

Article

Electroencephalogram Based Emotion Recognition Using Hybrid Intelligent Method and Discrete Wavelet Transform

Duy Nguyen ¹, Minh Tuan Nguyen ²  and Kou Yamada ^{1,*} ¹ Graduate School of Science and Technology, Gunma University, Kiryu 376-8515, Gunma, Japan; t232b604@gunma-u.ac.jp² Posts and Telecommunications Institute of Technology, Hanoi 100000, Vietnam; nmtuan@ptit.edu.vn

* Correspondence: yamada@gunma-u.ac.jp

Abstract: Electroencephalography-based emotion recognition is essential for brain-computer interface combined with artificial intelligence. This paper proposes a novel algorithm for human emotion detection using a hybrid paradigm of convolutional neural networks and a boosting model. The proposed algorithm employs two subsets of 18 and 14 features extracted from four sub-bands using discrete wavelet transform. These features are identified as the optimal subsets of the most relevant, among 42 original input features extracted from two subsets of 8 and 6 productive channels using a dual genetic algorithm combined with a wise-subject 5-fold cross validation procedure in which the first and second genetic algorithms address the efficient channels and optimal feature subsets. The feature subsets are estimated by differently intelligent models and wise-subject 5-fold cross validation procedure on the validation set. The proposed algorithm produces an accuracy of 70.43%/76.05%, precision of 69.88%/74.57%, recall of 98.70%/99.17%, and F1 score of 81.83%/85.13% for valence/arousal classifications, which suggest that the frontal and left regions of the cortex associate especially to human emotions.

Keywords: electroencephalogram; bio-signal processing; genetic algorithm; wavelet transform; deep learning; machine learning



Academic Editor: Andrea Prati

Received: 24 December 2024

Revised: 13 February 2025

Accepted: 19 February 2025

Published: 21 February 2025

Citation: Nguyen, D.; Nguyen, M.T.; Yamada, K. Electroencephalogram Based Emotion Recognition Using Hybrid Intelligent Method and Discrete Wavelet Transform. *Appl. Sci.* **2025**, *15*, 2328. <https://doi.org/10.3390/app15052328>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, affective computing has attracted significant interest among researchers from various fields, which requires human knowledge from multiple disciplines, including psychology, biology, and computer science [1,2]. A major challenge in these areas is emotion recognition, which focuses on enabling computer systems to process, recognize, and understand human emotional expressions accurately. Moreover, emotions play a key role in the fields of brain-computer interaction and artificial intelligence. Advances in emotion recognition technology can propel the progress of numerous disciplines, such as computer science, robotics, psychology, neuroscience, medicine, education, entertainment, and criminal investigation [3].

Human emotion detection can be conducted using either non-physiological or physiological signals' data [4]. Non-physiological signals, such as facial expressions [5], are unreliable due to inaccurate emotions caused by illness or artificial control leading to unclear or incorrect signal acquisition. In contrast, physiological electroencephalograph (EEG) signals are generated naturally by the mental states and endocrine systems in response to emotions, which are resistant to human manipulation. EEG uses scalp electrodes to directly record brain signals, which are nonlinear, non-stationary, and noisy [6]. Indeed, EEG signals are considered a more effective yet challenging method for emotion classification due to

characteristics of reliability, high temporal resolution, cost-effectiveness, and sensitivity to emotional changes. Obviously, the EEG are internal signals generated naturally from the brain without human interferences. Then, identification of the best method for emotion recognition is the big challenge for the biomedical experts. Moreover, EEG data collection presents various difficulties essentially due to specialized knowledge and high quality capture methods. It is clear that additional noise is existed in the collected EEG signals, which definitely reduces the signal quality. Because of the EEG complexity, the noise removal is researched continuously. Another challenge is the input feature dimension reduction while maintaining relatively high performance related to emotion detection. Furthermore, the selection of affective classifier is also an important challenge. Indeed, a small but effective input features and the optimal classifier produce better emotion detection performance and are easily applied in the practical environment [7].

Emotion recognition using the EEG signals employs discrete and dimensional models in which the former sorts emotions into positive and negative categories [8], while the latter adopts the 2-dimensional valence-arousal Russell's circumplex model, which delimits emotions on a scale of varying emotional intensity ascending from left to right and bottom to top on the coordinate axis system [9]. Additionally, different human states are linked closely to brain activities in various regions and frequency bands. Therefore, electrodes are positioned at various locations on the scalp to capture EEG signals from the corresponding brain regions [2]. Furthermore, EEG signals originate from the electrical wave frequencies of the human brain, which range from 1 to 100 Hz corresponding to 5 distinct sub-bands of delta, theta, alpha, beta, and gamma [10]. In addition, the human brain is constructed from the cerebrum, cerebellum, and brain stem, in which the cerebrum is divided into the frontal, parietal, temporal, central, and occipital lobes. The individual lobes provide unique EEG signals, which are crucial for emotion recognition. As a result, the electrode order is labeled as low, medium, and high corresponding to the first 32, from 33 to 128, and higher than 128 electrodes in resolution, respectively [11].

The rapid advancements in Machine Learning (ML) and Deep Learning (DL) have promoted a large number of extensive studies on the effective designs of emotion detection applied in practical environments. Obviously, channel and feature selection plays a crucial role in the optimization of the intelligent model. However, existing works have not considered properly the channel selection for the model optimization. Indeed, the entire channels are commonly used for feature extraction in most of the existing works [2,4,5,10]. Furthermore, a small amount of the original features is considered for the feature extraction from various channels and signal sub-bands, which results in a huge feature space used as the input of different models. It is noteworthy that the identification of the most relevant feature subset extracted from different sub-bands is an especially time-consuming procedure. Additionally, signal decomposition plays a crucial role in EEG analysis. The study in [12] presents the Empirical Wavelet Transform (EWT) as a potential method. This method decomposes signals into their frequency components, enabling detailed analysis by constructing a customized wavelet basis. Another approach, as demonstrated in [13], highlights the effectiveness of decomposing signals using Discrete Wavelet Transform (DWT). DWT effectively decomposes signals into sub-bands, including theta, delta, beta, and gamma, which are valuable for extracting key features in EEG analysis applications. Indeed, DWT is a crucial method for capturing both slow-varying low-frequency components and rapidly changing high-frequency components, making it well-suited for analyzing discrete EEG signals.

Motivated by the above analysis, we propose an emotion detection algorithm using Machine Learning (ML) and Deep Learning (DL) techniques. The proposed method uses Discrete Wavelet Transform (DWT) to decompose the preprocessed EEG signals into

different sub-bands of frequencies, which are then adopted for the extraction of numerous original features. Moreover, the Genetic Algorithm (GA)-based exhaustive search method in combination with the wise-subject cross-validation (CV) and an ML model as the fitness function is implemented to address the optimal channels and feature subsets. Finally, the proposed algorithm is validated for its detection performance by the wise-subject CV method on a separate dataset to improve the reliability of the applications in practical environments. The main contributions of this work are as follows:

- Investigation of expanded input features extracted from various sub-bands constructed from preprocessed EEG signals, which is effective in addressing the most relevant feature subsets in terms of detection performance improvement.
- Quality improvement of the final feature subsets by the application of dual GA procedures, in which the first GA is used to select the most productive channels using the total input features, and the second GA identifies the most informative features extracted from the selected channels.
- Proposal of a reliable algorithm for emotion detection, which is based on performance result comparisons of various models using a wise-subject CV-based statistical method.

This paper is organized as follows: Section 2 presents a literature review. Section 3 describes the materials and methods used in the research. Section 4 provides the simulation results, followed by the discussion in Section 5. Finally, Section 6 summarizes the key findings and concludes the paper.

2. Literature Review

The effectiveness of different ML models in [4,14–23] has been proven in the applications of emotion detection using EEG signals. Obviously, the utility of ML models requires feature extraction and selection using EEG signals to identify the most informative feature subsets for the improvement of the final detection performance. Indeed, the Support Vector Machine (SVM) is used as the main model for the method development in [14] using features extracted from EEG signals in time, frequency, and time-frequency domains. Here, the most relevant feature subset is selected by a modified particle swarm optimization method with multistage linearly-decreasing inertia weight for the input of the SVM model, which contributes to the significant improvement of multiple emotion recognition. In [15], 2 models, namely K-nearest neighbors (KNN) and Artificial Neural Networks (ANN) use the entropy and energy features extracted from the sub-bands, which are generated by the Fourier-Bessel series expansion-based empirical wavelet transform. The preselection of 10 channels from the frontal lobes in combination with 3 feature selection algorithms such as neighborhood component analysis, ReliefF, and minimum-Redundancy-Maximum-Relevance (mRMR) results in the input feature dimension reduction while the final detection performance of the proposed algorithm is improved certainly.

A GA based evolution algorithm is applied for the feature selection, which releases the most relevant feature subsets used as the input of KNN, Random Forest (RF), ANN models [19] and SVM [20] for the estimation of their performance in terms of emotion detection. In [19], the effectiveness of the GA has been proven clearly with a small percentage of the selected features, approximately 2% out of the total, corresponding to the model, which shows better detection performance than that using the entire input features. The authors of [20] consider 3 original feature sets for the GA-based feature selection, which leads to the feature proportions of 10%, 12%, and 25% chosen from the 3 above feature sets. In [21], the performance investigation of the variational mode decomposition and the DWT is indicated by the effectiveness comparisons of sub-signals in different frequency bands for feature extraction. Additionally, the non-dominated sorting genetic algorithm (NSGA-II) is implemented for the identification of the most effective channels and features in com-

bination with different ensemble learning algorithms. As a result, a subset of 8 features extracted from 7 channels, which are selected by the variational mode decomposition-based feature extraction combined with the NSGA-II, and a SVM model are proposed as the final emotion recognition algorithm.

Furthermore, the principal component analysis is deployed for the dimension reduction of the feature matrix, which is constructed by the power spectral density of the individual sub-bands. In [22], the performance superiority is confirmed definitely for the quantum-based SVM in comparison with the conventional SVM models. The authors of [4] also employ the principal component analysis for the dimension reduction of the input feature space including sub-band power, energy, and others in the time domain. Here, the wavelet-based atomic function and ML models are considered for the signal decomposition and classification. To improve the detection performance and generalization ability of the ensemble learning models, the overlapping 4s EEG segments are used for the extraction of numerous features, which are then reduced by the SVM model using L1-norm penalty-based feature selection. In [23], the detection performance of various ensemble learning models is compared for the proposal of the MOSNK model as the final algorithm using all of the extracted features.

There is no requirement of feature extraction, and selection using human expert knowledge or conventional expertise for the application of the DL models. Moreover, representative features are reinforced by the capability of the deep learning process, which makes the DL models outstanding in comparison with the ML techniques. Indeed, the authors of [24] use 2-dimension power topographic maps constructed by the power spectral density features of 5 sub-bands, which are then used as the input of the spatial-temporal information learning network including spatial feature extraction based-convolutional neural network (CNN) and temporal context learning based-long short term memory (LSTM). The power topographic maps effectively represent the most vital brain areas while the temporal context learning based-LSTM captures significant dependencies of EEG frames. The study in [25] introduces a hybrid unsupervised deep convolutional recurrent generative adversarial network designed for EEG feature characterization and fusion namely EEGFuseNet. This network plays a role in automatic deep feature extraction, which captures spatial and temporal dynamics, making them more generic and independent of specific EEG tasks. In [26], the adjacency matrices calculated by the phase locking value (PLV) method are adopted for the interaction representative of sub-bands generated from 20s EEG segments. Then, the CNN model is fed with the above matrices as the input for further classification. Clearly, the final detection performance of the proposed algorithm is improved significantly with the use of group PLV compared with individual PLV. Besides, the authors in [27] propose a novel deep learning framework for EEG-based emotion recognition, which includes a spatio-temporal representation of multichannel EEG signals namely TQWT-feature block sequences and a hybrid convolutional recurrent neural network. Here, a lightweight CNN and a LSTM network capture spatial information and temporal dependencies, respectively, are the 2 main elements of the hybrid convolutional recurrent neural network, which proves the essential ability of emotion recognition. Various ensemble learning and DL models are investigated in [28] using a large number of features extracted from 5 sub-bands. Superior detection performance is released for a hybrid paradigm in comparison with that produced by ensemble learning or individual DL models. In [29], the proposed method includes a graph convolutional network using differential entropy-based 3-dimension spatial-spectral features extracted from various sub-bands by the Welch method. Here, the adjacency matrix calculation is based on the above features and contextual loss computed by trainable adjacency relation in combination with the graph convolutional network model. The detection comparisons between the ML and bidirectional LSTM are provided in [30] using 4 features

extracted in time, frequency, and entropy domains. The structure of the bidirectional LSTM is optimized by the investigation of hyperparameters, which produce better performance compared with that released by the ML models.

To improve the emotion detection performance of the proposed algorithm, feature and channel selection are considered for the optimization of the extracted features as the input of DL models. In [2,17], adjustable channel selection based on attention distribution in the graph structure and 10 frontal lobe channels are given for the construction of scalogram images, and intra- and inter-channel EEG features. Moreover, the graph convolution models [2] and the deep CNN [17] are proposed as the emotion classifiers in which the classification performance of the former is slightly higher than that of the latter. The scalogram images are also used in [31] generated by the wavelet transform to convert EEG signals from time-scale representation into these images using different sub-bands. Then, the preselection model namely GoogLeNet is applied for emotion recognition using the scalogram images, which produces a relatively high performance for the multiple emotion classification. In [32], images are generated by the synchrosqueezing wavelet transform, which are then used as the input of a pretrained model namely ResNet-18 for binary emotion classification. Similarly, the authors of [33] consider a fused 2-dimension image combined from the connectivity measures and different overlapping time windows, which represent effectively brain connectivity. Then, a hybrid DL technique of CNN and LSTM is adopted for the recognition of four emotions using the above images. Moreover, the CWT is considered in [34,35] for the construction of the scalogram images, which improve the representation of the input EEG signals. Then, the final emotion recognition is decided by the majority voting algorithm [34] using the output of different transfer learning models and by the multiclass SVM model [35] using the best feature layers of various transfer learning models.

In [36], the binary gray wolf optimizer is implemented for the selection of the most relevant input features from statistic, wavelet, and Hurst exponent features. Additionally, the differential evolution algorithm is also employed to identify the optimal hyperparameters for the bi-directional LSTM model structure proposed as the final algorithm. Channel optimization is also considered in [37], where NSGA-II is implemented to balance the two-objective problem in terms of detection performance and channel number. Besides, the CNN is applied for the proposed algorithm namely EEGNet, which uses the feature extracted from the sub-bands using the DWT. The authors of [38] introduce binary many-objective particle swarm optimization with cooperative agents in combination with a ConvLSTM model as the channel selection method. Here, the proposed model is based on the autoencoder, which shows better performance in capturing the spatio-temporal information of the preprocessed EEG signals.

From previous studies, two key research gaps have been identified: (1) the lack of proper channel selection for model optimization and (2) a small number of extracted features used as input. Many studies have not adequately considered channel selection, which may compromise the efficiency of the EEG-based emotion recognition system. In addition, a small number of extracted features can limit the model's ability to capture comprehensive patterns in EEG signals. To address these limitations, this study implements an optimized channel selection process using GA to identify the most relevant EEG channels, thereby improving signal quality and model efficiency. Furthermore, we expand the extracted feature set to 42 features across time-frequency, entropy, and complexity domains. Increasing the number of extracted features is essential because EEG signals exhibit intricate variations that require a richer feature set for comprehensive pattern recognition.

3. Materials and Methods

In our research, to enhance the performance of models used in EEG-based emotion recognition, we employ a variety of models including ML, DL, and hybrid models. The optimization algorithms to fine-tune the models and identify the most effective parameter values are also implemented in terms of emotion detection performance improvement. Moreover, the DWT is adopted for the decomposition of the EEG signals, which are generated from various brain areas, into multiple sub-bands for the quality improvement of the extracted feature related to the human emotions in time-frequency, entropy, and complexity domains. By combining the advanced model optimization with the productive feature extraction, the research purpose is to significantly enhance both the accuracy and robustness of emotion recognition using the EEG signals.

The proposed method consists of 4 phases namely data preprocessing, channel and feature selection, model selection, and model validation as shown in Figure 1. In the first phase, EEG signals are preprocessed by the bandpass filter and downsampling method. Moreover, 4 bands namely theta, alpha, beta, and gamma are constructed from the pre-processed EEG signals by the utility of DWT. The total input features (TIF) are extracted from various signal bands in time, entropy, and complexity domains in the second phase. The optimal channels (OCH) are then selected by the GA using the TIF and the wise-subject 5-fold CV procedure. Moreover, the most informative features, as selected combined features (SCFs), are also addressed from the optimal channels (OCH) using the above GA method. In the third phase, a grid-search-based method is implemented for the selection of the optimal ML and DL models using the TIF extracted from the OCH and the wise-subject 5-fold CV method on the training set. In the last phase, the selected ML and DL models are validated for their detection performance on the validation set using the wise-subject 5-fold CV procedure. It is noteworthy that the 5-fold CV procedure is selected to ensure 20% of the total subjects used for the validation set.

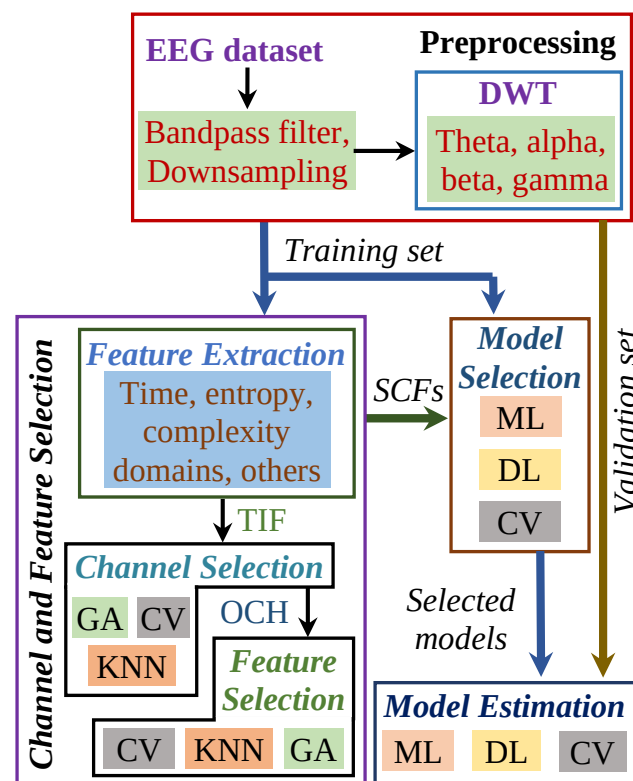


Figure 1. Method diagram.

3.1. Data Descriptions

Materials for emotion analysis used in this work are the database of the Physiological Signals (DEAP) [39], which is a collection of EEG signals from 32 subjects, ranging from 19 to 37 years old with equal numbers of males and females. A total of 40 videos with a length of 63 s are watched and listened to by the subjects to stimulate various emotions, which are measured by valence, arousal, liking, and dominance scales based on participant self-report. EEG signals are recorded with 40 AgCl wet electrodes during the period of video time with a sample frequency of 512 Hz.

3.2. Materials Preprocessing

The dataset is preprocessed with different techniques for further feature extraction as follows:

(a) Preprocessing

We use the DEAP dataset for this work, which was preprocessed by [39] as follows:

1. The data was downsampled to 128 Hz.
2. EOG artefacts were removed.
3. A bandpass frequency filter from 4–45 Hz was applied.
4. The data was averaged to the common reference.
5. The EEG channels were reordered so that they all follow the Geneva order as above.
6. The data was segmented into 60-s trials and a 3-s pre-trial baseline removed.

After preprocessed by [39], the individual EEG signals have a length of 63 s including the first 3 s of the pretrial baseline, which are then removed to correct the stimulus changes of EEG signals. As a result, 32 (subjects) $\times$ 40 (videos) $\times$ 32 (channels) $\times$ 7860 (data points) as considered as the preprocessed EEG signals using for our work. It is noteworthy that we do not implement a filter to denoise EOG and EMG because this technique has already completed by [39].

(b) Labelling

Two dimensions of emotions are selected namely valence and arousal for further classifications in the DEAP dataset. Each dimension ranges from 1 to 9 in which a threshold of 5 is employed. Binary classifications are for two scenarios using the above threshold, which are high/low valence and high/low arousal. As a result, four emotions namely Joy, Calm, Disgust, and Sadness are identified by High Valence High Arousal (HVHA), High Valence Low Arousal (HVLA), High Arousal Low Valence (HALV), and Low Valence Low Arousal (LVLA), respectively.

(c) Signal band construction

The DWT is an effective method known as a nonredundant sampled continuous wavelet transform (CWT) technique [40], which generates a set of wavelet coefficients of the discrete time series. Then, an orthogonal set of basic function is constructed by the use of the above wavelet coefficients, which ensures the nonredundant representation. By breaking down a signal into components at different frequencies and scales, DWT is particularly well-fitted in application for the non-stationary EEG signals [13], which makes DWT as a widely applied method for time-frequency analysis in the fields of neuroscience and biomedical engineering. The DWT of $S(t)$ is defined as [41]:

$$W_{\varnothing}(i_0, k) = \frac{1}{\sqrt{N}} \sum_t S(t) \varnothing_{i_0, k}(t) \quad (1)$$

and

$$W_{\psi}(i, k) = \frac{1}{\sqrt{N}} \sum_k S(t) \psi_{i, k}(t), \quad (2)$$

where, $\varnothing(t)$ and $\psi(t)$ are wavelet function and scaling function, respectively. N is power of 2. i is the scale parameter with $i = 0, 1, 2, \dots, I - 1$, and k is the shift parameter with $k = 0, 1, 2, \dots, 2^i - 1$. $S(t)$ is reconstructed by

$$S(t) = \frac{1}{N} \sum_k W_{\varnothing}(i_0, k) \varnothing_{i_0, k}(t) + \frac{1}{\sqrt{N}} \sum_{i=i_0}^{I-1} \sum_k W_{\psi}(i, k) \psi_{i, k}(t). \quad (3)$$

There are two components constructed by the DWT namely the low-pass and high-pass filters, which produce the approximate and detail signals using a scaling and a wavelet functions as shown in (1) and (2), respectively. The signal decomposition process is implemented recursively through various levels using repeatedly the high- and low-pass filters in which the detail signal is kept in current level while the approximate signal is continuously decomposed by the above filters in following level. We use four sub-bands constructed from preprocessed EEG signals using the DWT method with the fourth Daubechies mother wavelet function (db4). Here, the preprocessed EEG signals are decomposed into different sub-bands corresponding to frequency ranges such as theta of [4–8] Hz, alpha of [8–16] Hz, beta of [16–32] Hz, and gamma of [32–45] Hz. Figure 2 shows an example of different band signals.

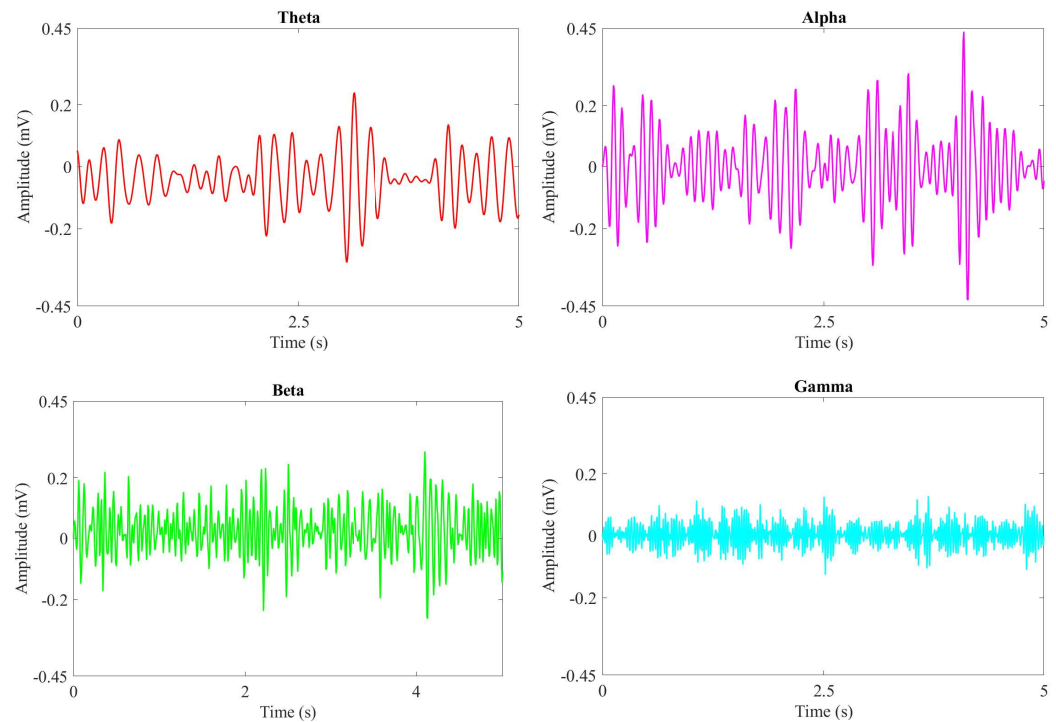


Figure 2. Examples of Theta, Alpha, Beta, and Gamma signals.

3.3. Channel and Feature Selection

The optimization method is applied for the selection of channels and features in this work as follows:

(a) Feature Extraction

A total of 42 input features are extracted from four sub-bands generated by DWT using different techniques in time-frequency, entropy, and complexity domains as shown in Table 1 panels (a), (b) and (c). All input features are extracted and used for the next step of channel selection.

Table 1. (a) The descriptions of the time-frequency domain features. (b) The descriptions of the entropy domain features. (c) The descriptions of the complexity domain features.

(a)	
Time-Frequency Domain Features	
Features Name	Description
Hjorth activity (HA) [23]	Representation of the signal power information
Hjorth complexity (HC) [23]	Reflection of the bandwidth and the change in frequency
Hjorth mobility (HM) [23]	The square root of the ratio between the variance of the signal's first derivative and the variance of the signal
Maximum of power spectral density (mPSD) [23]	The peak of spectral energy at a particular frequency
Maximum frequency (MF) [23]	The maximum frequency corresponding to mPSD
Lyapunov exponent of the uniformly sampled time-domain signal (LyaExp) [23]	Quantification of periodic behavior of chaotic systems
First difference (FD) [14]	Representation of the relationship between the current data point and its preceding one, highlighting changes in the waveform's dimensionality over time
Second difference (SD) [14]	Representation of the relationship among 3 consecutive data points and is a measure that is highly sensitive to changes in the signal's amplitude
Normalized first difference (NFD) [14]	Reflection of the change between the current data and the previous data in the normalized EEG signal
Normalized second difference (NSD) [14]	Description of the relationship between 3 sequential data points in the normalized EEG signal
Mean absolute value (MEA) [42]	Measurement of the average magnitude of the signal
Mean value of EEG amplitude (MA) [42]	The average value computed from all points in the signal
Median value of EEG (Med) [42]	The middle value in a signal's data set after arranging all values in ascending or descending order
Average power (BPA) [42]	The mean of the squared amplitudes of the signal
Energy (Ene) [42]	The sum of the squared magnitudes of all the signal's components
Slope sign change (SSC) [43]	Measurement of the number of times the slope of the waveform changes sign
Zero crossing rate (ZCR) [43]	A measure of EEG waveform frequency cross the zero axis in a unit time
(b)	
Entropy Domain Features	
Features Name	Description
Approximate entropy (AE) [23]	Quantification of fluctuation regularity and unpredictability
Permutation entropy (PE) [23]	Analysis of the relative occurrence of each of these patterns
Singular entropy (SSE) [23]	Computation of applying singular value decomposition to the trajectory, which is reconstructed a one-dimensional time series into a multidimensional phase space
Shannon entropy (ShE) [23]	Measurement of uncertainty, commonly to evaluate the degree of chaos in EEG signals

Table 1. Cont.

Dispersion entropy (DE) [44]	Quantification of the regularity of a time series
Fuzzy entropy (FE) [45]	Measurement of the degree of similarity between two vectors based on their shape
Renyi entropy (RE) [46]	Measurement of the uncertainty in a probability distribution using a parametric family of indices
Sample entropy (SE) [47]	A metric of the underlying or complexity
Maximum of Spectral entropy (Spe) [47]	Measurement of the irregularity or complexity of the signal's power distribution across frequencies
Tsallis entropy (TE) [48]	Description of the physical behavior of a system and aids in distinguishing between bursts, continuous rhythms, and spikes in EEG signals
(c)	
Complexity Domain Features	
Features Name	Description
Co-dimension (CoD) [23]	Identification of the number of independent variables required to characterize the system's dynamics
C0 complexity (C0) [23]	The proportion of stochastic components in a signal, assumed to have both regular and stochastic elements
Auto regressive model (AR) [49]	Modeling the characteristics and information present in a signal, where each sample is calculated by adding up the weighted previous samples
Higuchi (HFD) [50]	Definition of the minimum number of coordinates required to identify any point in phase space
Hurst (H) [50]	Measurement of the degree of self-similarity and predictability within a time series under this nonlinear statistical model
Root mean square of amplitude (AM) [42]	Representation of the square root of the average of the sum of the signal's squared values
Variance (Var) [42]	A measure of signal data dispersion, determined by the average of the squared differences from the mean
Skewness (Sw) [50]	Measurement of the asymmetry of a real-valued distribution function relative to its mean
Phase space reconstruction (Psr) [51]	Transformation of time series properties into topological features of a geometric object, preserving the original space's topological properties
Hilbert transform (Hilb) [51]	Representing frequency as a phase change rate
Complexity measure (CM) [52]	Measurement of the complexity and irregularity in EEG data
Covariance calculation (Cvbin) [53]	Quantification of the variance in the EEG signal and the corresponding binary signal
Frequency calculation (Frqbin) [53]	The count of binary signal transitions between '0' and '1', followed by division the result by 10 to obtain transitions for 1s
Area calculation (Abin) [53]	Representation of the maximum between the sum of binary signal values and the sum of inverted binary signal values
Kurtosis (Kurt) [50]	Measurement of the peakedness or flatness of a signal relative to its normal distribution

(b) Channel and Feature Selection

The original EEG dataset is collected over the AgCl wet electrodes and Biosemi ActiveTwo system [39], which possibly represent high correlation of the EEG signals. Hence, the final detection performance of human emotions is decreased due to the low-quality feature extracted from the above EEG signals. Moreover, few irrelevant features are certainly included in the TIF extracted from improper channels, which results in a decrease in the final performance of the emotion detection algorithm. Therefore, the GA-based channel and feature ranking methods are employed for the selection of the most informative channels and features in terms of emotion detection [20]. The chromosome includes a string of 32 binary bits representing 32 channels in which 1 and 0 bits are assigned to channel presence and absence, respectively. An exhaustive search is performed in 100 generations of GA in which each consists of 200 chromosomes, known as population, to generate various channel combinations with crossover of 0.8, and mutation of 0.02. The TIFs are then extracted from individual combinations for further resubstitution loss estimation of the KNN model-based fitness function on the training set using the wise-subject 5-fold CV. A repetition of 100 times is deployed for the GA resulting in the number of times that channels are selected. Because the individual channels represent variously specific human brain regions, which generate definitely the emotions. Hence, we remove the channels which are omitted by the GA. As a result, the channels with the numbers of chosen times over zero are collected as the OCH.

A similar GA procedure is applied for the feature selection using the above OCH. Here, a string of 42 bits is used to represent a total of 42 input features of the individual OCH. Bits 1 and 0 indicate the presence and absence of features extracted from the OCH. Consequently, the input features are identical for all channels in the OCH at each GA repetition as given in Algorithm 1. A large number of 5376 input features extracted from 4 sub-bands for which each sub-band includes 32 channels makes it difficult to apply for the practical environment. Therefore, it is necessary to set a proper number of selections by the GA to ensure a relevant number of selected features while maintaining relatively high classification performance and small processing time of the proposed algorithm. In this work, we address the SCFs as a set of individual features with the numbers of selected times by the GA, which are larger than 50 known as a half of GA repetitions.

3.4. Model Selection

A grid search-based method in combination with the wise-subject 5-fold CV procedure is implemented to address the best structures and learning parameters of differently intelligent algorithms for binary classification of valence and arousal scenarios. We use 5 intelligent models in this work, which are KNN [54], Bagging (BG), RF, Boosting (BS) [55], CNN [24].

A grid search used for hyper-parameter tuning of BG and RF includes 2 parameters with tree number of [3; 5; 10; 15; 25; 50; 75] and leaf number of [5; 10; 15; 20; 25; 30; 35; 45; 55]. The K parameter is found in a range of [3:2:99] for the KNN model. There are 3 parameters, namely learning rate (lr), iteration (iter), and minimum leaf (min_leaf), which make a grid search of [0.1; 0.5; 0.7; 0.9; 0.95], [25; 50; 100; 150], and [25; 50; 75; 95; 105] for the BS model. Hyper-parameter tuning for the CNN model contains structure and learning parameters. The former is defined by the network depths and sections. A network depth (Nd) is the number of continuous blocks in which each is constructed by convolutional, batchnormalization, and relu layers. The network section (Ns) is the number of network depths separated by the maxpooling layer. We search for optimal Nd and Ns in a range of [0–4]. Additionally, the latter including the learning rate of [0.005; 0.01], the momentum

of [0.8; 0.9], and the L2 regulation (l2) of [0.1; 0.15; 0.2] are considered to find the optimal values for the CNN model.

Algorithm 1 GA-based channel and feature selection

```

(1) Feature number  $f_n$ ,  $n = 42$ ; Channel number  $C_a$ ,  $a = 32$ .

(2) Channel selection based GA
    loop = 100; Population = 200; Chromosome = 32 of bit 0 and 1;
     $k = 1$ ;
    Repeat
      (a) Selection of  $C_b$  using chromosome of bit 1,  $b < a$ ;
      (b) Extraction of  $f_n$  from  $C_b$  channels;
      (c) Separation of training set into 5 folds  $T(i)$ ;
          for  $i = 1$  to 5
            • Training KNN model on  $T(j)$ ,  $j \neq i$ ;
            • Estimation of mean resubstitution loss on  $T(i)$ ;
          end
      (d) Calculation of mean resubstitution loss over CV process
       $k = k + 1$ ;
    Until  $k = \text{loop}$ 
    Calculation of the mean resubstitution loss over 100 repetitions;
  (3) Selection of  $C_b$  channels as the OCH
  (4) Feature selection based GA
    Chromosome = 42 of bit 0 and 1;
     $k = 1$ ;
    Repeat
      (a) Selection of  $f_m$  using chromosome of bit 1,  $m < n$ ;
      (b) Extraction of  $f_m$  from OCH;
      (c) Separation of training set into 5 folds  $V(i)$ ;
          for  $i = 1$  to 5
            • Training KNN model on  $V(j)$ ,  $j \neq i$ ;
            • Estimation of mean resubstitution loss on  $V(i)$ ;
          end
      (d) Calculation of mean resubstitution loss over CV process
       $k = k + 1$ ;
    Until  $k = \text{loop}$ 
    Calculation of the mean resubstitution loss over 100 repetitions;
  (5) Selection of  $f_m$  as the SCFs
  
```

3.5. Model Validation

Different sets of features, which are the TIF extracted from all channels, the entire features extracted from the OCH as selected combined all features (SCAF), and SCFs, are used as the input of the selected models using the wise-subject 5-fold CV-based method on the validation data for validation of the above model detection performance. Here, 5 subsets of subjects, known as folds, are divided arbitrarily from the validation set in which a subset is considered as the test data and the others are for the training data. The models are trained and validated 5 times using every subset as the test data to complete an integral process. In addition, the wise-subject 5-fold CV is repeated 50 times for the calculation of mean and standard deviation related to the selected model performance. Finally, the classification performance comparison is shown for different models using the above feature sets. The model and corresponding feature set, which produces the highest performance, are identified as the proposed algorithms for the classification of valence and arousal scenarios.

4. Simulation Results

The performance metrics, results of channel and feature selection, model selection, and model validation are provided in this section. The simulation is performed on a

computer with a 2.1 GHz Intel Core I7, a memory of 32 GB, an RTX 4070 graphic card, and Matlab 2023a.

4.1. Performance Metrics

We use four measures for the performance estimation of various models using different feature sets namely accuracy (Ac), precision (Pn), recall (Re), and F1-score ($F1$). The first measure is computed as a proportion of instances correctly identified. The proportions of relevant instances among all retrieved and relevant instances are measured by Pn and Re . $F1$ shows the harmonic mean related to Pn and Re , which implies the detection ability of relevant instances. Ac , Pn , Re and $F1$ are culculated by

$$Ac = \frac{TP + TN}{TP + TN + FP + FN}, \quad (4)$$

$$Pn = \frac{TP}{TP + FP}, \quad (5)$$

$$Re = \frac{TP}{TP + FN}, \quad (6)$$

and

$$F1 = \frac{2 \times Re \times Pn}{Re + Pn}, \quad (7)$$

respectively, where TP , FP , TN , and FN are true positive, false positive, true negative, and false negative values.

4.2. Channel and Feature Selection

There are 8 and 6 channels, which are selected by the GA using KNN as the fitness function for valence and arousal scenarios, respectively as shown in Figure 3. Firstly, the TIF extracted from the 4 signal bands using DWT is used for the OCH selection. Then, the SCFs, which include 18 and 14 features for the classification of high/low valence and arousal, are addressed by the GA using the SCAF. The names and times selected by the GA of the chosen channels and features are shown in Table 2.

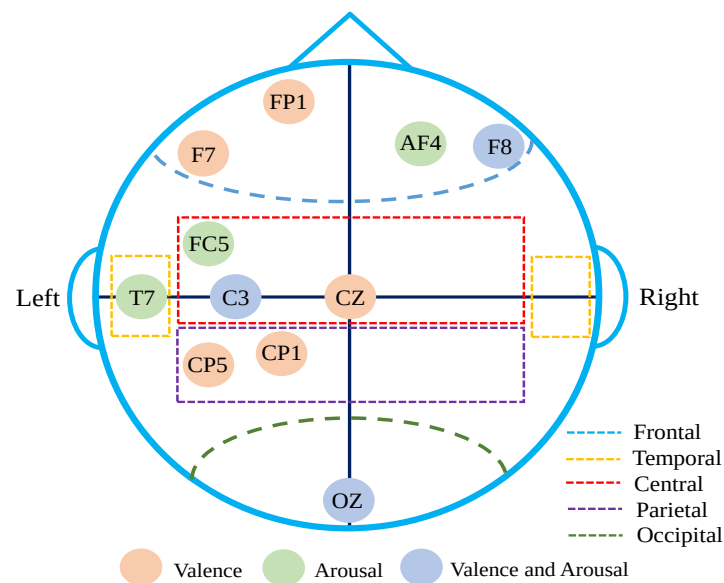


Figure 3. Selected channels for valence and arousal scenarios.

Table 2. Selected channels and features by the GA.

Channel			
Valence	Number of Selection	Arousal	Number of Selection
Fp1	52	FC5	20
F7	8	T7	94
C3	49	C3	1
CP1	1	Oz	5
CP5	42	F8	20
Oz	5	AF4	1
F8	42	Others	0
Cz	1		
Others	0		
Feature			
Valence	Number of Selection	Arousal	Number of Selection
FD	68	DE	81
NFD	70	FE	92
MEA	100	PE	91
SSC	52	RE	74
FE	68	Spe	98
PE	66	coD	96
RE	69	HFD	78
SE	100	H	100
ShE	77	AM	89
AR	54	Sw	100
H	65	Cm	57
AM	70	Cvbin	88
Var	64	Kurt	100
Cm	51	HA	92
Cvbin	83	Others	<50
HC	77		
BPA	70		
Ene	75		
Others	<50		

4.3. Model Selection

We select 2 CNN and 8 ML models on the training set using the grid search-based method for the classification of the valence and arousal scenarios. Moreover, the CNN model is also considered as the extractor to generate the deep features used as the input of the KNN model. As a result, 2 structures of the CNN extractor are selected for further estimation in combination with different ML models for the classification of the high/low

valence and arousal. Table 3 represents the optimal structure of the CNN model known as the feature extractor.

Table 3. The selected structure of the CNN extractor.

Layer	Number of Filter	Filter	Stride	Padding
Input	[576 1 1] of valence; [336 1 1] of arousal			
Convolutional (+BN,Relu)	10	[100 1]	[1 1]	[50 50 0 0]
Convolutional (+BN,Relu)	11	[100 1]	[1 1]	[50 50 0 0]
Convolutional (+BN,Relu)	11	[100 1]	[1 1]	[50 50 0 0]
Convolutional (+BN,Relu)	11	[100 1]	[1 1]	[50 50 0 0]
Maxpooling		[10 1]	[2 2]	[4 5 0 0]
Convolutional (+BN,Relu)	20	[100 1]	[1 1]	[50 50 0 0]
Convolutional (+BN,Relu)	20	[100 1]	[1 1]	[50 50 0 0]
Convolutional (+BN,Relu)	20	[100 1]	[1 1]	[50 50 0 0]
Convolutional (+BN,Relu)	20	[100 1]	[1 1]	[50 50 0 0]
Maxpooling		[10 1]	[2 2]	[4 5 0 0]
Convolutional (+BN,Relu)	40	[100 1]	[1 1]	[50 50 0 0]
Convolutional (+BN,Relu)	40	[100 1]	[1 1]	[50 50 0 0]
Convolutional (+BN,Relu)	40	[100 1]	[1 1]	[50 50 0 0]
Convolutional (+BN,Relu)	40	[100 1]	[1 1]	[50 50 0 0]
Maxpooling		[10 1]	[2 2]	[4 5 0 0]
Convolutional (+BN,Relu)	80	[100 1]	[1 1]	[50 50 0 0]
Convolutional (+BN,Relu)	80	[100 1]	[1 1]	[50 50 0 0]
Convolutional (+BN,Relu)	80	[100 1]	[1 1]	[50 50 0 0]
Convolutional (+BN,Relu)	80	[100 1]	[1 1]	[50 50 0 0]
Maxpooling		[10 1]	[2 2]	[4 5 0 0]
First Fully connected	Output sizes: 100			

BN: BatchNomalization layer.

The grid searches are implemented to address the optimal parameter values of the models as shown in Table 4. For valence scenario, KNN with K of 37; BG with tree number of 15 and leaf number of 5; BS with learning rate of 0.9, iteration of 50, and minimum leaf of 75; RF with tree number of 10 and leaf number of 5; CNN with network depth of 3, network section of 2; learning rate of 0.01, momentum of 0.9, L2 regulation of 0.1 are the optimal ML models. The optimal CNN extractor structure is shown in Table 3 with network depth of 4, network section of 4, learning rate of 0.01, momentum of 0.8, and L2 regulation of 0.2 for both valence and arousal scenarios. The optimal parameter values of the ML models in combination with the above extractor are K of 3 for the KNN; tree number of 5 and leaf number of 75 for the BG; learning rate of 0.1, iteration of 25, and minimum leaf of 75 for the BS; tree number of 3 and leaf number of 5 for the RF.

Table 4. The highest detection performance of different models with their optimal parameter values corresponding to the input feature sets.

Valence					
Model	Feature Set	Ac (%)	Pn (%)	Re (%)	F1 (%)
KNN (K = 37)	SCFs	65.35 ± 0.53	65.62 ± 0.23	98.96 ± 0.97	78.69 ± 0.50
BG tree = 15, leaf = 5	SCAF	40.65 ± 2.61	63.42 ± 7.71	29.41 ± 7.25	34.79 ± 6.74
BS lr = 0.9, iter = 50, min_leaf = 75	TIF	62.00 ± 2.40	66.86 ± 1.22	83.15 ± 4.12	73.63 ± 1.96
RF tree = 10, leaf = 5	SCFs	47.15 ± 2.75	61.27 ± 2.62	54.87 ± 6.69	55.15 ± 4.56
CNN Nd = 3, Ns = 2, lr = 0.01, mo = 0.9, l2 = 0.1	TIF	67.75 ± 6.33	67.03 ± 7.11	99.17 ± 1.14	79.83 ± 5.12
CNN + KNN K = 3	TIF	68.38 ± 0.98	68.24 ± 0.43	96.76 ± 2.32	80.03 ± 0.92
CNN + BG tree = 5, leaf = 75	TIF	58.63 ± 3.45	76.53 ± 2.51	53.13 ± 5.86	62.72 ± 4.86
CNN + BS lr = 0.1, iter = 25, min_leaf = 75	SCFs	70.43 ± 1.48	69.88 ± 0.33	98.70 ± 2.44	81.83 ± 1.23
CNN + RF tree = 3, leaf = 5	SCFs	60.61 ± 2.89	86.42 ± 5.74	48.06 ± 5.82	61.77 ± 3.82
Arousal					
KNN K = 5	TIF	63.70 ± 0.41	63.78 ± 0.25	96.66 ± 1.26	76.38 ± 0.28
BG tree = 15, leaf = 5	SCFs	55.17 ± 3.53	68.17 ± 2.17	53.72 ± 7.02	57.87 ± 5.27
BS lr = 0.9, iter = 50, min_leaf = 75	SCFs	63.62 ± 0.87	67.13 ± 1.79	79.56 ± 7.77	71.52 ± 3.80
RF tree = 10, leaf = 5	TIF	49.90 ± 3.26	65.91 ± 6.87	34.65 ± 7.83	41.81 ± 7.52
CNN Nd = 2, Ns = 3, lr = 0.005, mo = 0.9, l2 = 0.15	SCFs	65.00 ± 6.90	65.24 ± 8.37	91.46 ± 12.93	75.76 ± 8.85
CNN + KNN K = 53	TIF	70.48 ± 1.89	70.39 ± 0.82	94.77 ± 4.13	80.79 ± 1.67
CNN + BG tree = 3, leaf = 5	SCFs	46.27 ± 3.06	63.26 ± 3.63	35.91 ± 5.46	45.12 ± 5.05
CNN + BS lr = 0.1, iter = 50, min_leaf = 105	SCFs	76.05 ± 1.05	74.57 ± 0.23	99.17 ± 3.02	85.13 ± 1.25
CNN + RF tree = 3, leaf = 5	SCFs	50.33 ± 2.89	68.73 ± 3.97	38.81 ± 4.91	49.61 ± 4.93

For the arousal scenario, the optimal values are K of 5 for the KNN; tree number of 15 and leaf number of 5 for the BG; learning rate of 0.9, iteration of 50, and minimum leaf of 75 for the BS; tree number of 10 and leaf number of 5 for the RF; network depth of 2, network section of 3, learning rate of 0.005, momentum of 0.9 and L2 regulation of 0.15 for the CNN. The four ML models, which are combined with the above extractor, are given the optimal parameter values by the grid search method, such as KNN with K of 53; BG and RF with tree number of 3 and leaf number of 5; BS with learning rate of 0.1, iteration of 50, and minimum leaf of 105.

4.4. Model Validation

A total of 18 models are validated for their detection performance using 3 input feature sets and wise-subject 5-fold CV procedure on the validation set of valence and arousal scenarios in which 2, 8, and 8 of CNN, ML, and hybrid models using CNN as the feature extractor, respectively. Table 4 shows the highest classification accuracy of the above models corresponding to the input feature set. The performance comparisons between the proposed algorithm and the existing works for valence and arousal scenarios are given in Table 5.

4.5. Proposed Emotion Recognition Algorithm

The highest accuracy of 70.43% and 76.05% for valence and arousal scenarios is produced by the BS model using CNN as the feature extractor and the SCFs as represented in Table 5. Therefore, we propose an intelligent algorithm for emotion recognition as shown in Figure 4.

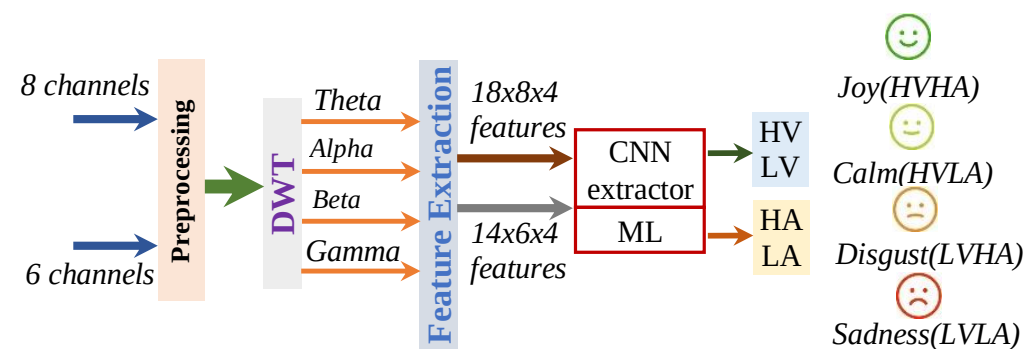


Figure 4. Proposed flow of emotion detection algorithm.

Table 5. Performance comparison of the proposed algorithm to the existing works for valence/arousal on the DEAP dataset.

Ref., Year	Method	Ac (%)	F1 (%)	Cons	Pros
[14], 2021	- Scalogram images constructed by CWT and preselected 10 channels. - Deep CNN model.	61.5/58.5	NA	- Preselection of 10 channels - Low detection accuracy	- Effective scalogram images as the input of DL models. - Reliability improved by cross-dataset validation.
[19], 2023	- Welch method for signal decomposition. - PSD of bands as extracted features. - PCA for dimension reduction and quantum SVM for classification.	65.6/75.0	77.0/60.0	- Only PSD considered for feature extraction. - Time-consuming for kernel selection of quantum SVM.	- Productive quantum-based SVM classifier in comparison with conventional SVM model. - Effective PCA for feature selection.
[21], 2023	- Generation of different sub-bands for the construction of topographic maps. - Spatial feature extraction using CNN - Temporal context learning using bi-directional LSTM for classification.	67.5/68.3	68.0/68.2	- Only a method of Cartesian coordinates for sub-band construction. - Time-consuming for DL parameter and structure optimization	- Effective topographic maps converted from sub-bands as the input. - Better deep learning representatives of CNN and LSTM.
[22], 2021	- Segmentation of EEG signals. - Combination of CNN, RNN, GAN for emotion detection.	56.4/58.5	70.8/72.0	- Low detection accuracy. - Time-consuming for training. - Difficulties of parameter selection	- A combination of different models to construct a detection network. - Deep features improved by multiple learning processes.
[26], 2023	- 3D spatial-spectral feature extraction using the Welch method applied for different sub-bands. - Generation of adjacency matrix using contextual loss computed by trainable adjacency relation in combination with GCN.	57.7/58.3	NA	- Preselection of DL model. - Time-consuming for adjacency matrix calculation	- Effective 3D spatial-spectral feature with adjacency matrix using trainable adjacency relation for optimization.
[27], 2021	- Extraction of 4 features. - Utility of different DL and ML for the performance comparison.	58.4/NA	56.0/NA	- Low detection accuracy. - Time consuming for DL parameter optimization	- Proposal of BiLSTM with hyper-parameter optimization. - Utility of a feature for the proposed algorithm.
Proposed algorithm	- Dual GA for feature and channel selection. - CNN as the feature extractor. - BS as the classifier	70.4/76.1	81.8/85.1	- Time-consuming for GA and CV procedures. - Small dataset.	- Feature quality improvement by dual GA for channel and feature selection and hybrid learning process. - Grid search for hyperparameter optimization. - Investigation of a large number of input features. - Similarly optimal features extracted from individual sub-bands. - Performance comparison of various ML and DL models.

NA: Not Available.

5. Discussion

The main purpose of this paper is to propose an efficient emotion recognition algorithm, which is reliable, simple, and easy to apply in the practical environment but remains relatively high performance.

The method development of the previous studies requires the decomposition techniques for the generation of various sub-bands, which are theta, alpha, beta, gamma, and delta corresponding to different frequency bands, from the preprocessed EEG signals. Clearly, 4 sub-bands and 32 channels are the most widely used for feature extraction, which results in a huge input feature space. Therefore, the existing works consider a small number of the original features to reduce the feature dimension and complexity for the method development such as 9, 10, 4, 6, 4, 1, 13, 1 features of [4,14,15,18,20,22,23,29], respectively. However, numerous features, which are omitted certainly, are able to contribute significantly to the final emotion detection performance of the proposed algorithm. Consequently, a large set of 42 original features is investigated in our research for better selection of the informative feature subsets.

Another significant characteristic is that the combined activities of different human brain areas represent various fundamental brain functions including human emotions. Obviously, the interaction between brain regions related to emotions is shown by the channel correlation in a specific group. As a result, the EEG recorded from the individual brain areas using electrodes contains significant information on human emotions. In other words, the selected electrodes represent informative brain regions, which generate high-quality EEG signals for emotion recognition. Therefore, the identification of effective human brain regions plays a vital role in the noble EEG generation for the extraction of high-quality features. In this work, the dual GA-based channel and feature selection are employed to select the most relevant channels and features in terms of improvement of the final emotion detection. Moreover, the wise-subject 5-fold CV-based statistic manner is combined with the GA method resulting in reinforcement of the performance reliability of the proposed algorithm and avoidance of overfitting problem.

The cerebral cortex is divided into frontal, temporal, parietal, central, and occipital lobes as shown in Figure 3. It is noteworthy that motor and sensitive functions are of the central lobe, which makes it part of the frontal and parietal cortex [56]. The simulation results of channel selection prove that the left and frontal cortex are productive for high-quality EEG generation related to emotion recognition. Although selected lobes are located over the entire cortex, most of them are on the left cortex as shown in Figure 3. Moreover, higher execution including emotional regulation, planning, reasoning, and problem-solving is implemented in the frontal brain area. Therefore, the EEG signals generated from this brain region using 4 channels contain a high proportion of relevant information related to human emotions. There are 3 and 2 EEG channels selected from the central and parietal cortex, which is responsible for motion and sensory information. Visual information and language processing are produced by the occipital and temporal cortex corresponding to OZ and T7 channels selected by the GA method. Clearly, all brain lobes are related to human emotions but only a few specific areas, where electrodes are located selectively for the EEG collection, are highly significant for emotion detection.

A total of 11 channels corresponding to different brain regions are addressed as the most relevant channels for feature extraction. In addition, the GA combined with wise-subject 5-fold CV is adopted for the feature selection in this research. Consequently, the small numbers of 18 and 14 features namely SCFs are selected for classification of the valence and arousal scenarios from 42 original features as shown in Table 2. There are 4 and 6 out of 9 models using SCFs as the optimal input feature subsets for valence and arousal classifications as given in Table 4, which means that SCFs are productive for

emotion detection. Moreover, the highest emotion classification accuracy of 70.43%/76.05% for valence/arousal classification released by a hybrid algorithm of CNN and BS shows the effectiveness of the selected feature subsets known as SCFs by the GA method.

It is noteworthy that an input feature set is extracted from 4 sub-bands generated by the DWT from the preprocessed EEG signals, which is totally different from Pearson Correlation Coefficient (PCC) feature as shown in [56]. Here, the authors combine CNN, Sparse autoencoder, and deep neural network (DNN) models as the proposed algorithm in which CNN and Sparse autoencoder (SAE) are used to construct the 2-dimension features from the PCC, and the DNN is as the classifier. Similarly, a 2-dimensional matrix of differential entropy features extracted from different sub-bands of a one-second EEG signal is provided in [57]. Indeed, The spatial location of the channel is employed for the construction of the 2-dimension features, which are then fed into a combination algorithm of CNN and improved transformer encoders. Obviously, the most important contribution of the above studies, which results in the relatively high performance of the final algorithm, is the utility of 2-dimension deep features, data reconstruction and transformation by the CNN, SAE [58] and CNN, improved transformer encoders [57]. However, the valuation method such as cross-valuation has not been used for the models and input datasets, which possibly make these researches obtain higher performance but less reliable for the applications in the practical environments.

Table 5 shows the limitations of this work and compares the proposed algorithm with recent studies related to human emotion recognition performance using the DEAP dataset. Obviously, our proposed algorithm outperforms the previous methods, which makes it potential for practical environment applications.

Various advanced decomposition techniques are investigated for the future research to improve the decomposed signal quality resulting in better extracted features, which certainly increase the final human emotion recognition performance.

6. Conclusions

Prompt detection of human emotions to enable efficient brain-computer interface played a very important role in different fields such as clinics, entertainment, and biometric security. Therefore, the proposal of a novel algorithm for human emotion recognition have been paid intensive attention from biomedical experts, clinic technicians, and game developers. Moreover, the proposed algorithm needs to show a significantly high emotion detection performance and significant reliability for practical environment applications using advanced techniques such as artificial intelligence.

In this paper, we have proposed a novel and effective human emotion recognition algorithm for practical applications. The proposed algorithm is addressed with the CNN and the BS models used as a hybrid paradigm of feature extractor and classifier using two subsets of 18 and 14 features extracted from 4 sub-bands by the DWT for valence and arousal classifications. The investigation of a large number of the original input features allows a productive search for the relevant feature subsets to improve the final emotion detection performance. Furthermore, we implement a dual GA using a ML model as the fitness function in combination with a wise-subject 5-fold CV procedure for the channel and feature selection in which the first and second GAs identified the optimal channels and feature subsets. The dual GA confirmed the superiority in terms of its reliability and effectiveness over the other feature selection methods due to the validated performance of the proposed hybrid algorithm using SCFs. Indeed, the statistically validate performance results with Ac of 70.43%/76.05%, Pn of 69.88%/74.57%, Re of 98.70%/99.17%, and F1 of 81.83%/85.13% for valence/arousal classifications proven that the frontal and left areas of human cortex are significant for the collection of the high-quality EEG signals related

to emotion recognition. Hence, our algorithm is well-fitted for practical applications in real environments.

Author Contributions: Conceptualization, M.T.N.; Methodology, D.N.; Validation, M.T.N.; Writing—original draft, D.N.; Writing—review and editing, M.T.N. and K.Y.; Visualization, D.N.; Supervision, K.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: DEAP Database at <https://www.eecs.qmul.ac.uk/mmv/datasets/deap/>.

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

References

1. Khare, S.K.; Blanes-Vidal, V.; Nadimi, E.S.; Acharya, U.R. Emotion recognition and artificial intelligence: A systematic review (2014–2023) and research recommendations. *Inf. Fusion* **2024**, *102*, 1–36. [\[CrossRef\]](#)
2. Lin, X.; Chen, J.; Ma, W.; Tang, W.; Wang, Y. EEG emotion recognition using improved graph neural network with channel selection. *Comput. Methods Programs Biomed.* **2023**, *231*, 1–11. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Wang, X.; Ren, Y.; Luo, Z.; He, W.; Hong, J.; Huang, Y. Deep learning-based EEG emotion recognition: Current trends and future perspectives. *Front. Psychol.* **2023**, *14*, 1–16. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Nandini, D.; Yadav, J.; Rani, A.; Singh, V. Design of subject independent 3D VAD emotion detection system using EEG signals and machine learning algorithms. *Biomed. Signal Process. Control* **2023**, *85*, 1–15. [\[CrossRef\]](#)
5. Hassouneh, A.; Mutawa, A.M.; Murugappan, M. Development of a Real-Time Emotion Recognition System Using Facial Expressions and EEG based on machine learning and deep neural network methods. *Inform. Med. Unlocked* **2020**, *20*, 1–9. [\[CrossRef\]](#)
6. Prabowo, D.W. A Systematic Literature Review of Emotion Recognition Using EEG Signals. *Cogn. Syst. Res.* **2023**, *40*, 101152. [\[CrossRef\]](#)
7. Abdulrahman, A.; Muhammet, B. A Comprehensive Review for Emotion Detection Based on EEG Signals: Challenges, Applications, and Open Issues. *Trait. Du Signal* **2021**, *38*, 1189–1200. [\[CrossRef\]](#)
8. Ekman, P. An Argument for Basic Emotions. *Cogn. Emot.* **1992**, *6*, 169–200. [\[CrossRef\]](#)
9. Russell, A. Core Affect and the Psychological Construction of Emotion. *Psychol. Rev.* **2003**, *110*, 145–172. [\[CrossRef\]](#)
10. Abdel-Hamid, L. An Efficient Machine Learning-Based Emotional Valence Recognition Approach towards Wearable EEG. *Sensors* **2023**, *23*, 1255. [\[CrossRef\]](#)
11. Aldawsari, H.; Al-Ahmadi, S.; Mohammad, F. Optimizing 1D-CNN-Based Emotion Recognition Process through Channel and Feature Selection from EEG Signals. *Diagnostics* **2023**, *13*, 2624. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Elouaham, S.; Dliou, A.; Jenkal, W.; Louzazni, M.; Zougagh, H.; Dlimi, S. Empirical Wavelet Transform Based ECG Signal Filtering Method. *J. Electr. Comput. Eng.* **2024**, *2024*, 9050909. [\[CrossRef\]](#)
13. Kamble, K.; Sengupta, J. A Comprehensive Survey on Emotion Recognition Based on Electroencephalograph (EEG) Signals. *Multimed. Tools Appl.* **2023**, *82*, 27269–27304. [\[CrossRef\]](#)
14. Li, Z.; Qiu, L.; Li, R.; He, Z.; Xiao, J.; Liang, Y.; Wang, F.; Pan, J. Enhancing BCI-Based Emotion Recognition Using an Improved Particle Swarm Optimization for Feature Selection. *Sensors* **2020**, *20*, 3028. [\[CrossRef\]](#)
15. Anuragi, A.; Singh Sisodia, D.; Bilas Pachori, R. EEG-Based Cross-Subject Emotion Recognition Using Fourier-Bessel Series Expansion Based Empirical Wavelet Transform and NCA Feature Selection Method. *Inf. Sci.* **2022**, *610*, 508–524. [\[CrossRef\]](#)
16. Yang, L.; Chao, S.; Zhang, Q.; Ni, P.; Liu, D. A Grouped Dynamic EEG Channel Selection Method for Emotion Recognition. In Proceedings of the 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Houston, TX, USA, 9–12 December 2021.
17. Pandey, P.; Seeja, K.R. Subject Independent Emotion Recognition System for People with Facial Deformity: An EEG Based Approach. *J. Ambient Intell. Humaniz. Comput.* **2020**, *12*, 2311–2320. [\[CrossRef\]](#)
18. Yin, Z.; Liu, L.; Chen, J.; Zhao, B.; Wang, Y. Locally Robust EEG Feature Selection for Individual-Independent Emotion Recognition. *Expert Syst. Appl.* **2020**, *162*, 113768. [\[CrossRef\]](#)

19. García-Hernández, R.A.; Celaya-Padilla, J.M.; Luna-García, H.; García-Hernández, A.; Galván-Tejada, C.E.; Galván-Tejada, J.I.; Gamboa-Rosales, H.; Rondon, D.; Villalba-Condori, K.O. Emotional State Detection Using Electroencephalogram Signals: A Genetic Algorithm Approach. *Appl. Sci.* **2023**, *13*, 6394. [\[CrossRef\]](#)
20. Saibene, A.; Gasparini, F. Genetic Algorithm for Feature Selection of EEG Heterogeneous Data. *Expert Syst. Appl.* **2023**, *217*, 119488. [\[CrossRef\]](#)
21. Aljalal, M.; Aldosari, S.A.; Molinas, M.; Alturki, F.A. Selecting EEG Channels and Features Using Multi-Objective Optimization for Accurate MCI Detection: Validation Using Leave-One-Subject-out Strategy. *Sci. Rep.* **2024**, *14*, 12483.
22. Garg, D.; Verma, G.K.; Singh, A.K. EEG-Based Emotion Recognition Using Quantum Machine Learning. *SN Comput. Sci.* **2023**, *4*, 480. [\[CrossRef\]](#)
23. Li, R.; Ren, C.; Zhang, X.; Hu, B. A Novel Ensemble Learning Method Using Multiple Objective Particle Swarm Optimization for Subject-Independent EEG-Based Emotion Recognition. *Comput. Biol. Med.* **2022**, *140*, 105080. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Tang, Y.; Wang, Y.; Zhang, X.; Wang, Z. STILN: A Novel Spatial-Temporal Information Learning Network for EEG-Based Emotion Recognition. *Biomed. Signal Process. Control* **2023**, *85*, 104999. [\[CrossRef\]](#)
25. Liang, Z.; Zhou, R.; Zhang, L.; Li, L.; Huang, G.; Zhang, Z.; Ishii, S. EEGFuseNet: Hybrid Unsupervised Deep Feature Characterization and Fusion for High-Dimensional EEG with an Application to Emotion Recognition. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2021**, *29*, 1913–1925. [\[CrossRef\]](#)
26. Cui, G.; Li, X.; Touyama, H. Emotion Recognition Based on Group Phase Locking Value Using Convolutional Neural Network. *Sci. Rep.* **2023**, *13*, 3769. [\[CrossRef\]](#)
27. Zhong, M.; Yang, Q.; Liu, Y.; Zhen, B.; Zhao, F.; Xie, B. EEG Emotion Recognition Based on TQWT-Features and Hybrid Convolutional Recurrent Neural Network. *Biomed. Signal Process. Control* **2023**, *79*, 104211. [\[CrossRef\]](#)
28. Iyer, A.; Das, S.S.; Teotia, R.; Maheshwari, S.; Sharma, R.R. CNN and LSTM Based Ensemble Learning for Human Emotion Recognition Using EEG Recordings. *Multimed. Tools Appl.* **2022**, *82*, 4883–4896. [\[CrossRef\]](#)
29. Li, W.; Wang, M.; Zhu, J.; Song, A. EEG-Based Emotion Recognition Using Trainable Adjacency Relation Driven Graph Convolutional Network. *IEEE Trans. Cogn. Dev. Syst.* **2023**, *15*, 1656–1672. [\[CrossRef\]](#)
30. Joshi, V.M.; Ghongade, R.B.; Joshi, A.M.; Kulkarni, R.V. Deep BiLSTM Neural Network Model for Emotion Detection Using Cross-Dataset Approach. *Biomed. Signal Process. Control* **2022**, *73*, 103407. [\[CrossRef\]](#)
31. Almanza-Conejo, O.; Almanza-Ojeda, D.L.; Contreras-Hernandez, J.L.; Ibarra-Manzano, M.A. Emotion Recognition in EEG Signals Using the Continuous Wavelet Transform and CNNs. *Neural Comput. Appl.* **2022**, *35*, 1409–1422. [\[CrossRef\]](#)
32. Bagherzadeh, S.; Norouzi, M.R.; Hampa, S.B.; Ghasri, A.; Kouroshi, P.T.; Hosseiniinasab, S.; Zadeh, M.A.; Nasrabadi, A.M. A subject-independent portable emotion recognition system using synchrosqueezing wavelet transform maps of EEG signals and ResNet-18. *Biomed. Signal Process. Control* **2024**, *90*, 105875. [\[CrossRef\]](#)
33. Bagherzadeh, S.; Shalbaf, A.; Shoeibi, A.; Jafari, M.; San Tan, R.; Acharya, U.R. Developing an EEG-based emotion recognition using ensemble deep learning methods and fusion of brain effective connectivity maps. *IEEE Access* **2024**, *12*, 50949–50965. [\[CrossRef\]](#)
34. Bagherzadeh, S.; Maghooli, K.; Shalbaf, A.; Maghsoudi, A. Emotion recognition using continuous wavelet transform and ensemble of convolutional neural networks through transfer learning from electroencephalogram signal. *Front. Biomed. Technol.* **2023**, *10*, 47–56. [\[CrossRef\]](#)
35. Bagherzadeh, S.; Maghooli, K.; Shalbaf, A.; Maghsoudi, A. A Hybrid EEG-based emotion recognition approach using wavelet convolutional neural networks and support vector machine. *Basic Clin. Neurosci.* **2023**, *14*, 87–101. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Algarni, M.; Saeed, F.; Al-Hadhrani, T.; Ghabban, F.; Al-Sarem, M. Deep Learning-Based Approach for Emotion Recognition Using Electroencephalography (EEG) Signals Using Bi-Directional Long Short-Term Memory (Bi-LSTM). *Sensors* **2022**, *22*, 2976. [\[CrossRef\]](#)
37. Moctezuma, L.A.; Abe, T.; Molinas, M. Two-Dimensional CNN-Based Distinction of Human Emotions from EEG Channels Selected by Multi-Objective Evolutionary Algorithm. *Sci. Rep.* **2022**, *12*, 3523. [\[CrossRef\]](#)
38. Kouka, N.; Fourati, R.; Fdhila, R.; Siarry, P.; Alimi, A.M. EEG Channel Selection-Based Binary Particle Swarm Optimization with Recurrent Convolutional Autoencoder for Emotion Recognition. *Biomed. Signal Process. Control* **2023**, *84*, 104783. [\[CrossRef\]](#)
39. Koelstra, S.; Muhl, C.; Soleymani, M.; Lee, J.S.; Yazdani, A.; Ebrahimi, T.; Pun, T.; Nijholt, A.; Patras, I. DEAP: A Database for Emotion Analysis Using Physiological Signals. *IEEE Trans. Affect. Comput.* **2012**, *3*, 18–31. [\[CrossRef\]](#)
40. DSundararajan, D. *Discrete Wavelet Transform: A Signal Processing Approach*, 1st ed.; John Wiley & Sons: Singapore, 2015.
41. Furht, B. (Ed.) *Discrete Wavelet Transform (DWT)*. In *Encyclopedia of Multimedia*; Springer: Boston, MA, USA, 2008; p. 188.
42. Yuvaraj, R.; Thangavel, P.; Thomas, J.; Fogarty, J.S.; Ali, F. Comprehensive Analysis of Feature Extraction Methods for Emotion Recognition from Multichannel EEG Recordings. *Sensors* **2023**, *23*, 915. [\[CrossRef\]](#)
43. Kalashami, M.P.; Pedram, M.M.; Sadr, H. EEG Feature Extraction and Data Augmentation in Emotion Recognition. *Comput. Intell. Neurosci.* **2022**, *2022*, 7028517. [\[CrossRef\]](#)

44. Rostaghi, M.; Azami, H. Dispersion Entropy: A Measure for Time-Series Analysis. *IEEE Signal Process. Lett.* **2016**, *23*, 610–614. [[CrossRef](#)]
45. Chen, W.; Wang, Z.; Xie, H.; Yu, W. Characterization of Surface EMG Signal Based on Fuzzy Entropy. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2007**, *15*, 266–272. [[CrossRef](#)] [[PubMed](#)]
46. Tripathy, R.K.; Sharma, L.; Dandapat, S. Detection of Shockable Ventricular Arrhythmia Using Variational Mode Decomposition. *J. Med. Syst.* **2016**, *40*, 79. [[CrossRef](#)] [[PubMed](#)]
47. Patel, P.; R, R.; Annavarapu, R.N. EEG-Based Human Emotion Recognition Using Entropy as a Feature Extraction Measure. *Brain Inform.* **2021**, *8*, 20. [[CrossRef](#)]
48. Khare, S.K.; Bajaj, V.; Sinha, G.R. Adaptive Tunable Q Wavelet Transform-Based Emotion Identification. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 9609–9617. [[CrossRef](#)]
49. Hatamikia, S.; Maghooli, K.; Nasrabadi, A. The Emotion Recognition System Based on Autoregressive Model and Sequential Forward Feature Selection of Electroencephalogram Signals. *J. Med. Signals Sens.* **2014**, *4*, 194. [[CrossRef](#)]
50. Taran, S.; Bajaj, V. Emotion Recognition from Single-Channel EEG Signals Using a Two-Stage Correlation and Instantaneous Frequency-Based Filtering Method. *Comput. Methods Programs Biomed.* **2019**, *173*, 157–165. [[CrossRef](#)]
51. Amann, A.; Tratnig, R.; Unterkofler, K. Detecting Ventricular Fibrillation by Time-Delay Methods. *IEEE Trans. Biomed. Eng.* **2007**, *54*, 174–177. [[CrossRef](#)]
52. Zhang, X.S.; Zhu, Y.S.; Thakor, N.V.; Wang, Z.Z. Detecting Ventricular Tachycardia and Fibrillation by Complexity Measure. *IEEE Trans. Biomed. Eng.* **1999**, *46*, 548–555. [[CrossRef](#)]
53. Jekova, I. Shock Advisory Tool: Detection of Life-Threatening Cardiac Arrhythmias and Shock Success Prediction by Means of a Common Parameter Set. *Biomed. Signal Process. Control* **2007**, *2*, 25–33. [[CrossRef](#)]
54. Duda, R.O.; Hart, R.O.; Stork, D.G. *Pattern Classification*, 2nd ed.; Wiley: New York, NY, USA, 2001.
55. Hastie, T.; Friedman, J.; Tibshirani, R. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer: New York, NY, USA, 2001.
56. Ribas, G.C. The Cerebral Sulci and Gyri. *Neurosurg. Focus* **2010**, *28*, E2. [[CrossRef](#)] [[PubMed](#)]
57. Zhang, X.; Cheng, X.; Liu, H. TPRO-NET: An EEG-Based Emotion Recognition Method Reflecting Subtle Changes in Emotion. *Sci. Rep.* **2024**, *14*, 13491. [[CrossRef](#)] [[PubMed](#)]
58. Liu, J.; Wu, G.; Luo, Y.; Qiu, S.; Yang, S.; Li, W.; Bi, Y. EEG-Based Emotion Classification Using a Deep Neural Network and Sparse Autoencoder. *Front. Syst. Neurosci.* **2020**, *14*, 43. [[CrossRef](#)] [[PubMed](#)]

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