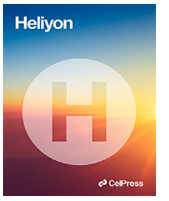
Heliyon 10 (2024) e38681

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Virtual reality-enabled high-performance emotion estimation with the most significant channel pairs

Yaşar Daşdemir

Department of Computer Engineering, Erzurum Technical University, Erzurum, 25050, Turkey

A R T I C L E I N F O

Keywords:
Phase-locking value (PLV)
Channel selection
Virtual reality
Cybersickness
Permutation test
Two-level classification

A B S T R A C T

Human-computer interface (HCI) and electroencephalogram (EEG) signals are widely used in user experience (UX) interface designs to provide immersive interactions with the user. In the context of UX, EEG signals can be used within a metaverse system to assess user engagement, attention, emotional responses, or mental workload. By analyzing EEG signals, system designers can tailor the virtual environment, content, or interactions in real time to optimize UX, improve immersion, and personalize interactions. However, in this case, in addition to the signals' processing cost and classification accuracy, cybersickness in Virtual Reality (VR) systems needs to be resolved. At this point, channel selection methods can perform better for HCI and UX applications by reducing noisy and redundant information in generally unrelated EEG channels. For this purpose, a new method for EEG channel selection based on phase-locking value (PLV) analysis is proposed. We hypothesized that there are interactions between EEG channels in terms of PLV in repeated tasks in different trials of the emotion estimation experiment. Subsequently, frequency-based features were extracted. The features were classified by dividing them into bags using the Multiple-Instance Learning (MIL) variant. This study provides higher classification performance using fewer EEG channels for emotion prediction. The performance rate obtained in binary classification with the Random Forests (RF) algorithm is at a promising level of 99%. The proposed method achieved an accuracy of 99.38% for valence using all channels on the new dataset (VREMO) and 98.13% with channel selection. The benchmark dataset (DEAP) achieved accuracies of 98.16% using all channels and 98.13% with selected channels.

1. Introduction

Multi-modal studies, such as virtual reality (VR) and electroencephalography (EEG) are essential for metaverse environments in many aspects, such as immersive experience, real-time feedback, personalized content and customization, Human-Computer Interaction (HCI), user experience (UX), and engagement [11].

Metaverse's primitives include augmented reality (AR), VR, mixed reality (MR), and extended reality (XR). Combining VR and EEG, which measure brain activity, capturing users' cognitive and emotional responses in real time makes it possible to capture a more immersive and personalized experience in the metaverse.

EEG can detect brain signals associated with different emotional states, such as excitement, stress, or relaxation. Integrating emotion recognition activities with VR and EEG allows real-time monitoring and analysis of users' emotional responses in the virtual

E-mail address: yasar.dasdemir@erzurum.edu.tr.
URL: https://www.erzurum.edu.tr.

https://doi.org/10.1016/j.heliyon.2024.e38681
Received 21 September 2024; Accepted 27 September 2024

Available online 9 October 2024
2405-8440/© 2024 The Author. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

environment. This integration allows the system to provide personalized feedback, adapt the experience, and complete more engaging and emotionally resonant interactions. The metadata warehouse can dynamically adjust its content according to users' emotional states using emotion recognition with VR and EEG. It provides personalized experiences tailored to individual preferences and emotional needs. For example, if a user gets stressed or bored, the system can adapt to the virtual environment or game to reduce stress or increase interaction. Combining VR and EEG opens new possibilities for natural and intuitive HCI. The system can interpret users' intentions, attention levels, or emotional involvement by analyzing EEG signals, providing smoother and more responsive interactions in the metadata store. It can improve the sense of presence and representation, making the experience more realistic and immersive.

Emotion recognition activities with VR and EEG contribute to improving user experience and engagement in metaverse environments. By capturing and responding to users' emotional states, the system can create emotionally rich and engaging experiences that resonate more deeply with the users. This can increase users' overall enjoyment, immersion, and satisfaction in the metadata store.

This study chose PLV as the basis for electrode channel selection due to its ability to capture phase synchronization between EEG signals, which is particularly relevant in emotion recognition. Unlike amplitude-based measures like Pearson correlation, PLV focuses on the consistency of phase differences between signals, providing insights into the functional connectivity of brain regions that are essential for emotional processing.

Some of the challenges in studies in this area include correctly interpreting EEG signals, developing powerful emotion recognition algorithms, providing real-time processing and feedback, and designing user interfaces that effectively integrate VR and EEG technologies while considering user comfort and usability. UX and HCI applications are rapidly becoming more widespread. The importance of affective computing [55], which has been studied for a long time, has become more evident as these applications have become more widespread. In particular, emotion classification methods from objective sources (such as EEG) are widely used in brain-computer interface (BCI) systems [12].

The role of activation of different brain regions in emotion recognition is essential. However, processing large amounts of data obtained from multi-channel systems is difficult and time-consuming. In addition, we cannot say that all these channels play a role in emotion recognition. Therefore, selecting effective channels for emotion estimation, eliminating noise, and removing irrelevant channels are issues that should be evaluated in terms of their performance. For this purpose, an experiment was conducted with 32 participants to classify the emotions obtained from VR environments. The phase-locking method selected the most significant channel pairs, and the effect of these channels on the classification performance was examined. A promising performance of 99 percent was obtained using channel selection and multiple-instance learning methods. In addition, the proposed method was validated using a publicly available DEAP dataset [46].

In summary, the significance of this study lies in its potential to profoundly enhance user experience and interaction within metaverse environments. By integrating EEG and VR, users' cognitive and emotional responses can be monitored in real-time, creating more personalized and immersive experiences. Specifically, EEG's ability to detect brain signals associated with different emotional states allows systems to provide feedback that aligns with users' emotional conditions, adapt virtual environments accordingly, and enhance engagement [45]. This, in turn, makes the metaverse experience more realistic and satisfying for users. Furthermore, advanced signal processing methods like PLV uncover the functional connectivity of brain regions involved in emotional processing, thereby improving the efficacy of emotion recognition algorithms. EEG facilitates the objective and precise measurement of users' emotional states, offering broad applications in brain-computer interface systems.

This paper is organized as follows: Related works are presented in Section 2. Datasets, pre-processing feature extraction, FC, and the MIL method are presented in Section 3. Section 4 presents experimental results. The discussion and conclusions are provided in Sections 5 and 6, respectively.

2. Related work

The electrical activity of the brain recorded as electroencephalography (EEG), provides objective information from individuals. At the same time, it contains rich information regarding emotional states and brain activity. Therefore, researchers have examined emotional states in the valence-arousal of the emotion space using EEG. In emotion recognition studies, common features extracted from EEG signals, including statistical, nonlinear, frequency domain, and deep learning features, are used. Also, the feature extraction process is an essential step in automatic detection systems that separate variables into different classes using machine learning (ML) or deep learning (DL) techniques. Various ML techniques such as SVM, KNN, and ANN and advanced DL models such as CNN, LSTM, GAN, and transformers have been studied [44,77,87,43,42].

Recently, physiological signals that are difficult, if not impossible, to trigger or control have gained attention in emotion recognition. Systems that can work in real-time for emotion recognition have also been proposed [56]. EEG studies on emotion recognition have primarily focused on single-channel approaches. However, many studies on animals and humans have shown that sensorimotor, visual, and cognitive tasks require integrating numerous functional areas widely distributed over the brain.

With the widespread use of mobile EEG systems (such as Emotiv Epoc Flex), the association of EEG-VR has been reflected in current studies [29,76]. With this combination, the limitations of traditional EEG experiments can be reduced to a certain extent. In particular, using mobile EEG systems instead of tethered EEG that limits the participant to a limited area provides both the use of the experience in VR and freedom of movement [76,79].

VR supported brain-computer interfaces (BCIs) are systems that enable users to interact with virtual environments by utilizing brain waves or other neurological signals. These systems aim to integrate neurological signals into VR environments, providing more natural and effective interactions. Such systems are commonly employed in rehabilitation, education, gaming, and other interactive

applications. In real-time, VR can simulate and observe user responses to brain signals, facilitating more immersive and engaging experiences [71,9,52,62].

As VR allows individuals to experience deep immersion, emotion recognition systems are being developed to evoke emotions in VR scenes [58]. In one of these systems, electroencephalogram (EEG) and electrocardiography (ECG) were recorded with 60 participants in a virtual environment. In the arousal-valence combination, they obtained 75% accuracy for arousal with an SVM and 71.21% for valence. In a study [37] conducted using VR and EEG, the stress level was classified using a convolutional neural network (CNN). With a Support Vector Machine (SVM), 96.42% success was achieved using all channels. In another study [18] that tested negative and neutral statuses with fear of heights in 75 participants, more right hemispheric lateralization was found in the negative VR condition. Therefore, VR technology has been reported to provide more ecologically valid ways to evoke emotions. Studies have used VREED, a new publicly available dataset obtained from 3D VR videos [78]. Related research has used differential entropy (DE) features for EEG-based emotion recognition. The classification result of two emotional states (positive/negative) in the gamma band was obtained using SVM with a score of 76.22%.

Presence is a prerequisite for emotional reactions in virtual reality environments [21]. One of the challenges associated with VR is cybersickness (CS). Regarding CS, it is worth noting that some methodological improvements were made to the VREMO dataset [10]. These studies provided valuable insights into managing and reducing CS by ensuring that the measured emotional responses were not confused with discomfort but reflected real emotional experiences triggered by VR stimuli. This demonstrates that the study addresses VR-specific challenges and contributes to advancing emotion recognition research in these innovative environments. Studies on CS have identified the conditions that affect these processes. In particular, teleportation is more comfortable and causes less CS than continuous locomotion [14]. A controlled multi-session motion sickness study [70] using an actuated rotating chair examined the potential of multi-sensory visual and auditory motion cues presented during a VR reading task to mitigate sickness. This study showed that visual cues are most effective in reducing symptoms, and auditory cues have some beneficial effects when combined with visual cues. Evidence has confirmed that the human vestibular system is involved in CS regarding balance and space orientation [51].

The critical frequency bands and channels were analyzed using trained deep belief networks (DBNs) weights. Three emotion models (positive, negative, and neutral) were used in the study [88]. Four profiles with four, six, nine, and twelve channels were selected, and the success rate of these profiles was obtained with the best accuracy rate of 86.65% compared with 62 channels. In another study [35], during the two-class classification of valence and arousal in DEAP, experimental results were obtained with classification rates of 98.93% and 99.10%, respectively. The channel feature used in this study was defined by a symmetric matrix calculated using the Pearson correlation coefficient between the two channel pairs. There are also studies that focus on channel selection [85,3] and salient features for emotion recognition [68,82,20,64]. In these studies, the number of features and channels was reduced using different algorithms, and promising results were obtained.

The proposed channel selection method offers a new approach based on PLV analysis, which differs from the existing literature methods (Table 1). The originality of this method is that PLV is used to determine the interactions between EEG channels more accurately, and these interactions are used to predict emotional states during repetitive tasks. While the channel selection method based on algorithms such as Swarm-Intelligence Algorithms [85] aims to increase the classification performance by reducing the number of channels in a similar way, the proposed PLV-based method provides a more comprehensive and sensitive analysis by considering not only the selection of essential channels but also the phase synchronization between channels. The proposed PLV-MIL method analyzes the neural interactions in more depth, basing the channel selection on signal strength and phase coherence between channel pairs. This enables the classification of more natural and ecologically valid emotional responses in the VR environment and, as a result, provides a higher classification performance.

Functional Connectivity (FC) is related to synchrony of brain activity. When brain regions oscillate in a coordinated fashion, there is a high correlation between their signal time series. Recently, the use of the functional connectivity methods in emotion recognition has become widespread [53,39,50,73,66]. EEG-based functional connectivity allows us to understand the brain areas involved in a particular task. Phase synchronization also plays an important role in many neurological diseases such as epilepsy [60], pathological tremors [75], [59], and schizophrenia [49]. EEG-based functional connectivity lets us understand which electrode positions are more critical in emotion recognition tasks. FC was calculated by considering the similarities between the time series or activation maps. Statistical dependencies in this calculation were used to investigate phase locking [7,63,17,8]. Many methods have been studied, including linear coherence estimation in the frequency domain. Among nonlinear methods, generalized synchronization [74] and phase synchronization have been investigated [75,48,60].

3. Methodology

An experiment was conducted to test the effects of the immersion properties of virtual scenes on emotions and the classification performance of the proposed method on EEG data in a virtual reality environment under various virtual task settings. Fig. 1 shows an overview of the experimental flow. This section provides detailed information regarding the materials and methods used.

3.1. Datasets

The method proposed in this study was used with both VREMO [10], and DEAP datasets and promising results were obtained. The VREMO dataset, which examines the effects of stimuli in the virtual environment on individuals' emotions, was used to test the proposed method. The publicly available DEAP dataset was used to validate the method. Table 2 provides important brief information about these datasets.

Table 1
Summarizing EEG channel selection methods.

Method	Advantages	Disadvantages
Fisher Score [54]	Simple and computationally efficient.	Ignores inter-channel dependencies.
Classifier-Based Methods [33]	Directly optimizes classification performance.	High computational cost; risk of overfitting.
Mutual Information [82]	Can assess dependencies and interactions.	Computationally expensive, especially with a large number of channels.
Genetic Algorithms (GA) [25]	Efficient in searching through a large number of channels.	Prone to local minima; time-consuming.
Feature Ranking and Selection [27]	Flexible in channel selection; evaluates individual channel contributions.	Does not consider interactions between features; results depend solely on ranking.
Principal Component Analysis (PCA) [61]	Reduces dimensionality, speeding up computations.	Potential loss of information; reduced interpretability.
Statistical Methods [5]	Simple and fast; low computational cost.	Fails to capture complex relationships or dependencies.
Swarm Intelligence Methods [85]	Effective in parallel search; potential to reach global optimum.	High computational cost; parameter tuning can be challenging.
Correlation-Based Methods [34]	Evaluates relationships between channels; can be combined with feature selection.	High computational cost; captures only linear relationships.

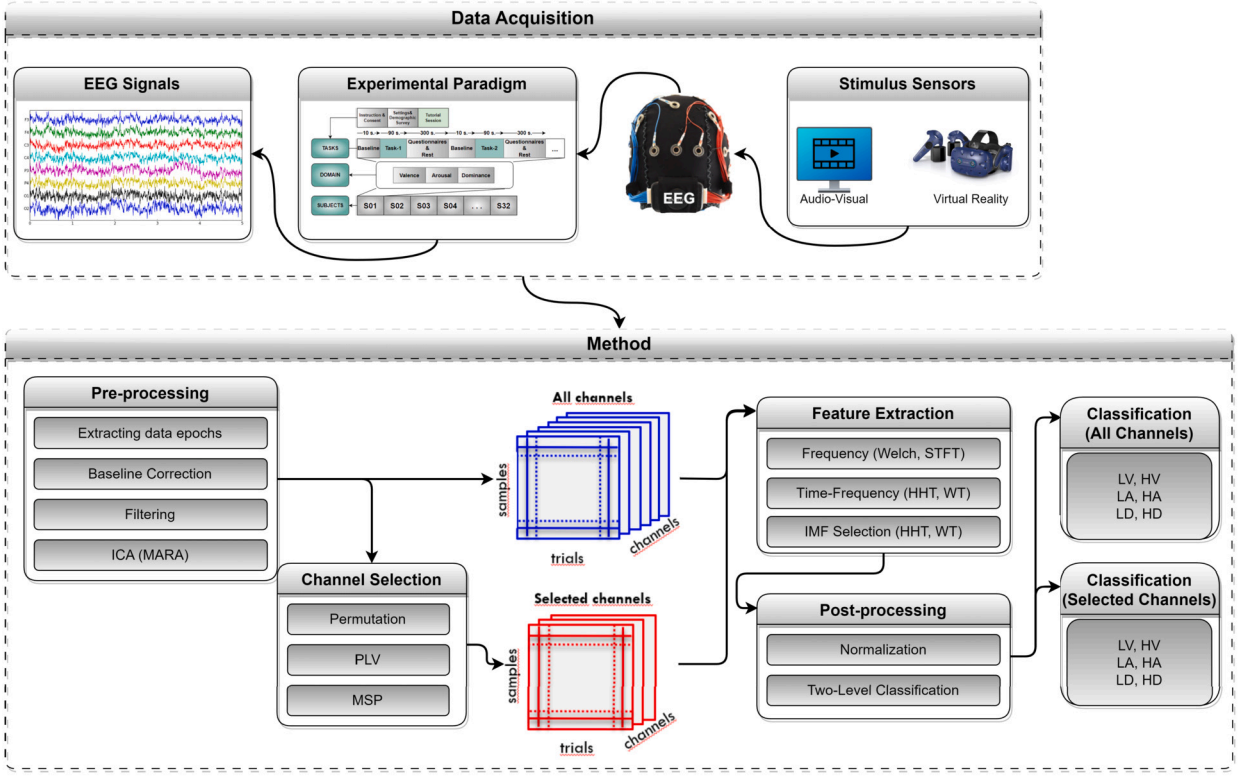


Fig. 1. Flow chart of the proposed method.

Five scenes from VR games were used in the VREEG dataset. The results were evaluated using the experimental protocol and the Self-Assessment Manikin (SAM), Virtual Reality Immersion (VRI), and Simulator Sickness Questionnaire (SSQ).

Table 2
Datasets used in this study.

Database	Number of subjects	Ages	Number of stimuli	Content	Duration	Recording device	Number of channels	Sampling frequency[a]	Dimensional states (1-9)[b]
DEAP	32 (16 Female)	19-37 (mean 26.9)	40	Music videos	60s	Biosemi Active II	32	128 Hz	V, A, D, L, F
VREMO	32 (11 Female)	18-30 (mean 21.4)	5	VR scenes	90s	Emotiv Epoc Flex	32	128 Hz	V, A, D

[a] downsampled.
[b] V: valence, A: arousal, D: dominance, L: liking, F: familiarity.

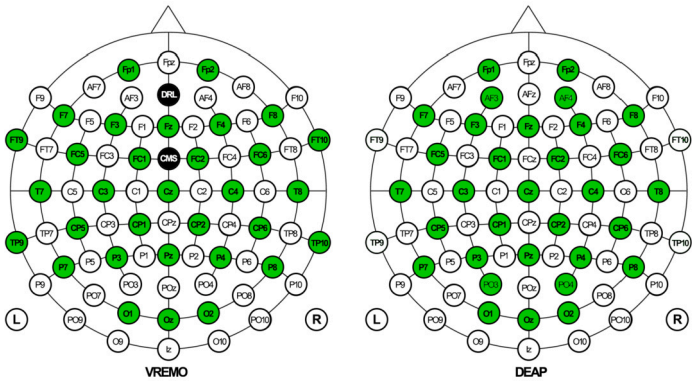


Fig. 2. Electrode positions of datasets.

Table 3
Descriptive statistics of datasets.

Dataset	Emo.Type	Valence	Arousal	Dominance
DEAP	Low	2.55 (0.93)[a]	2.58 (0.95)	2.68 (0.87)
	Neutral	4.84 (0.56)	5.01 (0.62)	4.99 (0.57)
	High	7.34 (0.84)	7.07 (0.84)	7.37 (0.99)
VREMO	Low	2.28 (0.83)	2.11 (0.79)	1.81 (0.84)
	Neutral	4.84 (0.95)	4.82 (0.92)	5.07 (0.88)
	High	8.41 (0.79)	8.19 (0.76)	7.67 (0.83)

[a] mean (std.dev.).

The EEG electrode configuration of both datasets is shown in Fig. 2. The electrodes of the Epoc Flex device were grounded on the forehead with reference to AFz (Driven-Right-Leg sensor, DRL) and FCz (Common-Mode Sensor, CMS). The electrodes in the cap were filled with saline solution to ensure the quality of EEG signals.

3.2. Descriptive statistics

The mean and standard deviations of the valence, arousal, and dominance of the low, neutral, and high emotion types are given in Table 3. This is illustrated in Fig. 3. Plotting a box-plot graph is essential to consider the dimensions' distribution. Therefore, according to the box-plot graph (Fig. 3), there were differences in the range of their evaluations. There are no dimensions with the same median for either dataset, and there are dimensions with different distributions. For example, the VREMO distribution in the valence dimension was closer to a more positive quarter than the DEAP distribution. On the other hand, VREMO assessments span a more comprehensive range, whereas DEAP assessments are closer to the center (neutral). This shows that for DEAP, many participants had similar views on certain parts of the scale, but VREMO's perspectives were more variable. In addition, this indicates that stimuli (such as VR) are subjectively more decisive to participants.

3.3. Pre-processing

The EEG data were processed in MATLAB using the EEGlab data analysis toolbox [16]. Poor contact with the subject's electrodes, sweat, or muscle tension may have caused recording artifacts during the experiments. To eliminate these artifacts, a "baseline interval" that records tens of milliseconds before the stimulus, during which the subject was asked to remain still. In this situation, EEG recordings are assumed to have nothing to do with the given stimulus. The mean signal in the baseline range was subtracted from each channel of the signal in the stimulus range. This process is called "baseline correction." This study uses 10-second baseline intervals

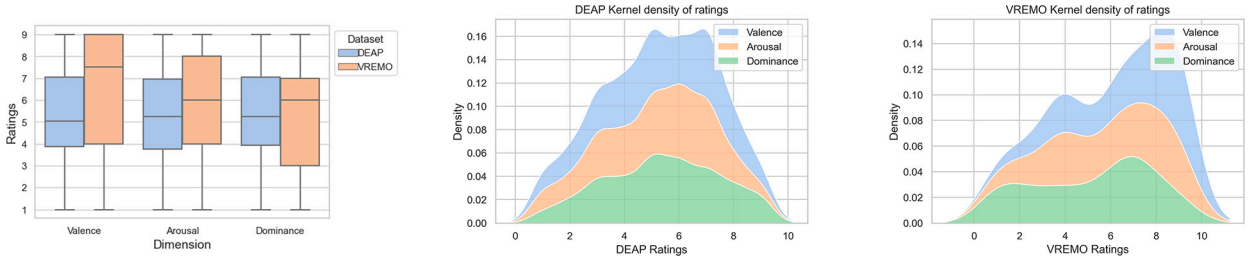


Fig. 3. (a) Boxplot representing valence, arousal, and dominance ratings classified according to the resulting dataset. (b,c) DEAP and VREMO affective dimensions density of ratings. Plot bandwidth adjusts set to 1.

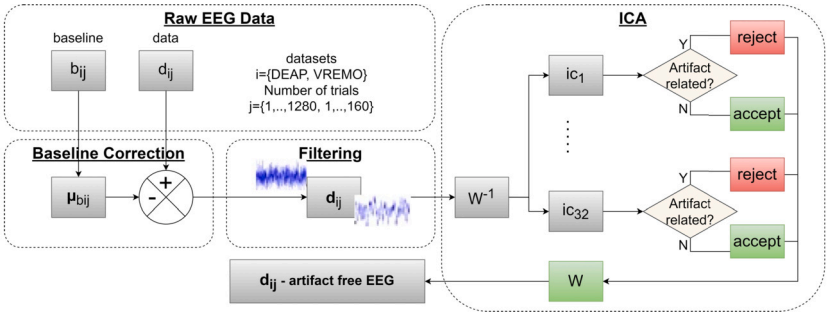


Fig. 4. Block diagram for artifact removal.

before each 90-second stimulus recording for baseline correction. The EEG signals in the datasets were obtained by downsampling and sampling with a sampling frequency of 128 Hz. The 4.0–45.0 Hz frequency range, including four bands (θ, α, β, and γ), was filtered with a 5th-order bandpass Butterworth filter. Infomax-based independent component analysis (ICA) was performed to remove artifacts using the spectra of the components and scalp maps. Muscle activity features related to horizontal/vertical eye movements and artificial components during blinking were excluded from the analysis. The multiple artifact rejection algorithm (MARA) [84] EEGlab plugin guides the ICA-based artifact removal. To preserve data quality, only cases reported above the 90% confidence level were removed, removing approximately to 1-2 components from each participant's data. The MARA is an algorithm used for binary classification problems. The independent components (IC) were evaluated as "accept" or "reject" (Fig. 4). It uses the following methods: Current Density Norm and Range Within Pattern, Fit Error, λ and 8-13 Hz, and Mean Local Skewness.

ICA aims to separate individual source signals from multi-dimensional mixed signals. The recorded EEG signals were collected from 32 channels

$$x(t) = [x_1(t), x_2(t), ..., x_{32}(t)]^T \tag{1}$$

which are linear mixtures of 32 independent sources, $s_i(t)$

$$s(t) = [s_1(t), s_2(t), ..., s_{32}(t)]^T \tag{2}$$

The estimated independent components (ICs) are obtained as

$$x(t) = W s(t) \text{ [31]} \tag{3}$$

where W is an invertible 32x32 linear-mixing matrix. The aim of ICA is to estimate the separation matrix W [65]. ICs are found using:

$$ic_i(t) = W^{-1} x(t) \quad \text{where} \quad i = 1, 2, ..., n \ (n \leq 32) \tag{4}$$

3.4. Feature extraction

Feature extraction plays a crucial role in the classification accuracy. To achieve an acceptable classification accuracy, fewer features representing the signal must be extracted from the raw data. This reduces the signal domain size and computational complexity by removing features that do not affect classification.

In the frequency domain, the relationship between the frequency and amplitude of the signal is established, and the frequency characteristics of the signal are evaluated analytically. Frequency-domain information is commonly used in EEG signals with temporal and spatial dimensions. The frequency-domain properties do not change over time and are less sensitive to noise [47]. Power Spectral Density (PSD) and Short-Time Fourier Transform (STFT) methods were used to extract frequency domain features. PSD was calculated

using Welch's method. The signal was divided into eight segments with a 50% overlap, and each component was windowed with a Hamming window (step size 32). The PSD of each element was averaged.

The short-time Fourier transform (STFT) [28] is a widely used technique for analyzing the frequency content of time-varying signals, such as EEG data. This provides a way to examine how the frequency components of a signal change over time. In the case of EEG signals, STFT allows researchers and clinicians to investigate dynamic changes in brain activity across different frequency bands. This information can be valuable for understanding brain function, identifying abnormal patterns, and diagnosing certain neurological conditions.

$$X(t,f)=\int_{-\infty}^{\infty} w(t-\tau)x(\tau)e^{-j2\pi f\tau}d\tau \text{ [57]} \tag{5}$$

Wavelet Transform (WT) and Hilbert-Huang Transform (HHT) were used in the time-frequency domain.

$$CWT_x^{\Psi'}(\tau,s)=\Psi(\tau,s)=\frac{1}{\sqrt{|s|}}\int_t x(t)\Psi^*\left(\frac{t-\tau}{s}\right)dt \tag{6}$$

$x(t)=$ given signal , tau $=$ translation parameter, $s=$ scaling parameter $=1/\mathrm{f}$, and phi$(t)=$ mother wavelet, all kernels are obtained by scaling and/or translating the mother wavelet.

This study used the Daubechies wavelet function ("db1"), considering 32 EEG channels. The features extracted from each channel consisted of the max/min value, standard deviation, mean value, mean power, and entropy of the coefficients obtained at each decomposition level.

Hilbert-Huang Transform (HHT) consists of empirical mode decomposition (EMD) and Hilbert Transform [30]. EMD breaks down data into components called Intrinsic Mode Functions (IMF). Processing all the IMFs is not required in a nonstationary data series. Therefore, the proper selection of the IMF plays a crucial role. IMFs with a threshold value (11) were selected from the IMFs obtained in this study [13].

Frequency domain methods such as Power Spectral Density (PSD) and Short Time Fourier Transform (STFT) analyze the signal's frequency content; PSD provides overall power distribution, and STFT provides time-resolved frequency information. However, time-frequency approaches such as Wavelet Transform (WT) and Hilbert-Huang Transform (HHT) provide more flexible analysis by capturing both time and frequency characteristics of non-stationary signals such as EEG. While frequency domain methods are helpful for static analysis, time-frequency approaches are more suitable for EEG signals with dynamic changes. In this study, the positive aspects of the effects in both domains were tried to be taken.

Normalization is essential for classification performance as it provides comparable scales, improves convergence speed, improves robustness against outliers, handles algorithm sensitivity, and simplifies the interpretation of feature importance. This enables the models to learn effectively from all features and make more accurate predictions. After feature extraction, z-score normalization was applied to the features in both domains in the next step.

$$x_{\text{scaled}}=\frac{x-\text{ mean}}{sd} \tag{7}$$

In this technique, the values were normalized to the mean and standard deviation of the x data.

3.5. Functional connectivity

Functional connectivity (FC) is the statistical dependency or correlation between brain regions or regions of interest (ROIs), based on neural activity. This provides insights into how different brain areas communicate and work together during various cognitive processes. FC research is an active area of study in neuroscience that contributes to our understanding of brain function and its alterations in multiple contexts. The EEG bands examined for FC in this study are θ (4-7), α (8-13), β (14-30), and γ (31-45).

The synchronization used in this study is called the PLV, which was introduced in [48]. A comparison of PLV with other potential methods for assessing channel importance is given in Table 4. This table overviews various methods for assessing channel significance in EEG analysis and their key strengths, limitations, and typical application areas. To compute the PLV between two signals, namely $s_x(t)$ and $s_y(t)$, instantaneous phase values at the target frequency should be extracted. For this purpose, the signals are first band-pass filtered in the desired frequency band. Then, instantaneous phase values were extracted using the Hilbert Transform (note that phases were extracted using the Gabor Wavelet Transform in [48]). The analytic signal of $s_x(t)$ is defined as

$$z_x(t)=s_x(t)+j\widetilde{s}_x(t)=A_x(t)e^{j\phi_x(t)} \text{ [6]} \tag{8}$$

The analytic signal of $s_y(t)$ is defined as:

$$z_y(t)=s_y(t)+j\widetilde{s}_y(t)=A_y(t)e^{j\phi_y(t)} \tag{9}$$

where A_x and A_y are the instantaneous amplitudes and $\phi_x(t)$ and $\phi_y(t)$ are the instantaneous phases (IPs). $\widetilde{s}_x(t)$ and $\widetilde{s}_y(t)$ are Hilbert transforms of $s_x(t)$ and $s_y(t)$, respectively.

$$\widetilde{s}_x(t)=\frac{1}{\pi}P.V.\int_{-\infty}^{\infty}\frac{s_x(\tau)}{t-\tau}d\tau \tag{10}$$

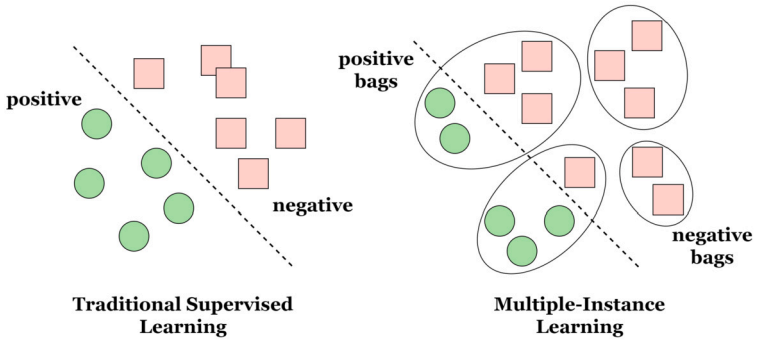


Fig. 5. Multiple-instance learning vs. traditional supervised learning.

$$\widetilde{s}_y(t) = \frac{1}{\pi} P.V. \int_{-\infty}^{\infty} \frac{s_y(\tau)}{t-\tau} d\tau$$

where P.V. denotes the Cauchy principal value. The instantaneous amplitudes are defined by Equation (11).

$$A_x(t) = \sqrt{s_x(t)^2 + \widetilde{s}_x(t)^2} \tag{11}$$
$$A_y(t) = \sqrt{s_y(t)^2 + \widetilde{s}_y(t)^2}$$

IPs are defined as (12).

$$\phi_x(t) = arg[z_x(t)] = arctan\left(\frac{\widetilde{s}_x(t)}{s_x(t)}\right) \tag{12}$$
$$\phi_y(t) = arg[z_y(t)] = arctan\left(\frac{\widetilde{s}_y(t)}{s_y(t)}\right)$$

Finally, PLV, between two signals over an interval of N samples, is computed at time t, as in (13).

$$PLV_t = \frac{1}{N}\left|\sum_{n=1}^{N} exp(j\theta(t,n))\right| \text{ [48]} \tag{13}$$

where $\theta(t,n)$ is the phase difference between the signals $s_x(t)$, $s_y(t)$, $\phi_x(t,n) - \phi_y(t,n)$. The PLV measures how this phase difference changes across trials. If the phase difference is close to zero across trials, then PLV will be close to 1; otherwise, it will be smaller. The PLV is an important synchronization measure when working with biosignals (particularly electrical brain activity). PLV uses narrow-band signals because of the challenges in physically interpreting the instantaneous phase value for wideband signals.

3.6. Two-level classification

Multiple Instance Learning (MIL) is particularly well-suited for EEG emotion recognition due to the inherently variable and complex nature of EEG signals. In traditional classification tasks, each data instance is labeled individually, which can be challenging for EEG data where emotions may not be uniformly expressed across all time segments or channels. MIL addresses this by allowing sets of instances (e.g., multiple time segments of EEG data) to be considered collectively, with a label assigned to the entire set rather than individual instances. This approach is advantageous in emotion recognition, where the emotional state may only be reflected in specific segments of the EEG signal.

MIL is a subfield of machine learning that deals with classification problems in which training data are organized into sets or bags of instances rather than individual instances. In MIL, each bag contains multiple instances, and labels are assigned at the bag level rather than at the instance level. A bag is a collection of instances representing a single data point in MIL. An instance refers to the individual data points within a bag. In MIL, labels are assigned to bags rather than instances. A bag is labeled positive if at least one instance belongs to the positive class and negative if all instances belong to the negative class (Fig. 5). This learning method is computationally more complex than attribute-value learning. This is a natural fit for essential application areas of machine learning, such as the classification of EEG signals and image classification [19,86,83,72,15]. One method used to solve MIL problems is to apply propositionalization, in which bags of data are converted to vectors of attribute-value pairs. One MIL application is the Two-Level Classification (TLC) method presented in [22]. TLC is a proposition method that creates random forests to suggest multi-instance data. This method uses a single decision tree to obtain propositional data. An experimental form of TLC was used in this study. This method uses a random forest instead of a single decision tree to obtain the propositional data.

Propositionalization is the process of converting the complex multi-instance data structure into a standard single-instance format that can be used by traditional machine learning algorithms. Consider an example of synthetic multi-instance data with n Boolean attributes x0, x1, and xn for each bag. Each node and leaf are considered as a region (R), and an instance proposition vector of the length of the number of these two is created. The bag label is then assigned to this vector (Fig. 6). After the attribute set is allocated to the partitions, the membership values of the partitions are determined using the j48 algorithm, which is a decision tree classifier (an

Table 4
Comparison of methods for assessing EEG channel significance.

Method	Description	Advantages	Disadvantages	Application Areas
Phase Locking Value (PLV)	Measures the consistency of phase differences between two signals across trials or time.	- Captures phase synchrony effectively	- Sensitive to noise	Connectivity analysis, seizure prediction, cognitive neuroscience
		- Simple to compute	- Doesn't provide amplitude information	
		- Suitable for connectivity analysis	- May miss non-linear dependencies	
Mutual Information (MI)	Quantifies the amount of information shared between two signals.	- Captures non-linear dependencies	- Computationally intensive	Feature selection, BCI
		- Can handle complex relationships between channels	- May require large data samples for accuracy	
Granger Causality (GC)	Tests whether one-time series can predict another, implying a directional relationship.	- Provides directional information	- Assumes linearity	Causal inference, network analysis, time series analysis
		- Useful for causal inference	- Sensitive to signal length and noise	
Coherence	Measures the degree of correlation between the frequency components of two signals.	- Captures both amplitude and phase information	- Limited to linear relationships	Frequency analysis, neurofeedback, sleep studies
		- Suitable for frequency domain analysis	- Frequency-specific, may miss transient synchrony	
Correlation Coefficient	Measures the linear relationship between two signals.	- Simple to compute	- Only captures linear relationships	Basic statistical analysis, preliminary data exploration
		- Well-understood metric	- Sensitive to outliers and noise	
Cross-Correlation	Measures the similarity between two signals as a function of the time-lag applied to one of them.	- Identifies lagged relationships	- Assumes linearity	Time-lag analysis, signal synchronization, system identification
		- Simple to implement	- Sensitive to noise and outliers	
Canonical Correlation Analysis (CCA)	Assesses the relationship between two sets of variables, often used to find correlated patterns between multi-channel EEG data.	- Captures complex, multi-dimensional relationships	- Computationally intensive	BCI, multimodal data analysis, feature extraction
		- Can handle multiple channels simultaneously	- May require dimensionality reduction	
Independent Component Analysis (ICA)	Decomposes multivariate signals into statistically IC, often used to isolate sources of activity in EEG data.	- Effective in removing artifacts	- Assumes statistical independence	Artifact removal, source localization, signal decomposition
		- Can reveal underlying neural sources	- May miss weak sources or components	
Entropy Measures (e.g., Sample Entropy)	Quantifies the complexity or regularity of time series data, used to assess the non-linear characteristics of EEG signals.	- Sensitive to signal complexity	- Requires large data sets for stability	Complexity analysis, seizure detection, sleep stage classification
		- Can capture subtle changes in signal dynamics	- Computationally expensive	

implementation of the C4.5 algorithm). Once the data is propositionalized (i.e., converted into a single-instance format with a single vector representing each bag), it can be used as input for traditional machine-learning algorithms. At this point, the Random Forest (RF) classifier is used for propositional data. [22]. RF is an ensemble learning algorithm with multiple decision trees. Therefore, they exhibit high performance in MIL methods. The transformed instances were then used to train the RF classifier, which can predict new instances.

An exemplary pruned tree structure obtained from the valence classification in the time-frequency domain of the VREMO dataset is shown in Fig. 7. The resulting pruned tree had nine leaves and was 17 in length.

4. Results

Experimental results are presented in this section. Python, C#, MATLAB, and SPSS were used for the programming, statistical analysis, and graphing. First, the statistical significance was evaluated using PLV, and the most significant channel pairs were determined. The classification results are presented before and after the most significant channel pairs were removed from the datasets.

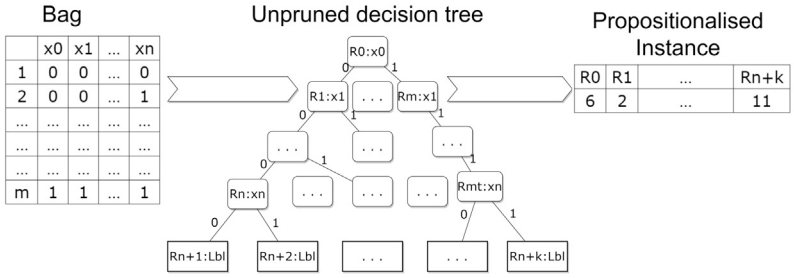


Fig. 6. Converting a bag of instances to a propositionalized vector.

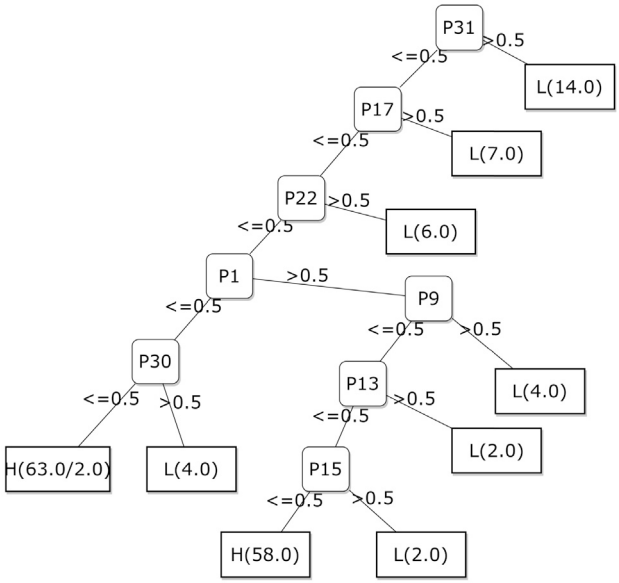


Fig. 7. Pruned decision tree used for VREMO dataset Valence classification (P: partition, L: low valence, H: high valence).

4.1. Statistical significance

The 3D emotion model consists of three dimensions: Valence, Arousal, and Dominance (in short, VAD). This study investigated the interactions between brain regions through PLV for VAD in three-dimensional space in θ, α, β, and γ bands.

The phase-locking values for all electrode pairs were obtained for the VAD. The valence was negative (V< 4), neutral ($4 \leq V \leq 6$), and positive (V> 6). For arousal, calm (A< 4), neutral ($4 \leq A \leq 6$), and excited (A> 6). Dominance was dominated (D< 4), neutral ($4 \leq D \leq 6$), and dominant (D> 6).

For VAD analysis, threshold values 0.55 were used for the DEAP dataset and 0.65 for the VREMO dataset. Both threshold values were determined to reveal at least one channel pair in each frequency band. This difference was used to determine the most significant channel pairs and frequency bands according to the characteristics of each dataset. Both datasets (DEAP and VREMO) were collected under different experimental conditions, which may lead to differences in signal noise, number of participants, devices used, and other environmental factors. Such factors may affect signals' connectivity strength (PLV) in different brain regions. Therefore, a particular threshold value may give different results in other datasets. While 0.55 was a sufficient threshold value to detect significant channel pairs in all frequency bands in the DEAP dataset, this value had to be increased to 0.65 for the VREMO dataset.

Permutation testing is a non-parametric statistical method used to determine the significance of an observed effect. Surrogate data are generated to evaluate whether the observed PLV significantly differs from what might occur by chance. These surrogates are created by shifting the phase of one of the signals (e.g., through temporal shifts) to break the actual temporal relationship between the signals while preserving their characteristics. This process generates PLV values that should represent what would be expected if there were no proper connectivity between the signals. The number of surrogates (in this case, 100) is often chosen to balance computational efficiency and statistical robustness. While using more surrogates could provide a more precise estimate of the null distribution (i.e., the distribution of PLV values under the assumption that there is no actual connection between the signals), 100 is often sufficient to give a reliable indication of statistical significance without high computational cost. The key is that the observed PLV must consistently exceed the surrogate PLVs to be considered significant.

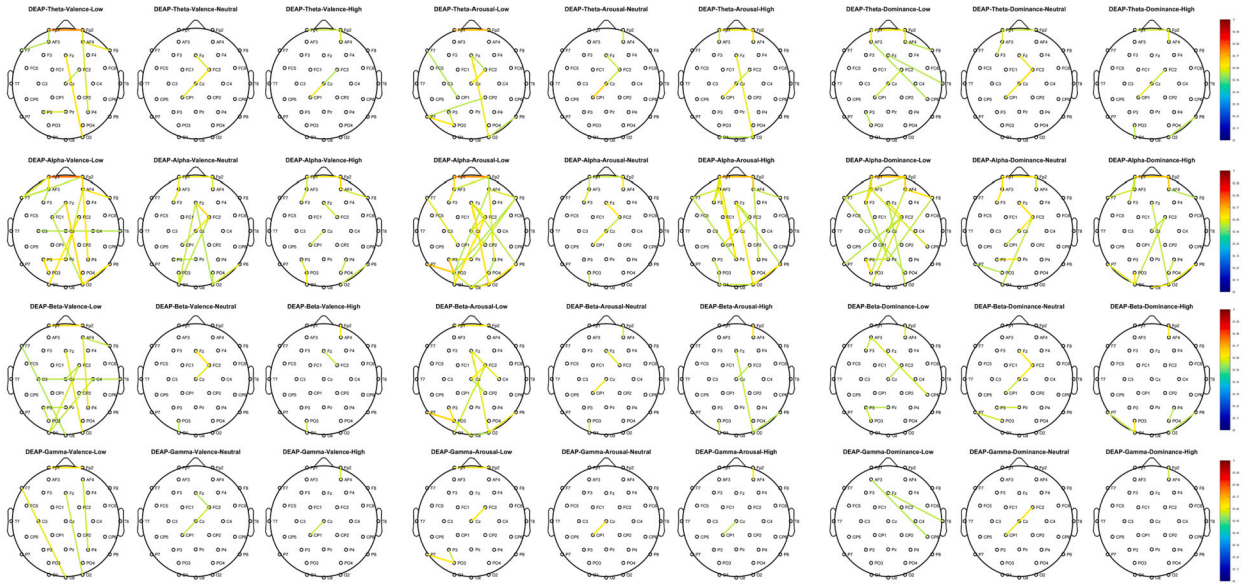


Fig. 8. (DEAP) Significant Phase Locking Values Between electrodes exceeding 0.55 threshold value for VAD. Rows represent oscillations in θ, α, β, and γ bands, and columns represent VAD in low, neutral, and high.

To determine the statistical significance of each PLV, they were compared with the PLVs obtained between shifted trials [48]. The surrogate values were obtained by calculating the phase differences over the shifted trials, as provided by Equation (14).

$$PLV_t^{surrogate(i)} = \frac{1}{N}\left|\sum_{n=1}^{N} e^{j(\theta_x(t,n)-\theta_y(t,n_{perm(i)}))}\right| \tag{14}$$

Using 100 surrogate values means the phase differences were computed 100 times with a different random shift to create 100 PLV values representing the "null distribution". The observed PLV values (from the actual data) are then compared against this null distribution to assess how likely the observed values could have occurred by chance. If the observed PLV is more significant than most or all of the surrogate values, it suggests that the observed PLV is statistically significant and not due to random chance.

Permutation testing with 100 surrogate values revealed that most channel pairs had PLV values that were significantly greater than chance ($p < 0.001$). Fig. 8 shows the strong PLVs ($PLV \geq 0.55$) for the DEAP dataset.

Similarly, for the VREMO dataset, permutation testing with 100 surrogate values revealed that most channel pairs had PLV values that were significantly higher than chance ($p < 0.001$). Fig. 9 shows the PLVs ($PLV \geq 0.65$) for the VREMO dataset.

4.2. MSP and channel selection strategy

It is essential to clarify that the selection of channels in this study is not arbitrary but is guided by the differences in PLV corresponding to the emotional states' valence, arousal, and dominance parameters. The PLV-based approach allows for identifying channels that exhibit significant connectivity differences across these emotional dimensions, ensuring that the selected channels are those most relevant to the nuanced dynamics of emotion recognition. This method is grounded in the principle that different emotional states manifest distinct connectivity patterns in the brain, which PLV can effectively capture, leading to more accurate and meaningful channel selection for emotion recognition tasks.

PLV values highlight the functional connectivity between brain regions for specific tasks. This study explored connection patterns for channel selection and negative and positive emotions through PLV. Previous studies have shown that most reactive bands have higher event-related desynchronization success rates [69,80]. In these studies, the most reactive band is defined as the frequency band corresponding to the largest average power difference. Similarly, the most reactive channel pair is defined as the channel pair with the largest difference between a task-related PLV and the rest state [26,38]. Here, we define the most significant channel pair (MSP) as the channel pair for which the difference between PLV values for negative/positive emotional state vs. neutral state has the largest value.

The PLVs for each channel pair were tested for significance by permutation test. The original PLVs were compared with PLVs calculated using phase differences calculated over randomly shuffled trials. Most channel pairs were confirmed to be significant in the permutation test, with alpha = 0.01, indicating that the PLVs were significantly greater than chance.

$$MSP_e = \arg\max_i |PLV_e^i - PLV_n^i| \tag{15}$$

where PLV_e^i and PLV_n^i show the PLV values for the emotional and neutral cases for i^{th} channel, respectively.

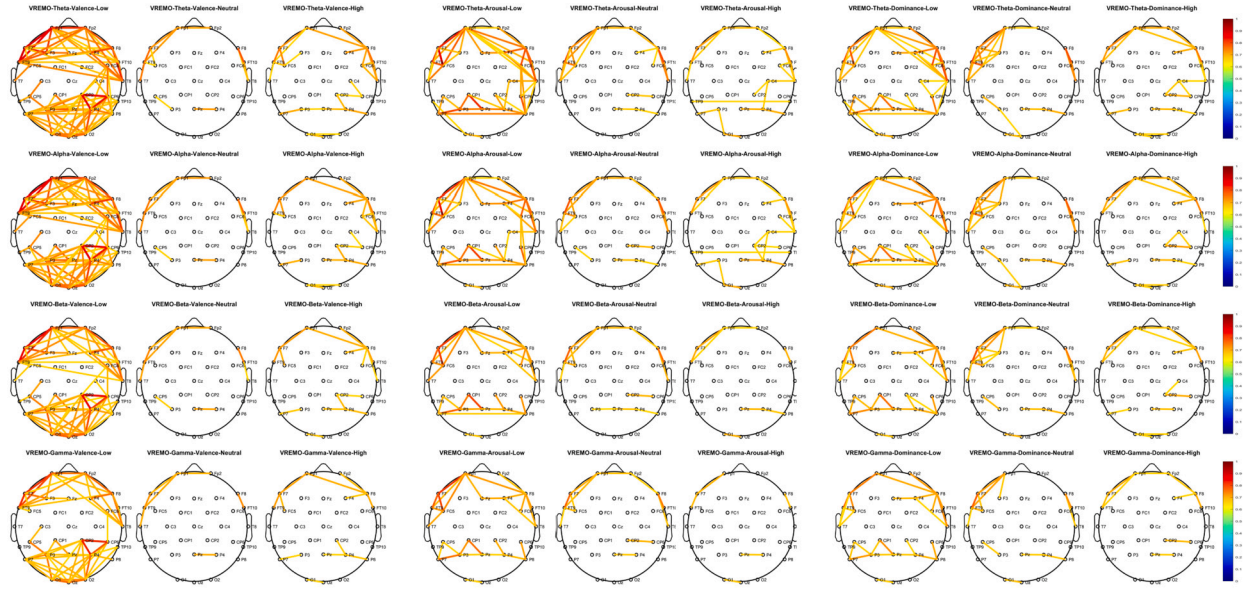


Fig. 9. (VREMO) Significant Phase Locking Values Between electrodes exceeding 0.65 threshold value for VAD. Rows represent oscillations in θ, α, β, and γ bands, and columns represent VAD in low, neutral, and high.

For the DEAP and VREMO datasets, the most significant channel pairs were determined separately for each emotional state in the θ, α, β, and γ bands, and the PLV values for these pairs are shown in Fig. 10 and 11.

For DEAP, the negative and positive emotions were usually close together. However, in the α band, there was a distinct difference between the valence and arousal. Fig. 10 shows that the most considerable difference between the PLV values of MSPs was observed for negative emotions in the α-band. For DEAP, 'Fp1', 'F7', 'C3', 'CP1', 'P7', 'O1', 'Pz', 'Fp2', 'AF4', 'Fz', 'F8', 'FC2', 'Cz', 'T8', 'CP2', 'P8', 'O2' channels revealed as MSPs.

For VREMO, in most cases, negative and positive emotions were similar (Fig. 11). However, the difference in the valence dimension between positive and negative emotions was found to be significant. Because stimuli in an immersive environment were used as stimuli in the VREMO dataset, the initial experiences may have had adverse effects on the participants. For VREMO, 'Cz', 'Fp1', 'F7', 'FC1', 'C3', 'T7', 'TP9', 'P3', 'P7', 'O2', 'P8', 'P4', 'CP6', 'TP10', 'T8', 'FC6', 'F8' channels revealed as MSPs.

4.3. Cybersickness status

The commonly used questionnaire in experiments measuring CS is the Simulator Sickness Questionnaire (SSQ) [41]. The SSQ consists of 16 items on a 4-point Likert scale. The subjects rated each item on a scale of 0 (none), 1 (slight), 2 (moderate), and 3 (severe). The SSQ is based on the following three sub-components:

- (N) Nausea factors (for example, general discomfort, increased salivation, sweating, nausea, difficulty concentrating, stomach awareness, burping).
- (O) Oculomotor factors (for example, general discomfort, fatigue, headache, eyestrain, difficulty focusing, difficulty concentrating, blurred vision).
- (D) Disorientation factors (for example, difficulty focusing, nausea, fullness of the head, blurred vision, dizziness, and vertigo).

Each of the three sub-components was summed. The total SSQ score was calculated as: $(N + O + D) \times 3.74$.

The descriptive values of the five VR scenes used in this study are listed in Table 5. While the most positive emotions were found in the TL and PVR excerpts, the most negative emotions were found in the TWD excerpts. Arousal levels were highest for the PNI, PVR, and RC. The dominance value of most VR scenes was gathered at the middle points.

TL caused less nausea among the excerpt sources than the other scenes (Fig. 12). The RC and PNI scenes caused more cybersickness in the vestibular system.

4.4. Classification results

The binary classification results were very promising due to the MIL method's success with RF (Table 6). RF was used with the following hyperparameters: the number of iterations was 100, the maximum depth of the tree was unlimited, the random number seed to be used was 1, and the percentile size of each bag was chosen as 100. After configuring the RF classifier, a 10-fold cross-validation method was used to estimate the model's success.

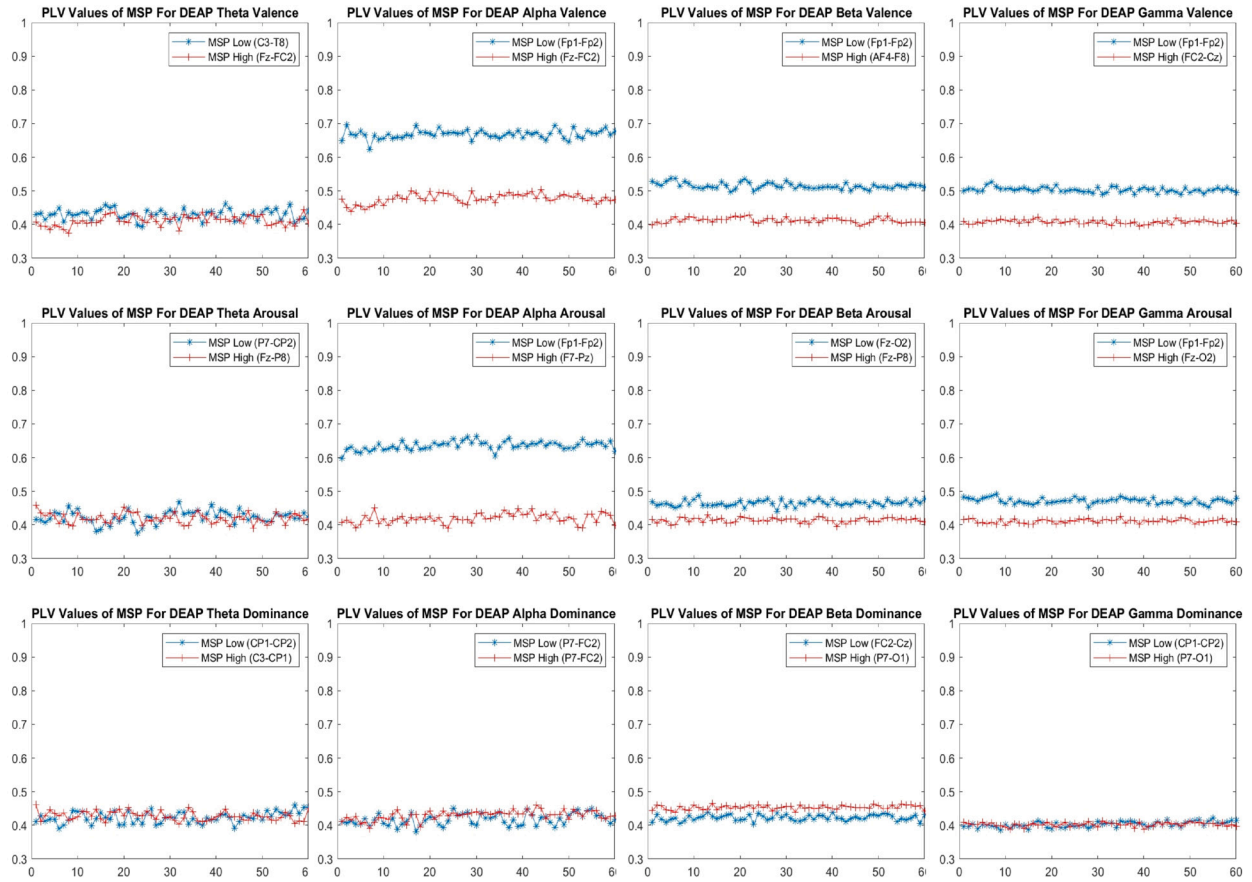


Fig. 10. (DEAP)) Most significant channel pairs between electrodes. Rows represent VAD dimensions; columns represent oscillations in θ, α, β, and γ bands (Vertical axes show PLV values, while horizontal axes represent trial durations in seconds).

Table 5
Comparison of descriptive statistics between VR scenes.

Excerpt's source	Short names	V[a]	A	D
The Walking Dead	TWD	3.156 (1.050)[b]	5.531 (2.229)	5.906 (2.557)
Plank Not Included	PNI	7.063 (1.831)	6.719 (2.261)	5.438 (2.422)
The LAB	TL	8.313 (0.998)	4.156 (2.503)	3.188 (2.348)
Powder VR	PVR	8.063 (1.366)	6.188 (2.416)	5.500 (2.423)
Epic Roller Coaster	RC	7.219 (1.963)	6.813 (1.857)	5.781 (2.324)

[a] V: valence, A: arousal, D: dominance.
[b] mean (std.dev.).

Table 6
Comparison of accuracies for all and selected channels.

		DEAP				VREMO			
Domain		All (Acc)[a]	All (F1)[b]	Sel[c] (Acc)	Sel (F1)	All (Acc)	All (F1)	Sel (Acc)	Sel (F1)
Frequency	V	97.969	0.977	**98.125**	0.978	96.875	0.940	**98.125**	0.976
	A	97.656	0.972	97.891	0.974	**99.375**	0.991	98.750	0.983
	D	98.203	0.976	97.891	0.972	98.750	0.984	96.875	0.984
Time-Frequency	V	**98.359**	0.981	96.406	0.918	99.375	0.988	98.125	0.963
	A	98.203	0.978	97.656	0.971	98.125	0.974	**98.750**	0.983
	D	97.734	0.970	97.266	0.964	**100.000**	1.000	96.875	0.961

[a] Acc.: Accuracy (%).
[b] F1: F1-Score.
[c] Sel: Selected.

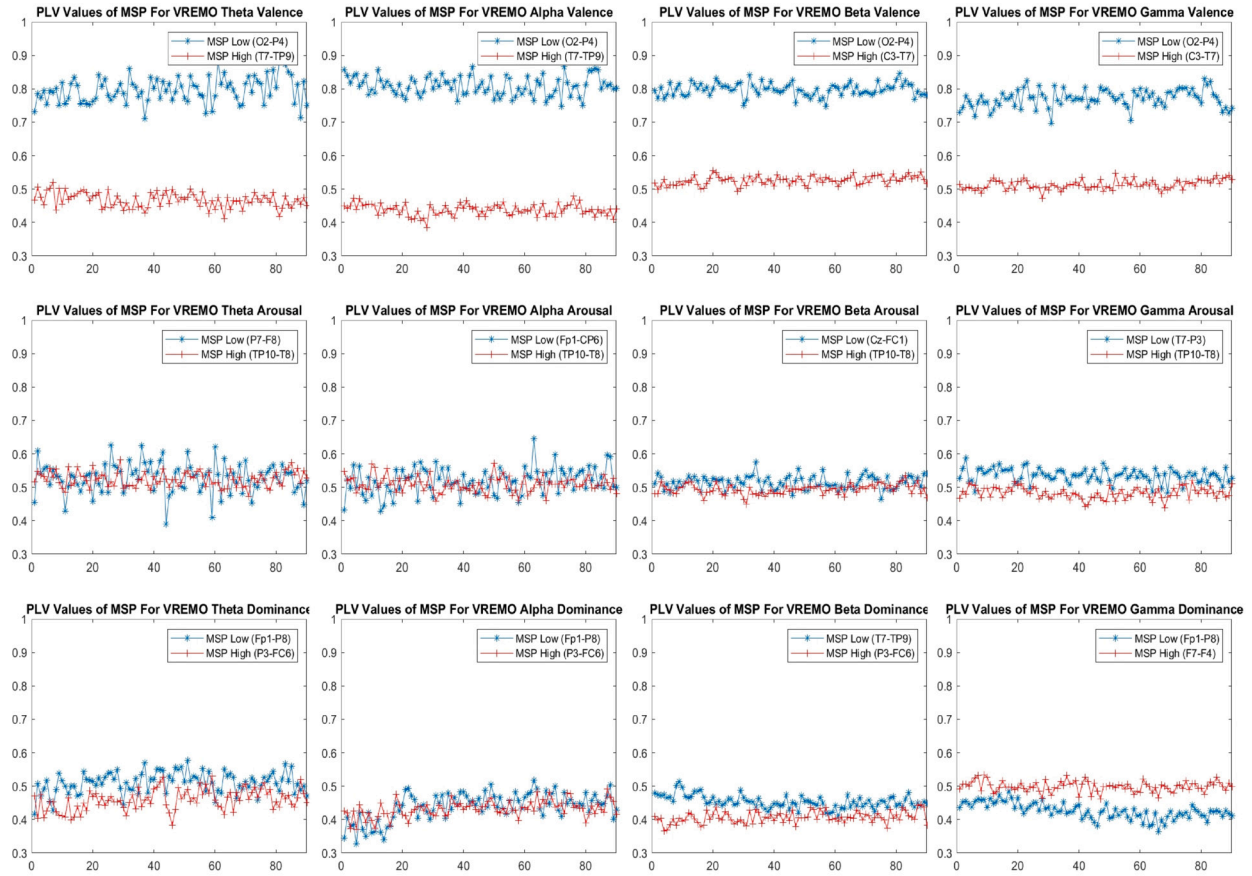


Fig. 11. (VREMO)) Most significant channel pairs between electrodes. Rows represent VAD dimensions; columns represent oscillations in θ, α, β, and γ bands (Vertical axes show PLV values, while horizontal axes represent trial durations in seconds).

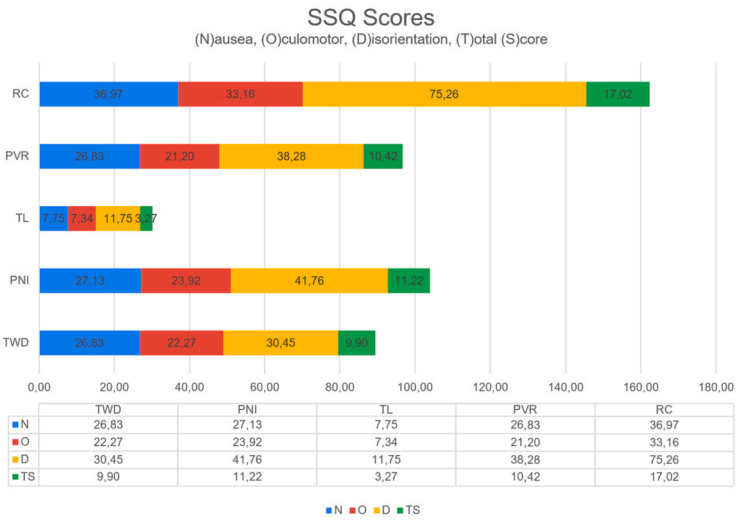


Fig. 12. Total simulator sickness scores.

The features obtained were as follows: 8736 attributes for DEAP and 12576 for VREMO were extracted when all channels were used in the frequency domain. When the selected channels were used, 4641 for DEAP and 6681 for VREMO were extracted. In the time-frequency domain, 1024 features were extracted when all channels were used, and 544 features were removed when only selected channels were used, which was the same for both datasets. The HHT and WT features were combined in the time-frequency domain.

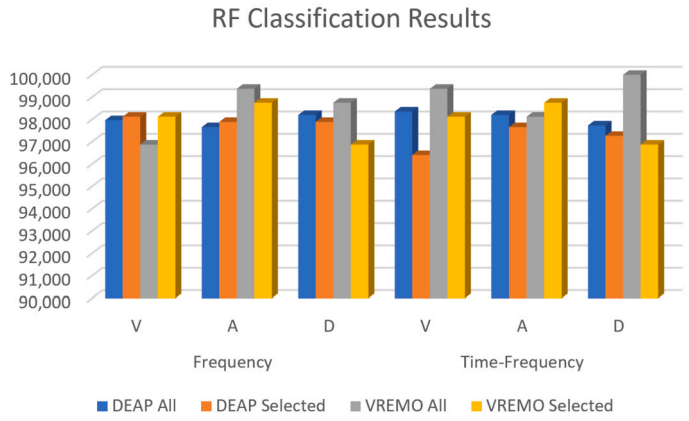


Fig. 13. Random Forest (RF) classification accuracies.

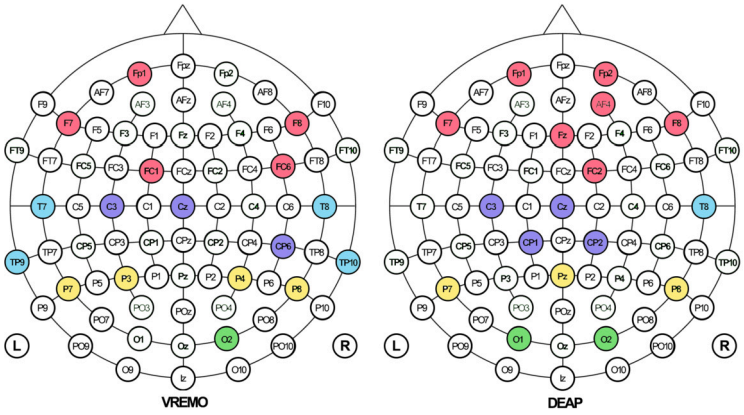


Fig. 14. Active channels selected in classification for both datasets.

For the DEAP dataset, channel selection increased the performance in the valence and arousal dimensions in the frequency domain but decreased it in all dimensions in the time-frequency domain. However, this decrease remained reasonable, and the success achieved at these levels with 17 channels was promising. Channel selection achieved high performance in the frequency domain for the VREMO dataset only in the valence dimension. In the time-frequency domain, high performance was obtained in the arousal dimension. For both datasets, 17 channels were selected, and the classification performance of these channels was remarkable (Fig. 13).

5. Discussion

The method aims to reduce unnecessary and noisy data in HCI and UX applications by determining the interactions between EEG channels through PLV analysis. The validity of the study is verified, especially with permutation tests. The study shows that high classification performance can be achieved in emotion prediction even with fewer EEG channels. The study investigated the interactions between the EEG channels for positive (high), negative (low), and neutral emotions in the θ, α, β, and γ bands, respectively, using the phase-locking values. For this purpose, we created an emotional EEG database using stimuli from a VR environment. ANOVA analysis was conducted between the conditions (low, neutral, and high). In addition, the performance of the database we created was tested on a widely used database for verification.

The permutation test evaluated the PLVs of the channel pairs in both datasets for significance. The actual PLVs were compared with PLVs calculated from randomly shuffled trial tags. Most channel pairs were confirmed to be significant in the permutation test, with alpha = 0.01. An alpha value of 0.01 indicated that the PLVs were significantly greater than chance. As a result, 17 channels were chosen for both datasets (Fig. 14).

Network attributes were extracted using brain activity connectivity analysis in emotion recognition [81]. This study obtained more significant channel pairs for all dimensions of the α channel in the DEAP dataset. For the VREMO dataset, more significant channel pairs were obtained in the θ band at the low valence level.

Table 7
Impact of the method on classification performance based on the Valence parameter (with all channels).

	Non-MIL		MIL	
Datasets	Acc (All)	F1 (All)	Acc (All)	F1 (All)
DEAP	70.625	0.634	98.359	0.984
VREMO	73.125	0.728	99.375	0.994

Table 8
Comparison with the state-of-the-art in the time-frequency domain.

Work	Database	Elect[a]	Classifier	Features	V[b] (%)	A (%)	D (%)
[23]	DEAP	32	kNN	Power spectral densities, entropy, band powers	89.83	89.84	N/A
[2]	DEAP	32	RNN	Temporal patterns, spectral features	85.65	85.45	N/A
[24]	DEAP	Adaptive	RNN	Temporal and frequency domain features	85.88	87.71	86.63
[78]	VREED	59	SVM	Differential entropy	76.22	N/A	N/A
[58]	OWN	N/A	SVM	EEG power, heartbeat dynamics	71.21	75.00	N/A
[35]	DEAP	32	LSTM	Channel-wise features	98.93	99.10	N/A
[36]	IDEA	24	BiLSTM	Spectral features, time-domain features	73.50	75.00	N/A
[4]	DEAP	32	SVM	Temporal and frequency domain features	91.20	93.7	N/A
[40]	DREAMER	14	SVM	EEG and ECG signal features	62.49	62.17	61.84
W[c]	VREMO	32	RF	EEG frequencies, time-frequency features	**99.38**	98.13	100.00
W	VREMO	17[d]	RF	EEG frequencies, time-frequency features	**98.13**	98.75	96.88
W	DEAP	32	RF	EEG frequencies, peripheral signals, multimedia content analysis	**98.36**	98.20	97.73
W	DEAP	17[d]	RF	EEG frequencies, peripheral signals, multimedia content analysis	**96.41**	97.66	97.27

[a] Electrodes#.
[b] V: valence, A: arousal, D: dominance.
[c] W: This work.
[d] Selected channels.

5.1. Summary of proposed classifier's performance

We achieved high classification performance with MIL from multi-channel EEG signals collected with emotional stimuli. We eliminated irrelevant channels and channels carrying noise and extracted the most significant channel pairs with the PLV. When statistical significance was used in the channel selection, the classification performances were close to those obtained with all channels. Higher performance was achieved in the valence dimensions in the time-frequency domain (Table 7). To avoid complex classification processes and eliminate time costs, we aimed to make the proposed channel selection a high-performance algorithm adaptable to real-time applications. The proposed system also performed well on the widely used benchmark dataset (DEAP).

For the most significant channel pairs, when the threshold PLV value was 0.55 for DEAP (Fig. 8) and 0.65 for VREMO (Fig. 9), the channel pairs appeared in all bands. Channel pairs did not appear in some bands at values below these threshold values. For this reason, because each band is thought to contribute to the classification, these threshold values containing at least one channel pair in all bands were selected. Analyzes resulted in 17 channels selected for both datasets (Fig. 14); For DEAP, Fp1, Fp2, F7, Fz, F8, and FC2 in the frontal region, C3, Cz, CP1, and CP2 in the central region, T8 in the temporal region; P7, Pz, and P8 in the parietal region, and O1 and O2 in the occipital regions. For VREMO, Fp1, F7, F8, FC1, and FC6 in the frontal region, C3, Cz, and CP6 in the central region; T7, T8, TP9, and TP10 in the temporal region, P7, P3, P4, and P8 in the parietal region; and O2 in the occipital regions. These results confirm the importance of emotion recognition in multiple brain areas [67]. Previous studies have demonstrated the relationship between these selected channels and emotions [32,1].

A comparison of the data with other binary classification studies is presented in Table 8. The effect of the classifier type on the final classification accuracy created a noticeable difference between the RF used in our study and the methods used in other studies. The fact that the RF classifier performed well on certain features and channels shows that combining these methods plays an important role in the final accuracy. Although other methods have achieved similar results using different features and channels, the high accuracy rates of RF reveal how critical a factor the choice of classifier is. This situation emphasizes that channel selection, feature extraction methods, and classifier preferences all impact classification accuracy. [23] can be examined in future research.

5.2. Limitations of the current study

Because VR is a topic that includes up-to-date and uncommon datasets, there are limited datasets in this area. The dataset used in this study was created to test classification performance and contained a limited number of stimuli. However, on a benchmark dataset (DEAP), the method was also tested on datasets containing only standard audio/visual stimuli. Another limitation is that the study was not tested using EEG bands. The classification performance with the signal containing all bands is promising; therefore, these bands can be considered for future studies.

Binary classification allowed the basic principles and performance of the method to be demonstrated. The binary classification provided an ideal starting point to illustrate the effectiveness and applicability of the method more understandably. However, the method can also be applied to more complex multiple-classification problems in the future.

5.3. Recommendations for future research

RF is an ensemble learning algorithm with multiple decision trees. Therefore, they exhibit high performance in MIL methods. However, the implementation should also be measured using classifier algorithms with different structures. Additionally, this method can be applied to multiple classifications. However, HHT is not a suitable feature extraction method in real time. Instead, strategies that respond faster can be tested. In addition, CS in VR systems is a challenge that must be solved. Therefore, determining and classifying the causes of CS are necessary for future studies.

This study used the method with commonly employed features such as STFT, PSD, and HHT. The success of the process can be tested with techniques like Wigner-Ville Distribution (WVD), Cohen's Class (CC), Variational Mode Decomposition (VMD), Adjustable Q Wavelet Transform (AQWT), and Fractional Fourier Transform (FAWT) in future studies.

6. Conclusion

It is important to emphasize that using VR stimuli for emotion recognition adds significant value to the field by providing a more immersive and ecologically valid environment for evoking emotions. While traditional datasets often rely on less engaging stimuli, VR allows for a richer, more dynamic interaction that can evoke more natural emotional responses. Furthermore, while PLV-MIL's methodology is effective across different datasets, including VR-induced emotions, it offers unique insights that extend beyond traditional approaches.

This study investigated the binary classification performance of channel selection from EEG signals to classify a subject's emotional states in a VR environment. For this purpose, virtual scenes containing stimuli falling in four regions of the valence-arousal space were used in a 6DoF-supported VR environment. The experiment was performed on 32 healthy adult volunteers using the Emotiv EPOC Flex wireless EEG head system for the EEG signal. Next, we used conventional machine learning algorithms to classify the binary emotion status with features in both the frequency and time-frequency domains. All the channels were used in the initial classification process. Only the MSP was used in the second classification process. The results from both cases were promising. These results greatly encourage further study of VR solutions' popularity and widespread use.

The classification results in the frequency and time-frequency domains, and the binary classification performance exceeded 98% for DEAP and 99% for VREMO. With the use of the RF classifier, the positive contribution of the MIL method to performance was high in both cases [22]. The overall results were optimistic and surpassed those found in similar studies on emotion classification using EEG signals. Therefore, it can be concluded that VR induces significant emotional engagement, leading to a better classification rate. The proposed method shows high classification performance in emotion prediction even with fewer EEG channels. In particular, the proposed method performs 99% in binary classification with the RF algorithm, revealing the effectiveness and innovation of the method.

CRediT authorship contribution statement

Yaşar Daşdemir: Conceptualization, Data curation, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] S.M. Alarcao, M.J. Fonseca, Emotions recognition using eeg signals: a survey, IEEE Trans. Affect. Comput. 10 (2017) 374–393.
[2] S. Alhagry, A.A. Fahmy, R.A. El-Khoribi, Emotion recognition based on eeg using lstm recurrent neural network, Int. J. Adv. Comput. Sci. Appl. 8 (2017).
[3] Z.A.A. Alyasseri, O.A. Alomari, S.N. Makhadmeh, S. Mirjalili, M.A. Al-Betar, S. Abdullah, N.S. Ali, J.P. Papa, D. Rodrigues, A.K. Abasi, Eeg channel selection for person identification using binary grey wolf optimizer, IEEE Access 10 (2022) 10500–10513.
[4] M. Aslan, Cnn based efficient approach for emotion recognition, J. King Saud Univ, Comput. Inf. Sci. 34 (2022) 7335–7346.
[5] V.P. Balam, S. Chinara, Statistical channel selection method for detecting drowsiness through single-channel eeg-based bci system, IEEE Trans. Instrum. Meas. 70 (2021) 1–9.
[6] R.N. Bracewell, The Fourier Transform and Its Applications, 2nd edition ed., McGraw-Hill, New York, 1986.
[7] S.L. Bressler, Large-scale cortical networks and cognition, Brains Res. Rev. 20 (1995) 288–304, https://doi.org/10.1016/0165-0173(94)00016-I.
[8] A. Brovelli, M. Ding, A. Ledberg, Y. Chen, R. Nakamura, S.L. Bressler, Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality, Proc. Natl. Acad. Sci. USA 101 (2004) 9849–9854, https://doi.org/10.1073/pnas.0308538101.
[9] W. Chen, Y. Zhang, W. Li, A novel brain-computer interface system integrated with virtual reality for cognitive rehabilitation, J. Neural Eng. 19 (2022) 036023, https://doi.org/10.1088/1741-2552/ac556e.
[10] Y. Daşdemir, Classification of emotional and immersive outcomes in the context of virtual reality scene interactions, Diagnostics 13 (2023) 3437.
[11] Y. Daşdemir, Locomotion techniques with eeg signals in a virtual reality environment, Displays 80 (2023) 102538.

[12] Y. Daşdemir, A brain-computer interface with gamification in the metaverse, Dicle Üniv. Diş Hekim. Fak. Derg. 13 (2022) 645–652.
[13] Y. Daşdemir, Cognitive investigation on the effect of augmented reality-based reading on emotion classification performance: a new dataset, Biomed. Signal Process. Control 78 (2022) 103942.
[14] Y. Daşdemir, Impact of artificial and physical locomotion techniques on cybersickness, usability, immersion, in: 2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), IEEE, 2023, pp. 1–5.
[15] Y. Daşdemir, R. Özakar, Affective states classification performance of audio-visual stimuli from eeg signals with multiple-instance learning, Turk. J. Electr. Eng. Comput. Sci. 30 (2022) 2707–2724.
[16] A. Delorme, S. Makeig, EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, J. Neurosci. Methods 134 (2004) 9–21, https://doi.org/10.1016/j.jneumeth.2003.10.009.
[17] M. Ding, S.L. Bressler, W. Yang, H. Liang, Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data preprocessing, model validation, and variability assessment, Biol. Cybern. 83 (2000) 35–45, https://doi.org/10.1007/s004229900137.
[18] Y. El Basbasse, J. Packheiser, J. Peterburs, C. Maymon, O. Güntürkün, G. Grimshaw, S. Ocklenburg, Walk the plank! Using mobile electroencephalography to investigate emotional lateralization of immersive fear in virtual reality, R. Soc. Open Sci. 10 (2023) 221239.
[19] A. Erdem, E. Erdem, Multiple-instance learning with instance selection via dominant sets, in: International Workshop on Similarity-Based Pattern Recognition, Springer, 2011, pp. 177–191.
[20] L. Farokhah, R. Sarno, C. Fatichah, Simplified 2d cnn architecture with channel selection for emotion recognition using eeg spectrogram, IEEE Access (2023).
[21] A. Felnhofer, O.D. Kothgassner, M. Schmidt, A.K. Heinzle, L. Beutl, H. Hlavacs, I. Kryspin-Exner, Is virtual reality emotionally arousing? Investigating five emotion inducing virtual park scenarios, Int. J. Hum.-Comput. Stud. 82 (2015) 48–56.
[22] E. Frank, B. Pfahringer, Propositionalisation of multi-instance data using random forests, in: Australasian Joint Conference on Artificial Intelligence, Springer, 2013, pp. 362–373.
[23] F. Galvão, S.M. Alarcão, M.J. Fonseca, Predicting exact valence and arousal values from eeg, Sensors 21 (2021) 3414.
[24] S. Gannouni, K. Belwafi, A. Aledaily, H. Aboalsamh, A. Belghith, Software usability testing using eeg-based emotion detection and deep learning, Sensors 23 (2023) 5147.
[25] G. Ghorbanzadeh, Z. Nabizadeh, N. Karimi, P. Khadivi, A. Emami, S. Samavi, Dgaff: deep genetic algorithm fitness formation for eeg bio-signal channel selection, Biomed. Signal Process. Control 79 (2023) 104119.
[26] V. Gonuguntla, Y. Wang, K.C. Veluvolu, Phase synchrony in subject-specific reactive band of eeg for classification of motor imagery tasks, in: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2013, pp. 2784–2787.
[27] J. Han, Y. Zhao, H. Sun, J. Chen, A. Ke, G. Xu, H. Zhang, J. Zhou, C. Wang, A fast, open eeg classification framework based on feature compression and channel ranking, Front. Neurosci. 12 (2018) 217.
[28] F. Hlawatsch, F. Auger, Time-Frequency Analysis, John Wiley & Sons, 2013.
[29] S.M. Hofmann, F. Klotzsche, A. Mariola, V. Nikulin, A. Villringer, M. Gaebler, Decoding subjective emotional arousal from eeg during an immersive virtual reality experience, eLife 10 (2021) e64812.
[30] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, H.H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proc. R. Soc., Math. Phys. Eng. Sci. 454 (1998) 903–995.
[31] A. Hyvärinen, E. Oja, A fast fixed-point algorithm for independent component analysis, Neural Comput. 9 (1997) 1483–1492, https://doi.org/10.1162/neco.1997.9.7.1483.
[32] T. Iwaki, M. Noshiro, Eeg activity over frontal regions during positive and negative emotional experience, in: 2012 ICME International Conference on Complex Medical Engineering (CME), IEEE, 2012, pp. 418–422.
[33] R. Jenke, A. Peer, M. Buss, Feature extraction and selection for emotion recognition from eeg, IEEE Trans. Affect. Comput. 5 (2014) 327–339, https://doi.org/10.1109/TAFFC.2014.2339834.
[34] J. Jin, Y. Miao, I. Daly, C. Zuo, D. Hu, A. Cichocki, Correlation-based channel selection and regularized feature optimization for mi-based bci, Neural Netw. 118 (2019) 262–270.
[35] L. Jin, E.Y. Kim, Interpretable cross-subject eeg-based emotion recognition using channel-wise features, Sensors 20 (2020) 6719.
[36] V.M. Joshi, R.B. Ghongade, Idea: intellect database for emotion analysis using eeg signal, J. King Saud Univ, Comput. Inf. Sci. 34 (2022) 4433–4447.
[37] D. Kamińska, K. Smółka, G. Zwoliński, Detection of mental stress through eeg signal in virtual reality environment, Electronics 10 (2021) 2840.
[38] J.S. Kang, U. Park, V. Gonuguntla, K. Veluvolu, M. Lee, Human implicit intent recognition based on the phase synchrony of eeg signals, Pattern Recognit. Lett. 66 (2015) 144–152.
[39] K.S. Kassam, A.R. Markey, V.L. Cherkassky, G. Loewenstein, M.A. Just, Identifying emotions on the basis of neural activation, PLoS ONE 8 (2013) e66032.
[40] S. Katsigiannis, N. Ramzan, Dreamer: a database for emotion recognition through eeg and ecg signals from wireless low-cost off-the-shelf devices, IEEE J. Biomed. Health Inform. 22 (2017) 98–107.
[41] R.S. Kennedy, N.E. Lane, K.S. Berbaum, M.G. Lilienthal, Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness, Int. J. Aviat. Psychol. 3 (1993) 203–220.
[42] S. Khare, A. Nishad, A. Upadhyay, V. Bajaj, Classification of emotions from eeg signals using time-order representation based on the s-transform and convolutional neural network, Electron. Lett. 56 (2020) 1359–1361.
[43] S.K. Khare, V. Bajaj, An evolutionary optimized variational mode decomposition for emotion recognition, IEEE Sens. J. 21 (2020) 2035–2042.
[44] S.K. Khare, V. Bajaj, Time–frequency representation and convolutional neural network-based emotion recognition, IEEE Trans. Neural Netw. Learn. Syst. 32 (2020) 2901–2909.
[45] S.K. Khare, V. Blanes-Vidal, E.S. Nadimi, U.R. Acharya, Emotion recognition and artificial intelligence: a systematic review (2014–2023) and research recommendations, Inf. Fusion 102019 (2023).
[46] 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 (2011) 18–31.
[47] A. Konar, A. Chakraborty, Emotion Recognition: A Pattern Analysis Approach, John Wiley & Sons, 2015.
[48] J.P. Lachaux, E. Rodriguez, J. Martinerie, F.J. Varela, Measuring phase synchrony in brain signals, Hum. Brain Mapp. 8 (1999) 194–208.
[49] M. Le Van Quyen, J. Foucher, J. Lachaux, E. Rodriguez, a. Lutz, J. Martinerie, F.J. Varela, Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony, J. Neurosci. Methods 111 (2001) 83–98, https://doi.org/10.1016/S0165-0270(01)00372-7.
[50] Y.Y. Lee, S. Hsieh, Classifying different emotional states by means of eeg-based functional connectivity patterns, PLoS ONE 9 (2014) e95415.
[51] G. Li, M. McGill, S. Brewster, C.P. Chen, J.A. Anguera, A. Gazzaley, F. Pollick, Multimodal biosensing for vestibular network-based cybersickness detection, IEEE J. Biomed. Health Inform. 26 (2021) 2469–2480.
[52] J. Li, R. Wang, M. Zhou, Application of brain-computer interfaces in virtual reality environments: a comprehensive review, Front. Neurosci. 17 (2023) 873224, https://doi.org/10.3389/fnins.2023.873224.
[53] K.A. Lindquist, T.D. Wager, H. Kober, E. Bliss-Moreau, L.F. Barrett, The brain basis of emotion: a meta-analytic review, Behav. Brain Sci. 35 (2012) 121–143.
[54] Y. Luo, W. Mu, L. Wang, J. Wang, P. Wang, Z. Gan, L. Zhang, X. Kang, An eeg channel selection method for motor imagery based on Fisher score and local optimization, J. Neural Eng. 21 (2024) 036030.
[55] 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 (2021) 5135.

[56] N.D. Mai, H.T. Nguyen, W.Y. Chung, Real-time on-chip machine-learning-based wearable behind-the-ear electroencephalogram device for emotion recognition, IEEE Access (2023).
[57] S. Mallat, A Wavelet Tour of Signal Processing, 2nd edition ed., Academic Press, San Diego, CA, USA, 1999.
[58] J. Marín-Morales, J. Higuera-Trujillo, A. Greco, et al., Affective computing in virtual reality: emotion recognition from brain and heartbeat dynamics using wearable sensors, Sci. Rep. 8 (2018) 13657.
[59] J. McAuley, C. Marsden, Physiological and pathological tremors and rhythmic central motor control, Brain 123 (2000) 1545–1567.
[60] F. Mormann, K. Lehnertz, P. David, C.E. Elger, Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients, Phys. D, Nonlinear Phenom. 144 (2000) 358–369, https://doi.org/10.1016/S0167-2789(00)00087-7.
[61] M. Naeem, C. Brunner, G. Pfurtscheller, Dimensionality reduction and channel selection of motor imagery electroencephalographic data, Comput. Intell. Neurosci. 2009 (2009) 537504.
[62] T. Nguyen, H. Kim, J. Yoon, Development of a virtual reality-based brain-computer interface for real-time cognitive state monitoring, Sensors 20 (2020) 4587, https://doi.org/10.3390/s20164587.
[63] P.L. Nunez, R. Srinivasan, a.F. Westdorp, R.S. Wijesinghe, D.M. Tucker, R.B. Silberstein, P.J. Cadusch, EEG coherency. I: statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales, Electroencephalogr. Clin. Neurophysiol. 103 (1997) 499–515, https://doi.org/10.1016/S0013-4694(97)00066-7.
[64] C.O. Okreghe, M. Zamani, A. Demosthenous, A deep neural network-based spike sorting with improved channel selection and artefact removal, IEEE Access 11 (2023) 15131–15143.
[65] J. Onton, S. Makeig, Information-based modeling of event-related brain dynamics, Progress in Brain Research 159 (2006) 99–120.
[66] J. Pan, R. Liang, Z. He, J. Li, Y. Liang, X. Zhou, Y. He, Y. Li, St-scgnn: a spatio-temporal self-constructing graph neural network for cross-subject eeg-based emotion recognition and consciousness detection, IEEE J. Biomed. Health Inform. (2023).
[67] J. Pan, L. Zhan, C. Hu, J. Yang, C. Wang, L. Gu, S. Zhong, Y. Huang, Q. Wu, X. Xie, et al., Emotion regulation and complex brain networks: association between expressive suppression and efficiency in the fronto-parietal network and default-mode network, Front. Human Neurosci. 12 (2018) 70.
[68] Y. Park, W. Chung, Optimal channel selection using correlation coefficient for csp based eeg classification, IEEE Access 8 (2020) 111514–111521.
[69] G. Pfurtscheller, F.L. Da Silva, Event-related eeg/meg synchronization and desynchronization: basic principles, Clin. Neurophysiol. 110 (1999) 1842–1857.
[70] K.M.T. Pöhlmann, G. Li, M. Mcgill, R. Markoff, S.A. Brewster, You spin me right round, baby, right round: examining the impact of multi-sensory self-motion cues on motion sickness during a vr reading task, in: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, 2023, pp. 1–16.
[71] J. Ramirez, L. Martinez, A. Smith, Virtual reality-based brain-computer interfaces: a review of recent advances, IEEE Access 9 (2021) 30576–30587, https://doi.org/10.1109/ACCESS.2021.3055878.
[72] L. Romeo, A. Cavallo, L. Pepa, N. Bianchi-Berthouze, M. Pontil, Multiple instance learning for emotion recognition using physiological signals, IEEE Trans. Affect. Comput. 13 (2019) 389–407.
[73] H. Shahabi, S. Moghimi, Toward automatic detection of brain responses to emotional music through analysis of eeg effective connectivity, Comput. Hum. Behav. 58 (2016) 231–239.
[74] C. Stam, B. van Dijk, Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets, Phys. D, Nonlinear Phenom. 163 (2002) 236–251, https://doi.org/10.1016/S0167-2789(01)00386-4.
[75] P. Tass, M. Rosenblum, J. Weule, J. Kurths, A. Pikovsky, J. Volkmann, A. Schnitzler, H.J. Freund, Detection of n: m phase locking from noisy data: application to magnetoencephalography, Phys. Rev. Lett. 81 (1998) 3291.
[76] C. Tremmel, C. Herff, D.J. Krusienski, Eeg movement artifact suppression in interactive virtual reality, in: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2019, pp. 4576–4579.
[77] L. Trujillo, D.E. Hernandez, A. Rodriguez, O. Monroy, O. Villanueva, Effects of feature reduction on emotion recognition using eeg signals and machine learning, Expert Syst. 41 (8) (2024) e13577.
[78] H. Uyanık, S.T.A. Ozcelik, Z.B. Duranay, A. Sengur, U.R. Acharya, Use of differential entropy for automated emotion recognition in a virtual reality environment with eeg signals, Diagnostics 12 (2022) 2508.
[79] W.E. Wang, R.L. Ho, B. Gatto, S.M. Van Der Veen, M.K. Underation, J.S. Thomas, A.B. Antony, S.A. Coombes, A novel method to understand neural oscillations during full-body reaching: a combined eeg and 3d virtual reality study, IEEE Trans. Neural Syst. Rehabil. Eng. 28 (2020) 3074–3082.
[80] Y. Wang, K.C. Veluvolu, J.H. Cho, M. Defoort, Adaptive estimation of eeg for subject-specific reactive band identification and improved erd detection, Neurosci. Lett. 528 (2012) 137–142.
[81] Z. Wang, F. Wang, C. Liang, J. Zhang, A time-varying method for brain effective connectivity analysis of emotional eeg data, in: 2021 IEEE 6th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA), IEEE, 2021, pp. 131–139.
[82] Z.M. Wang, S.Y. Hu, H. Song, Channel selection method for eeg emotion recognition using normalized mutual information, IEEE Access 7 (2019) 143303–143311.
[83] N. Weidmann, E. Frank, B. Pfahringer, A two-level learning method for generalized multi-instance problems, in: European Conference on Machine Learning, Springer, 2003, pp. 468–479.
[84] I. Winkler, S. Haufe, M. Tangermann, Automatic classification of artifactual ICA-components for artifact removal in EEG signals, Behav. Brain Funct. 7 (2011) 30, https://doi.org/10.1186/1744-9081-7-30.
[85] E. Yildirim, Y. Kaya, F. Kiliç, A channel selection method for emotion recognition from eeg based on swarm-intelligence algorithms, IEEE Access 9 (2021) 109889–109902.
[86] Y. Zhang, H. Zhang, H. Wei, J. Tang, S. Zhao, Multiple-instance learning with instance selection via constructive covering algorithm, Tsinghua Sci. Technol. 19 (2014) 285–292.
[87] Z. Zhang, S.h. Zhong, Y. Liu, Torcheegemo: a deep learning toolbox towards eeg-based emotion recognition, Expert Syst. Appl. 249 (2024) 123550.
[88] 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 (2015) 162–175.