad, neutral, and fear emotions.

### 3.5. Gender-based analysis

The gender-based analysis involves 20 subjects, in each case male and female, experiencing three emotional states: positive, negative, and neutral. Here, the model performance is checked across gender groups to understand how classification performance may vary among male and female subjects. Both power spectral density and differential entropy are employed, allowing the classifier to capture distinct aspects of EEG signal variation across genders. The results obtained from this analysis are discussed in Table 7.

### 3.6. Channel and band-wise analysis

In previous discussions, the performance analysis was done across all frequency bands and channels. The next step is to identify brain areas associated with different emotions and to select the electrode in these areas to minimize computational complexity and reduce irrelevant noise because some channels may not significantly contribute to emotion recognition retaining these extra electrodes can increase computational complexity, introduce noise, and degrade performance. The recorded signals with 32 channels out of these, some channels corresponding to brain lobes, selection followed the 10–20 international system and focused on the frontal, temporal, parietal, occipital, and central lobes. For example, channels Fp1, Fp2, F3, F4, F7, F8, Fz, FC1, FC2, FC5, and FC6 are associated with the frontal lobe; T7 and T8 with the temporal lobe; P3, P4, P7, P8, and Pz with the parietal lobe; O1, O2, and Oz with the occipital lobe; and C3, C4, and Cz with the central lobe, as shown in Fig. 15, each set of channels is associated with some neural activities and cognitive functions. The performance analyzed with these subsets of electrodes across frequency bands: delta ($\delta$) 0.5–4 Hz, theta ($\theta$) 4–8 Hz, alpha ($\alpha$) 8–12 Hz, beta ($\beta$) 12–25 Hz, gamma ($\gamma$) 25-45 Hz, and total ($\delta$, $\theta$, $\alpha$, $\beta$, and $\gamma$) combines information from all these bands with VREEG03 dataset. This analysis aims to evaluate the importance of the selected channels within these frequency bands, and it found that considering a subset of 11 channels specifically, Fp1, Fp2, F3, F4, F7, F8, Fz, FC1, FC2, FC5, and FC6 produced results that almost similar as obtained when using all 32 channels more efficiently, as discussed in Table 8. Each frequency band (wave) is associated with some brain activities: delta waves are associated with deep sleep and unconscious states, theta waves with light sleep and relaxation, alpha waves to relaxation while awake, beta waves to active thinking and focus, and gamma waves to higher cognitive functions and information processing.

**Table 7**
The table shows results obtained from gender-based analysis for positive, neutral, and negative.

| Method | Precision | Recall | F1-Score | Accuracy |
|---|---|---|---|---|
| Female Subjects | | | | |
| EEGNet+PSD | 81.54 | 65.50 | 55.37 | 65.50 |
| EEGNet+DE | 67.29 | 67.66 | 60.03 | 67.66 |
| Male Subjects | | | | |
| EEGNet+PSD | 40.67 | 61.90 | 40.67 | 61.90 |
| EEGNet+DE | 67.32 | 63.83 | 58.56 | 63.83 |

### 3.7. Emotions analysis in virtual and non-virtual environments

To assess the impact of virtual environments compared to non-virtual environments, data were first collected from 15 subjects without VR headsets. These subjects were exposed to three emotional states: negative, neutral, and positive, as discussed in Table 9, classification accuracy and F1 scores in the non-virtual environment are relatively low, indicating limitations in emotion detection under non-immersive conditions. While in Table 10 shows comparative analysis demonstrates the advantages of virtual settings for data collection where classification accuracy and F1 scores are consistently higher in the virtual environment than in the non-virtual setting. The immersive nature of the virtual environment appears to enhance emotional engagement, leading to more genuine emotional responses and improved classification results.

### 3.8. Statistical analysis

To validate performance differences between virtual and non-virtual environments, a one-way analysis of variance (ANOVA) was conducted to compare classification accuracy and F1 scores [11–13]. This test for accuracy yielded a $p$-value of 0.004, indicating a statistically significant difference ($p \leq 0.01$). This result suggests that the virtual environment significantly improves classification accuracy [55–58]. Similarly, the F1 score yielded a $p$-value of 0.00045, confirming a significant difference ($p \leq 0.01$). These statistical findings emphasize the effectiveness of the virtual environment in emotion recognition.

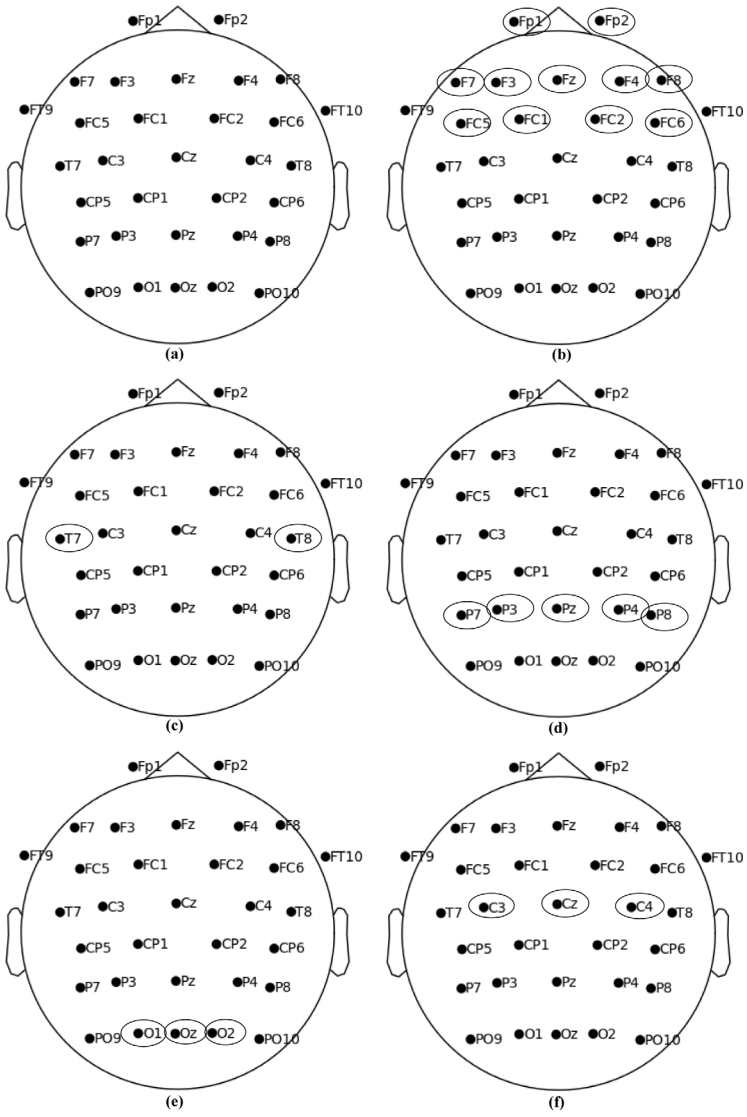



**Fig. 15.** In this figure, eclipse represents the location of selected channels using the 10-20 international system, where (a) represents the location of all 32 channels, (b) channels in the frontal lobe, (c) channels in the temporal lobe, (d) channels in the parietal lobe, (e) channels in the occipital lobe, and (f) channels in the central lobe.

### *3.9. Ablation study*

To comprehensively evaluate the proposed architecture, we conducted an ablation study to understand the contribution of individual components to the model's performance. Specifically, it investigated the effects of replacing the DepthwiseConv2D layer with Conv2D, removing the SeparableConv2D layer, replacing the ELU activation function with ReLU, and removing Batch Normalization layers. The results of these experiments are presented in Table 11 for each case.

### *3.10. Analysis with State-of-the-Art (SOTA) datasets*

Table 13 compares the classification performance between the SEED public dataset and the newly created dataset in virtual and non-virtual environments. This public dataset [12] comprises recordings from 15 subjects over three sessions. The classification accuracy and F1 scores combined in these sessions with PSD are 56.26%, and the F1 score is 51.63%, with DE 59.82% and the F1 score 58.22%, as discussed in Table 12.

For comparative analysis with public datasets, experiments were conducted in both non-virtual and virtual settings. The non-virtual setting on our dataset achieved an accuracy of 57.60% and an F1 score of 54.14%, similar to the public dataset. However, in the virtual setting, the VREEG dataset significantly outperformed both the SEED public dataset and our non-virtual setting, with an accuracy of 77.75%, an F1 score of 77.20% for three classes, and an accuracy of 62.67% and an F1 score of 60.62% for four classes. This analysis highlights the effectiveness of virtual environments in improving the performance and reliability of EEG-based emotion recognition systems. By providing a more immersive and controlled setting, virtual environments elicit stronger emotional responses, resulting in superior classification performance.

## 4. Conclusion

This paper introduces a new EEG emotion dataset collected in a virtual environment, along with a comprehensive analysis that considers various frequency bands and subsets of channels to improve computational efficiency. The findings show that using a subset of 11 channels achieved the same results as using all 32 channels. A comparative study was then conducted between virtual and non-virtual environments, using our dataset and a public dataset for validation. Since the public dataset was recorded in a non-virtual environment, we compared its results with those from our non-virtual dataset and

**Table 8**
The table shows the results of channel-wise and band-wise analysis for selected channels with various frequency bands corresponding to positive, neutral, and negative emotions.

| All 32 Channels | | | | | | | |
|---|---|---|---|---|---|---|---|
| Method | Metrics | $\delta$ | $\theta$ | $\alpha$ | $\beta$ | $\gamma$ | total |
| PSD | Accuracy | 38.78 | 63.53 | 61.83 | 63.74 | 69.19 | 75.44 |
| | F1-Score | 31.57 | 62.47 | 60.36 | 61.70 | 68.23 | 75.13 |
| DE | Accuracy | 55.94 | 62.67 | 59.14 | 64.36 | 56.61 | 77.75 |
| | F1-Score | 55.49 | 62.34 | 58.42 | 62.87 | 55.23 | 77.20 |
| FP1, FP2, F3, F4, F7, F8, Fz, FC1, FC2, FC5, FC6 | | | | | | | |
| Method | Metrics | $\delta$ | $\theta$ | $\alpha$ | $\beta$ | $\gamma$ | total |
| PSD | Accuracy | 35.06 | 59.06 | 62.69 | 64.58 | 61.86 | 74.17 |
| | F1-Score | 26.91 | 57.86 | 61.18 | 62.43 | 59.01 | 73.76 |
| DE | Accuracy | 52.47 | 63.03 | 57.64 | 57.94 | 57.03 | 75.28 |
| | F1-Score | 51.70 | 61.44 | 56.29 | 56.56 | 54.08 | 74.76 |
| T7, T8 | | | | | | | |
| Method | Metrics | $\delta$ | $\theta$ | $\alpha$ | $\beta$ | $\gamma$ | total |
| PSD | Accuracy | 45.64 | 44.75 | 40.06 | 50.67 | 57.72 | 66.58 |
| | F1-Score | 43.64 | 44.34 | 33.34 | 43.72 | 52.99 | 66.58 |
| DE | Accuracy | 53.03 | 50.03 | 45.78 | 54.00 | 60.39 | 69.14 |
| | F1-Score | 52.07 | 48.42 | 45.12 | 48.99 | 57.88 | 69.01 |
| P3, P4, P7, P8, Pz | | | | | | | |
| Method | Metrics | $\delta$ | $\theta$ | $\alpha$ | $\beta$ | $\gamma$ | total |
| PSD | Accuracy | 33.25 | 41.50 | 46.75 | 44.86 | 44.31 | 62.72 |
| | F1-Score | 22.26 | 31.66 | 39.98 | 39.07 | 39.34 | 62.14 |
| DE | Accuracy | 39.67 | 40.83 | 52.89 | 37.97 | 44.50 | 69.11 |
| | F1-Score | 38.54 | 36.84 | 50.78 | 34.05 | 41.97 | 69.23 |
| O1, O2, Oz | | | | | | | |
| Method | Metrics | $\delta$ | $\theta$ | $\alpha$ | $\beta$ | $\gamma$ | total |
| PSD | Accuracy | 33.81 | 33.33 | 37.17 | 32.61 | 34.36 | 48.44 |
| | F1-Score | 25.05 | 25.48 | 32.45 | 27.21 | 28.21 | 44.55 |
| DE | Accuracy | 34.47 | 35.69 | 40.17 | 34.03 | 34.89 | 54.22 |
| | F1-Score | 32.00 | 33.19 | 36.64 | 31.34 | 32.92 | 52.64 |
| C3, C4, Cz | | | | | | | |
| Method | Metrics | $\delta$ | $\theta$ | $\alpha$ | $\beta$ | $\gamma$ | total |
| PSD | Accuracy | 37.11 | 38.56 | 41.14 | 44.56 | 51.25 | 65.36 |
| | F1-Score | 29.87 | 34.94 | 34.84 | 38.54 | 51.14 | 64.47 |
| DE | Accuracy | 46.64 | 39.69 | 45.36 | 44.67 | 56.67 | 66.42 |
| | F1-Score | 46.11 | 38.45 | 43.48 | 42.58 | 55.56 | 65.81 |

**Table 9**
Table shows the analysis of the emotions in a non-virtual environment.

| Emotions (Positive, negative, and neutral emotions) | | | | |
|---|---|---|---|---|
| Method | Precision | Recall | F1-Score | Accuracy |
| PSD | 52.48 | 52.16 | 46.31 | 52.16 |
| DE | 54.07 | 57.60 | 54.14 | 57.60 |

**Table 10**
This table shows the results in virtual and non-virtual environments.

| Environments | Method | Classes | F1-Score | Accuracy |
|---|---|---|---|---|
| Without VR | PSD | Positive, neutral, | 46.31 | 52.16 |
| | DE | and negative | 54.14 | **57.60** |
| With VR | PSD | Positive, neutral, | 75.13 | 75.44 |
| | DE | and negative | 77.20 | **77.75** |

**Table 11**
Ablation study results for LEEGNet components.

| Component | Modification | Precision | Recall | F1-Score | Accuracy |
|---|---|---|---|---|---|
| LEEGNet | Original Architecture | 79.71 | 77.75 | 77.20 | 77.75 |
| DepthwiseConv2D | Replaced with Conv2D | 73.18 | 72.50 | 72.67 | 72.50 |
| SeparableConv2D | Removed | 74.49 | 70.41 | 70.82 | 70.41 |
| Activation Function | ELU → ReLU | 69.62 | 68.05 | 67.86 | 68.05 |
| Batch Normalization | Removed | 73.34 | 71.52 | 71.70 | 71.52 |

**Table 12**
Table shows classification accuracy and F1 scores obtained from the SEED dataset from different sessions and after combining all three sessions.

| Dataset | Method | F1-Score | Accuracy |
|---|---|---|---|
| SEED | PSD | 38.53 | 46.81 |
| session01 | DE | 44.58 | 49.28 |
| SEED | PSD | 45.81 | 50.72 |
| session02 | DE | 45.77 | 51.59 |
| SEED | PSD | 49.56 | 56.23 |
| session03 | DE | 58.76 | 61.74 |
| SEED | PSD | 51.63 | 56.26 |
| Combined | DE | 58.22 | 59.82 |

**Table 13**
Table shows classification accuracy and f1 scores obtained from the SEED and our datasets in non-virtual and virtual settings.

| Dataset | Method | Classes | F1-Score | Accuracy |
|---|---|---|---|---|
| SEED | PSD | Positive, neutral, | 51.63 | 56.26 |
| Combined | DE | and negative | 58.22 | **59.82** |
| Ours without | PSD | Positive, neutral, | 46.31 | 52.16 |
| VR (NVEEG) | DE | and negative | 54.14 | **57.60** |
| Ours with | PSD | Positive, neutral, | 75.13 | 75.44 |
| VR (VREEG03) | DE | and negative | 77.20 | **77.75** |
| Ours with | PSD | Happy, fear, | 57.79 | 59.10 |
| VR (VREEG04) | DE | neutral, and sad | 60.62 | **62.67** |

found similar performance. However, the dataset collected in the virtual environment yielded better results, highlighting the effectiveness and potential benefits of virtual environments for EEG-based emotion recognition.

This study presents a method for emotion recognition in virtual environments, with the potential for future enhancements. Incorporating advanced time–frequency techniques in future work could offer more detailed insights into signal characteristics associated with emotional states, while raw data with minimal preprocessing has some limitations, such as missing complex emotion-relevant information but this method ensures that the model remains computationally efficient and adaptable for real-time applications.

## CRediT authorship contribution statement

**Naseem Babu:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Udit Satija:** Writing – review & editing, Visualization, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Jimson Mathew:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Investigation, Formal analysis, Conceptualization. **A.P. Vinod:** Writing – review & editing, Validation, Investigation, Conceptualization.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Naseem Babu reports administrative support, equipment, drugs, or supplies, statistical analysis, and writing assistance were provided by Indian Institute of Technology Patna Department of Computer Science and Engineering. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## References


[1] S.M. Alarcão, M.J. Fonseca, Emotions recognition using EEG signals: A survey, IEEE Trans. Affect. Comput. 10 (3) (2019) 374–393, http://dx.doi.org/10.1109/TAFFC.2017.2714671.

[2] R. Jenke, A. Peer, M. Buss, Feature extraction and selection for emotion recognition from EEG, IEEE Trans. Affect. Comput. 5 (3) (2014) 327–339, http://dx.doi.org/10.1109/TAFFC.2014.2339834.

[3] Z. Zeng, M. Pantic, G.I. Roisman, T.S. Huang, A survey of affect recognition methods: Audio, visual, and spontaneous expressions, IEEE Trans. Pattern Anal. Mach. Intell. 31 (1) (2009) 39–58, http://dx.doi.org/10.1109/TPAMI.2008.52.

[4] L. Luis, J. Gomez-Gil, Brain computer interfaces, a review, Sensors (Basel, Switzerland) 12 (2012) 1211–1279, http://dx.doi.org/10.3390/s120201211.

[5] F. Lotte, M. Congedo, A. Lécuyer, L. Fabrice, B. Arnaldi, A review of classification algorithms for EEG-based brain-computer interfaces, J. Neural Eng. 4 (2007) http://dx.doi.org/10.1088/1741-2560/4/2/R01.

[6] A. Craik, Y. He, J. Contreras-Vidal, Deep learning for electroencephalogram (EEG) classification tasks: A review, J. Neural Eng. 16 (2019) http://dx.doi.org/10.1088/1741-2552/ab0ab5.

[7] E. Ramdinmawii, A. Mohanta, V.K. Mittal, Emotion recognition from speech signal, in: TENCON 2017 - 2017 IEEE Region 10 Conference, 2017, pp. 1562–1567, http://dx.doi.org/10.1109/TENCON.2017.8228105.

[8] A. El Haj, Emotions recognition in audio signals using an extension of the latent block model, Speech Commun. 161 (2024) 103092, http://dx.doi.org/10.1016/j.specom.2024.103092.

[9] Y. Ülgen Sönmez, A. Varol, In-depth investigation of speech emotion recognition studies from past to present –The importance of emotion recognition from speech signal for AI–, Intell. Syst. Appl. 22 (2024) 200351, http://dx.doi.org/10.1016/j.iswa.2024.200351.

[10] P. Tarnowski, M. Kołodziej, A. Majkowski, R.J. Rak, Emotion recognition using facial expressions, Procedia Comput. Sci. 108 (2017) 1175–1184, International Conference on Computational Science, ICCS 2017, 12–14 June 2017, Zurich, Switzerland, https://doi.org/10.1016/j.procs.2017.05.025.

[11] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, I. Patras, DEAP: A database for emotion analysis using physiological signals, IEEE Trans. Affect. Comput. 3 (1) (2012) 18–31, http://dx.doi.org/10.1109/T-AFFC.2011.15.

[12] W.-L. Zheng, B.-L. Lu, Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks, IEEE Trans. Auton. Ment. Dev. 7 (3) (2015) 162–175, http://dx.doi.org/10.1109/TAMD.2015.2431497.

[13] W.-L. Zheng, W. Liu, Y. Lu, B.-L. Lu, A. Cichocki, EmotionMeter: A multimodal framework for recognizing human emotions, IEEE Trans. Cybern. 49 (3) (2019) 1110–1122, http://dx.doi.org/10.1109/TCYB.2018.2797176.

[14] N.-D. Mai, B.-G. Lee, W.-Y. Chung, Affective computing on machine learning-based emotion recognition using a self-made EEG device, Sensors 21 (15) (2021) http://dx.doi.org/10.3390/s21155135.

[15] J. Atkinson, D. Campos, Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers, Expert Syst. Appl. 47 (2016) 35–41, http://dx.doi.org/10.1016/j.eswa.2015.10.049.

[16] C. Liu, X. Zhou, Y. Wu, R. Yang, Z. Wang, L. Zhai, Z. Jia, Y. Liu, Graph neural networks in EEG-based emotion recognition: A survey, 2024.

[17] X. Qiu, S. Wang, R. Wang, Y. Zhang, L. Huang, A multi-head residual connection GCN for EEG emotion recognition, Comput. Biol. Med. 163 (2023) 107126, http://dx.doi.org/10.1016/j.compbiomed.2023.107126.

[18] Z. Liu, B. Zhu, M. Hu, Z. Deng, J. Zhang, Revised tunable Q-factor wavelet transform for EEG-based epileptic seizure detection, IEEE Trans. Neural Syst. Rehabil. Eng. 31 (2023) 1707–1720, http://dx.doi.org/10.1109/TNSRE.2023.3257306.

[19] D.V. Puri, J.P. Gawande, J.L. Rajput, S.L. Nalbalwar, A novel optimal wavelet filter banks for automated diagnosis of Alzheimer's disease and mild cognitive impairment using Electroencephalogram signals, Decis. Anal. J. 9 (2023) 100336, http://dx.doi.org/10.1016/j.dajour.2023.100336.

[20] D.V. Puri, S.L. Nalbalwar, A.B. Nandgaonkar, J.P. Gawande, A. Wagh, Automatic detection of Alzheimer's disease from EEG signals using low-complexity orthogonal wavelet filter banks, Biomed. Signal Process. Control. 81 (2023) 104439, http://dx.doi.org/10.1016/j.bspc.2022.104439.

[21] R. Remesan, M. Bray, J. Mathew, Application of PCA and clustering methods in input selection of hybrid runoff models, J. Environ. Inform. 31 (2) (2018) 137–152.

[22] M.E. Torres, M.A. Colominas, G. Schlotthauer, P. Flandrin, A complete ensemble empirical mode decomposition with adaptive noise, in: 2011 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, 2011, pp. 4144–4147, http://dx.doi.org/10.1109/ICASSP.2011.5947265.

[23] C. Zhang, S. Wang, C. Yu, et al., A complete ensemble empirical mode decomposition with adaptive noise deep autoregressive recurrent neural network method for the whole life remaining useful life prediction of lithium-ion batteries, Ionics 29 (2023) 4337–4349, http://dx.doi.org/10.1007/s11581-023-05152-2.

[24] N.u. Rehman, H. Aftab, Multivariate variational mode decomposition, IEEE Trans. Signal Process. 67 (23) (2019) 6039–6052, http://dx.doi.org/10.1109/TSP.2019.2951223.
[25] A. Nalwaya, R.B. Pachori, Fourier–bessel domain adaptive wavelet transform-based method for emotion identification from EEG signals, IEEE Sensors Lett. 8 (2) (2024) 1–4, http://dx.doi.org/10.1109/LSENS.2023.3347648.
[26] A. Bhattacharyya, R.B. Pachori, A. Upadhyay, U.R. Acharya, Tunable-Q wavelet transform based multiscale entropy measure for automated classification of epileptic EEG signals, Appl. Sci. 7 (4) (2017) http://dx.doi.org/10.3390/app7040385.
[27] R.B. Pachori, Time-Frequency Analysis Techniques and their Applications, first ed., CRC Press, 2023, http://dx.doi.org/10.1201/9781003367987.
[28] Y. Wang, Y. Peng, M. Han, X. Liu, H. Niu, J. Cheng, S. Chang, T. Liu, GCTNet: a graph convolutional transformer network for major depressive disorder detection based on EEG signals, J. Neural Eng. 21 (3) (2024) 036042, http://dx.doi.org/10.1088/1741-2552/ad5048.
[29] J. Jia, B. Zhang, H. Lv, Z. Xu, S. Hu, H. Li, CR-GCN: Channel-relationships-based graph convolutional network for EEG emotion recognition, Brain Sci. 12 (8) (2022) http://dx.doi.org/10.3390/brainsci12080987.
[30] S.V. Bhalerao, R.B. Pachori, Sparse spectrum based swarm decomposition for robust nonstationary signal analysis with application to sleep apnea detection from EEG, Biomed. Signal Process. Control. 77 (2022) 103792, http://dx.doi.org/10.1016/j.bspc.2022.103792.
[31] B. Chakravarthi, S.-C. Ng, M.R. Ezilarasan, M.-F. Leung, EEG-based emotion recognition using hybrid CNN and LSTM classification, Front. Comput. Neurosci. 16 (2022) http://dx.doi.org/10.3389/fncom.2022.1019776.
[32] S.V. Bhalerao, R.B. Pachori, Clustering sparse swarm decomposition for automated recognition of upper limb movements from nonhomogeneous cross-channel EEG signals, IEEE Sensors Lett. 8 (1) (2024) 1–4, http://dx.doi.org/10.1109/LSENS.2023.3347626.
[33] S.V. Bhalerao, R.B. Pachori, Automated classification of cognitive visual objects using multivariate swarm sparse decomposition from multichannel EEG-MEG signals, IEEE Trans. Hum. Mach. Syst. 54 (4) (2024) 455–464, http://dx.doi.org/10.1109/THMS.2024.3395153.
[34] Z. Fan, F. Chen, X. Xia, Y. Liu, EEG emotion classification based on graph convolutional network, Appl. Sci. 14 (2) (2024) http://dx.doi.org/10.3390/app14020726.
[35] D. Pachori, R.K. Tripathy, T.K. Jain, Detection of atrial fibrillation from PPG sensor data using variational mode decomposition, IEEE Sensors Lett. 8 (3) (2024) 1–4, http://dx.doi.org/10.1109/LSENS.2024.3358589.
[36] S. Bhalerao, R. Pachori, Imagined speech-EEG detection using multivariate swarm sparse decomposition-based joint time-frequency analysis for intuitive BCI, 2024, http://dx.doi.org/10.36227/techrxiv.171624101.13954925/v1.
[37] A. Subasi, T. Tuncer, S. Dogan, D. Tanko, U. Sakoglu, EEG-based emotion recognition using tunable Q wavelet transform and rotation forest ensemble classifier, Biomed. Signal Process. Control. 68 (2021) 102648, http://dx.doi.org/10.1016/j.bspc.2021.102648.
[38] D. Pachori, T.K. Gandhi, Automated human emotion recognition system using TQWT based EEG sub-bands, IEEE Sensors Lett. (2024) 1–4, http://dx.doi.org/10.1109/LSENS.2024.3486708.
[39] S.V. Bhalerao, R.B. Pachori, ESSDM: An enhanced sparse swarm decomposition method and its application in multi-class motor imagery–based EEG-BCI system, TechRxiv (2024) http://dx.doi.org/10.36227/techrxiv.24132564.v2.
[40] D. Pachori, T.K. Gandhi, Automated emotion identification system utilizing EEG bands extracted via wavelet filter banks, in: 2024 15th International Conference on Computing Communication and Networking Technologies, ICCCNT, 2024, pp. 1–6, http://dx.doi.org/10.1109/ICCCNT61001.2024.10725942.
[41] H. Gao, X. Wang, Z. Chen, M. Wu, Z. Cai, L. Zhao, J. Li, C. Liu, Graph convolutional network with connectivity uncertainty for EEG-based emotion recognition, IEEE J. Biomed. Heal. Inform. 28 (10) (2024) 5917–5928, http://dx.doi.org/10.1109/JBHI.2024.3416944.
[42] C. Li, F. Wang, Z. Zhao, H. Wang, B.W. Schuller, Attention-based temporal graph representation learning for EEG-based emotion recognition, IEEE J. Biomed. Heal. Inform. 28 (10) (2024) 5755–5767, http://dx.doi.org/10.1109/JBHI.2024.3395622.
[43] B.C. Biswas, S.V. Bhalerao, A real time based wireless wearable EEG device for epilepsy seizure control, in: 2015 International Conference on Communications and Signal Processing, ICCSP, 2015, pp. 0149–0153, http://dx.doi.org/10.1109/ICCSP.2015.7322758.
[44] R.-N. Duan, J.-Y. Zhu, B.-L. Lu, Differential entropy feature for EEG-based emotion classification, in: 2013 6th International IEEE/EMBS Conference on Neural Engineering, NER, 2013, pp. 81–84, http://dx.doi.org/10.1109/NER.2013.6695876.
[45] L.-C. Shi, Y.-Y. Jiao, B.-L. Lu, Differential entropy feature for EEG-based vigilance estimation, in: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC, 2013, pp. 6627–6630, http://dx.doi.org/10.1109/EMBC.2013.6611075.
[46] M. Miao, L. Zheng, B. Xu, Z. Yang, W. Hu, A multiple frequency bands parallel spatial–temporal 3D deep residual learning framework for EEG-based emotion recognition, Biomed. Signal Process. Control. 79 (2023) 104141, http://dx.doi.org/10.1016/j.bspc.2022.104141.
[47] Z. Wang, X. Xue, Multi-class support vector machine, in: Y. Ma, G. Guo (Eds.), Support Vector Machines Applications, Springer International Publishing, Cham, 2014, pp. 23–48, http://dx.doi.org/10.1007/978-3-319-02300-7_2.
[48] J. Weston, C. Watkins, Support vector machines for multi-class pattern recognition, in: The European Symposium on Artificial Neural Networks, 1999, https://api.semanticscholar.org/CorpusID:15059183.
[49] F.F. Chamasemani, Y.P. Singh, Multi-class support vector machine (SVM) classifiers – An application in hypothyroid detection and classification, in: 2011 Sixth International Conference on Bio-Inspired Computing: Theories and Applications, 2011, pp. 351–356, http://dx.doi.org/10.1109/BIC-TA.2011.51.
[50] V.J. Lawhern, A.J. Solon, N.R. Waytowich, S.M. Gordon, C.P. Hung, B.J. Lance, EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces, J. Neural Eng. 15 (5) (2018) 056013, http://dx.doi.org/10.1088/1741-2552/aace8c.
[51] K. Chen, H. Jing, Q. Liu, Q. Ai, L. Ma, A novel caps-EEGNet combined with channel selection for EEG-based emotion recognition, Biomed. Signal Process. Control. 86 (2023) 105312, http://dx.doi.org/10.1016/j.bspc.2023.105312.
[52] W. Huang, Y. Xue, L. Hu, H. Liuli, S-EEGNet: Electroencephalogram signal classification based on a separable convolution neural network with bilinear interpolation, IEEE Access 8 (2020) 131636–131646, http://dx.doi.org/10.1109/ACCESS.2020.3009665.
[53] C.M. Köll˝ od, A. Adolf, K. Iván, G. Márton, I. Ulbert, Deep comparisons of neural networks from the EEGNet family, Electronics 12 (12) (2023) http://dx.doi.org/10.3390/electronics12122743.
[54] M. Hossin, S. M.N, A review on evaluation metrics for data classification evaluations, Int. J. Data Min. Knowl. Manag. Process. 5 (2015) 01–11, http://dx.doi.org/10.5121/ijdkp.2015.5201.
[55] M. Krzywinski, N. Altman, Significance, P values and t-tests, Nature Methods 10 (2013) 1041–1042, http://dx.doi.org/10.1038/nmeth.2698.
[56] G. Di Leo, F. Sardanelli, Statistical significance: p value, 0.05 threshold, and applications to radiomics—reasons for a conservative approach, Eur. Radiol. Exp. 4 (18) (2020) http://dx.doi.org/10.1186/s41747-020-0145-y.
[57] T. Dahiru, P-value, a true test of statistical significance? A cautionary note, Ann. Ib. Postgrad. Med. 6 (1) (2008) 21–26, http://dx.doi.org/10.4314/aipm.v6i1.64038.
[58] M. Krzywinski, N. Altman, Significance, P values and t-tests, Nature Methods 10 (2013) 1041–1042, http://dx.doi.org/10.1038/nmeth.2698.