An AI-empowered affect recognition model for healthcare and emotional well-being using physiological signals

Zijian Zhou1 • Muhammad Adeel Asghar2 • Daniyal Nazir3 • Kamran Siddique1,4 • Mohammad Shorfuzzaman5 • Raja Majid Mehmood1

Received: 23 February 2022 / Revised: 17 June 2022 / Accepted: 3 August 2022 / Published online: 3 November 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
Affective Computing is one of the central studies for achieving advanced human-computer interaction and is a popular research direction in the field of artificial intelligence for smart healthcare frameworks. In recent years, the use of electroencephalograms (EEGs) to analyze human emotional states has become a hot spot in the field of emotion recognition. However, the EEG is a non-stationary, non-linear signal that is sensitive to interference from other physiological signals and external factors. Traditional emotion recognition methods have limitations in complex algorithm structures and low recognition precision. In this article, based on an in-depth analysis of EEG signals, we have studied emotion recognition methods in the following respects. First, in this study, the DEAP dataset and the excitement model were used, and the original signal was filtered with others. The frequency band was selected using a butter filter and then the data was processed in the same range using min–max normalization. Besides, in this study, we performed hybrid experiments on sash windows and overlays to obtain an optimal combination for the calculation of features. We also apply the Discrete Wave Transform (DWT) to extract those functions from the preprocessed EEG data. Finally, a pre-trained k-Nearest Neighbor (kNN) machine learning model was used in the recognition and classification process and different combinations of DWT and kNN parameters were tested and fitted. After 10-fold cross-validation, the precision reached 86.4%. Compared to state-of-the-art research, this method has higher recognition accuracy than conventional recognition methods, while maintaining a simple structure and high speed of operation.

Keywords Cognitive computing • Discrete wavelet transform • Emotion recognition • Emotional healthcare • Well-being

1 Introduction
Emotion recognition is a process of identifying people’s emotional states by analyzing their physiological or non-physiological information. There are many directions in the research of emotion recognition and the common method is to study the recognition and classification of emotion through some physiological characteristics, such as human facial expression, behavior, and language intonation. Although these external physiological characteristics are easy to obtain, their physiological performance is easy to be controlled by people’s consciousness [1]. This study is more focused on bio-signals such as electroencephalogram (EEG) where cognitive thinking is complex but readable by brain sensors. The current advancement of sensors technologies made these challenges simpler to understand inner thinking of the human brain, only when sensors are implanted invasively. Studying human emotions using EEG signals helps humans interact with AI systems. With the advent of smart health technologies and systems, EEG-based sentiment analysis has enabled patient-related media processing to serve monitoring frameworks that engage stakeholders in smart cities [2]. Deep learning models also play an important role in quickly and accurately recognizing the effects of EEG-based classification models [3, 4].

Muhammad Adeel Asghar, Daniyal Nazir, Kamran Siddique, Mohammad Shorfuzzaman, Raja Majid Mehmood contributed equally to this work.

Extended author information available on the last page of the article
1.1 Emotions for healthcare and smart cities

Human emotions can be used effectively in healthcare and smart cities for real-time IoT applications [5]. Due to complex experimentation of invasive techniques, many researchers prefer to have a non-invasive approach [6]. But still, researchers are working very hard with non-invasive devices such as EEG where human brain signals were collected from the top of the scalp without any surgery. In recent years, many good studies have emerged on a range of EEG devices that are very easy to experiment with and inexpensive. Because of this potential, human motivation and robotics emotions could be the future direction for many researchers. Intelligent robotics is evolving field, having many applications in manufacturing, healthcare, and other services. But still, robots are not capable to think with motivations and emotions. Due to these limitations, robotics failed in many situations of uncertainty, situations where robots could not replace humans’ brain such as human wisdom, critical thinking or decision making, and so on. It will help introduce human traits into robots using Human-Robot Interaction (HRI) [7]. It can also play an important role in the development of complex and intelligent HRI-based systems, which will be a major innovation in modern robotics [8].

1.2 Research objective

The purpose of this work is to extract and classify the emotional features of participants based on the EEG dataset “Dataset for Emotion Analysis using Physiological signals” (DEAP) and achieve the highest recognition accuracy while maintaining a lightweight framework. A data set is a collection of physiological signals such as EEG brain signals and other artifacts such as eye movements. This dataset is used to train and test the proposed model. In this study, the Valence-Arousal emotion model [9] is considered, and discrete wavelet transform (DWT) is adopted to calculate emotional features, and kNN model is built for classification. EEG signal acquisition process is easy to be interfered with by various noises, this paper proposed a signal optimization method based on the combination of Butter-filter and min–max normalization in processing the original data. At the same time, to determine the sliding window size for feature extraction, the original data are divided into different sliding windows and considered the corresponding overlapping. Through extensive experiments, these results show that the DWT method can be used in an efficient feature extraction process without complex structures and that the kNN model has achieved the highest recognition speed. Recent trends and the pandemic situation have also motivated researchers to raise awareness of COVID-19 in smart healthcare. Therefore, the development of detection and diagnosis methods based on artificial intelligence (AI) will be very useful [10–12].

1.3 Organization

In this study, a single machine learning model was employed for emotion recognition and it is combined with all DEAP dataset (32 subjects). Overall, results show that this is a significant contribution to the affective computing research field. The proposed method achieved the highest recognition rate over state-of-art methods. Since maintaining the lightweight system, the main reason for improving the accuracy of this research is the test and optimization of data and parameters, which would be explained in below sections. The following part of this paper is organized as: Sect. 2 introduced some methods of EEG signal preprocessing, feature extraction and recognition, and made a comparison table of similar research in recent 5 years. Section 3 proposed our method. In Sect. 4, we provide a set of experiments with detailed steps and made some evaluations. Section 5 discussed the comparison between our method and the original author’s method and provided some details of comparison. Section 6 summarized our work and discussed the possible optimization direction.

1.4 Research objective

Since EEG signals are inherently nonlinear and subjective, the main goal of this study is to analyze the effects using EEG signals. To address this problem, we proposed a model with quality feature selection and extraction processes to classify high-performance effects in shorter processing times. We use preprocessing based on discrete wavelet transforms to solve the problem of nonlinearity in EEG signals. The central goal of this study is to extract and classify participants’ emotional activities while achieving the highest recognition accuracy. In this study, the valence arousal emotion model used in the DEAP dataset was considered. The proposed emotion recognition framework includes preprocessing, feature extraction, feature selection, and classification using butter filter and min–max normalization, energy entropy, discrete wavelet transform (DWT), and k-nearest neighbor (KNN) , respectively. The classification process is well described with the following algorithm, making it easy for the reader to understand how this experiment is used to obtain an optimal function of the human input signal.
2 Related work

2.1 Feature extraction

In the study of emotion recognition based on EEG signal, feature extraction plays an extremely important role. Only by choosing the correct features that can effectively express emotions can we proceed to the next classification research. Some international experts have done a lot of correlation experiments on feature extraction based on EEG signals, and combined with psychology, physics and neurology, they have found out some useful EEG features that can be used to identify emotions. On the basis of approximate entropy, Dhall et al. [13] proposed the sample entropy algorithm which is more convenient to calculate and retain the advantages of approximate entropy. In the calculation process, the sample entropy requires shorter data length than the approximate entropy, but the estimated value is more stable. Many domestic researchers have proved that sample entropy is more conducive to emotion classification of EEG signals [14, 15]; Kahou et al. [16] extracted P-QRS-T wave energy features for emotion recognition and classification, and through experiments, it is concluded that this feature has a good representation of emotional state, especially in the optimistic, happy and other normal emotions, it has a great effect on the classification results; Mirmamadi et al. [17] used Fourier transform power spectrum, nonlinear dynamic features and wavelet band energy features to identify and classify EEG emotions, and compared the classification accuracy under each feature. Through experiments, it was found that the classification effect of power spectrum features was better than other features; Poria et al. [18] used the frequency characteristics of EEG signals and extracted the asymmetry and incoherence of EEG signals as features to identify and classify three types of emotions (anger, fear and surprise), with an average recognition accuracy of 66.3%; Trigeorgis et al. [19] used differential entropy and energy spectrum for feature extraction of emotion classification, and used differential entropy feature to classify positive emotion and negative emotion. The classification results using differential entropy feature and energy feature reached 82.22% and 76.56%. With the advent of IoT and smart health systems, the authors of [20] used 5G and interconnected devices to create a smart health framework. The proposed framework will greatly contribute to seamless and personalized emotion recognition care services for 5G [20].

2.2 Emotion recognition

One method of EEG emotion recognition is unsupervised emotion recognition. Without knowing the sample category, the samples with similar characteristics are close to each other, and the samples with different characteristics are far away from each other. Common unsupervised learning methods include K-means clustering algorithm, fuzzy clustering algorithm and self-organizing mapping algorithm [21]. The other method is supervised emotion recognition. Compared with unsupervised emotion recognition, supervised emotion recognition needs to know the emotion tag of the EEG data to be recognized in advance. The EEG data training set with emotion tag is used to train the classification model, which is verified by the EEG signal test set. Common supervised learning methods include support vector machine (SVM), neural network (NN), decision tree (DT), Bayesian network (BN), KNN and hidden Markov model (KM) [22]. Tzirakis et al. [23] used the fuzzy clustering algorithm (fuzzy c-means and fuzzy k-means) in unsupervised learning to cluster three kinds of emotions, and observed the clustering results through experiments, and selected the feature set that is most conducive to emotion recognition; Zheng et al. [24] divided EEG emotions into six categories (happy, angry, disgusted, sad, surprised and afraid), and used Bayesian linear discriminant to classify them, and the accuracy rate reached more than 70%; Wright et al. [25] collected EEG signals of 16 healthy people through three channels. The extracted features were classified and compared by quadratic discriminant analysis, KNN, Mahalanobis distance and support vector machine. Emotions were divided into six categories: happiness, surprise, fear, anger, disgust and sadness. The classification accuracy of quadratic discriminant analysis and support vector machine reached 62.3% and 83.3% respectively.

2.3 DEAP-based state-of-art methods

This work compared our method and precision with those adopted by other scientists in the same data set. The articles were selected from the last 5 years, including their publication, description, source of accuracy, accuracy, and validation method used. The comparison table for two emotions is shown in Table 1.

3 Methods

The core objective of this study is to extract and classify the emotional activity of participants, and at the same time to achieve the highest recognition accuracy. In this study, we considered the valence-arousal emotion model which was employed by the DEAP dataset. The proposed emotion recognition framework contains preprocessing, features extraction, features selection, and classification with butter-filter and min–max normalization, energy-entropy, discrete
wavelet transform (DWT), and kth nearest neighbors (KNN), respectively. The classification process is well explained by the following algorithm where reader can easily understand how this experiment is used to acquire optimal features from input human signals.

### 3.1 Experiment dataset

DEAP [24] (database for emotion analysis using physical signals), the database is based on the physiological signals induced by music video materials. It records the physiological signals of 32 subjects who watched 40 min of music video (each music video for 1 min) and the psychological scale of valence, arousal, dominance and liking of the video, as well as the facial expression videos of the top 22 participants. The database can be used to study the physiological signals under multimodal conditions, which is of great significance to the study of emotional EEG. Table 2 shows some brief introductions about this experiment.

### 3.2 Data preprocessing

In this experiment, the EEG data of the first 20 films were watched by the first 32 subjects. Using MATLAB and python (Numpy) format to preprocess the physiological data in the experiment: down sampling, EOG removal, filtering, segmentation, etc. The data were divided into 3 s baseline and 60 s baseline. After preprocessing, EEG data of a participant is converted into 40

### 3.3 Wavelet transform—DWT

For continuous wavelets, scale ‘a’ and time ‘t’ are continuous. If the computer is used for calculation, it must be discretized to obtain DWT. All the equations below are processed to perform DWT [29]. Take the scale parameter ‘a’ and translation parameter ‘b’ of continuous wavelet transform as:

\[
  a = a_0^j, \quad b = k a_0^j b_0
\]

where \( J \in \mathbb{Z} \) and extension step \( a_0 \neq 1 \) are fixed values. For convenience, it is always assumed that \( a_0 > 1 \), and the corresponding discrete wavelet function \( \psi(t) \) can be written as:

| Reference, year & publication | Description, accuracy source | A or V avg. | Validation |
|-------------------------------|------------------------------|-------------|------------|
| Liu et al. [15], 2018 IEEE Transactions on Cognitive and Developmental Systems | Empirical Mode Decomposition (EMD) and KNN, Average Subjects | 86.4%/84.9% | 10-fold |
| Cao et al. [16], 2019 Chinese Control Conference | PCA and CNN, Average Subjects | 81.2%/83.3% | 10-fold |
| Byunet et al. [17], 2017 International Conference on Robotics and Automation Sciences | Relief algorithm and Bhattacharyya distance, SVM, Combined Subjects | 81.3%/81.9% | 10-fold |
| Shahnaz et al. [18], 2016 IEEE Region 10 Conference | DWT, PCA and SVM, Average Subjects | 74.1%/75.5% | 10-fold |
| Li et al. [19], 2015 IEEE International Conference on Systems, Man, and Cybernetics | KNN, CNN, Not Mentioned | 86.5%/81.9% | 8-fold |
| Xiang Li et al. [20], 2016 IEEE International Conference on Bioinformatics and Biomedicine | R-CNN, Combined Subjects | 72.1%/74.1% | 8-fold |
| Yang et al. [21], 2018 International Joint Conference on Neural Networks | Deep belief network (DBN), SVM, LSTM, Average Subjects | 90.8%/91.3% | 5-fold |
| Shao et al. [22], 2019 IEEE International Conference on Image Processing | SVM, baive byes (NB), SVM model with decision fusion (SVM-DF), Not Mentioned | 80.9%/75.9% | 5-fold |

Table 1 Comparison table for similar research on DEAP

| Feature | Description |
|---------|-------------|
| Number of subjects | 32 |
| Number of videos | 40 |
| Number of EEG channels | 32 |
| Labels | Valence, arousal |
| Sampling rate | 128 Hz |

Table 2 DEAP dataset experiment description
\[ \psi_{j,k}(t) = \frac{1}{\sqrt{2^j}} \psi(t - k) \]  

The coefficients of DWT can be expressed as:

\[ C_{j,k} = \langle f(t), \psi_{j,k}(t) \rangle = \int_{-\infty}^{\infty} f(t) \psi_{j,k}(t) \, dt \]  

Its reconstruction formula is:

\[ f(t) = A \sum_{j} \sum_{k} C_{j,k} \psi_{j,k}(t) \]  

\('A' is a constant independent of the signal. Like continuous wavelet transform, DWT also requires a filter to obtain a proportional representation of the digital signal in the time domain. In the DWT, the filter will cut off certain frequency components of the signal under different scale conditions: the signal passes through different high-pass filters to obtain a series of signal high-frequency components, and through different low-pass filters to obtain a series of high frequency components.

### 3.4 Normalization

There are many methods of data normalization. The commonly used methods in the field of EEG signal analysis are “min–max normalization method” “Z-score normalization method”, “mean value normalization method”, “logarithmic function method” [31], etc. This study adopted min–max normalization because achieves the best result, and it is defined as:

\[ y(k) = \frac{x(k) - \min(x(n))}{\max(x(n)) - \min(x(n))}, k = 1, 2, 3...n \]  

where \( \min(x(n)) \) represents the minimum value of sample data \( x(n) \), and \( \max(x(n)) \) represents the maximum value of sample data \( x(n) \). This formula can map an original value \( x(k) \) to a value in the interval [0, 1] through min max normalization. This is the method used in this research.

### 3.5 Wavelet packet feature extraction

In this article, we extract two important characteristics in EEG time-frequency analysis: energy in three frequency bands and wavelet entropy in three frequency bands. Calculate the characteristic using the following equation [20]. The three frequency bands represent \( \alpha \), \( \beta \), and \( \gamma \). The formula for calculating the signal energy for each frequency band is as follows.

\[ E = \sum_{i=0}^{n} x_i^2 \]  

where \( x \) is the amplitude of discrete point of time domain signal, \( n \) is the number of discrete signal point. Electrode energy difference of four frequency bands: the average energy difference of EEG signal obtained on 32 electrodes in each second of four frequency bands. Wavelet packet entropy: calculate the energy of four frequencies and their energy sum and calculate the wavelet entropy of EEG according to formulas 7 and 8.

\[ p(j) = \frac{E(j)}{E_{\text{Total}}} \]  

\[ \text{Entro} = - \sum_j p(j) \log(p(j)) \]  

In Eq. 7, \( E_{\text{total}} \) is the sum of energy of four frequency bands. According to the size of the window, the sum of energy will also change. \( J \) is the number of any frequency band in the four frequency bands, and \( \text{entro} \) represents the wavelet entropy value.

### 3.6 Classification—KNN

KNN is the main method used and classified in this study and is well suited to the characteristics of the EEG. As a non-parametric classification algorithm, the KNN algorithm is very powerful and easy to implement. It is widely used for classification, regression, and pattern recognition. In theory, it is one of the most mature and simple machine learning algorithms (Fig. 1).

There are several ways to measure the distance between points in space, such as calculating the Manhattan distance, calculating the Euclidean distance, etc. However, in general, the Euclidean distance [21] is used in KNN algorithms. Here is a brief introduction. Taking a two-dimensional plane as an example, the formula for calculating the Euclidean distance between two points in two-dimensional space is:

\[ \rho = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \]  

Expand to multi-dimensional space, and the distance formula becomes as follows:

\[ d(x, y) = \sqrt{(x_n - y_n)^2 + (x_{n-1} - y_{n-1})^2 + ... + (x_1 - y_1)^2} \]  

Knowing the data and the labels in the training set, we input the test data, compare the features of the test data with those of the training set, and find the first \( k \) most similar data in the training set. So those categories in the test data are the ones that appear the most in the data \( k \).
4 Experiments and result

4.1 Experiment process

This research selected the first experimenter’s EEG data to explain in detail, Fig. 2 shows the initial form of the participant’s EEG signal matrix. This section will analyze and process the matrix obtained from each video. Here, different colors represent different brain channels, and each rectangular box represents a video.

The initial matrix format is $40 \times 32 \times 8064$ (video/trial $\times$ channel $\times$ data). Next, the proposed method will divide the data into windows. To determine the window size, different window sizes from 1 to 4 s, and consider the situation of window overlapping is tested. After experiments, this study found that the original accuracy of 4S window size is the highest without overlapping. The follow-up accuracy optimization experiments will be based on this situation. The windowed feature matrix of the first participants is shown in Fig. 3.

After windowing, this research adopted a Butter filter to the above matrix and divide it into three frequency bands: $\alpha$ (8–13 Hz), $\beta$ (14–30 Hz) and $\gamma$ (30–60 Hz). After testing with different K values from 1 to 15, we find that using the beta frequency band can achieve the highest accuracy, 84.4% for Valence and 88.3% for Arousal, 86.4% in average. To calculate the characteristics of each channel efficiently, the filtered beta frequency band matrix is reshaped where each row represents a window (4 s and 512 sampled points) and every 15 columns represents a channel, see Fig. 4 for a reference.

The DWT is used to calculate the energy and entropy for each channel. The features of each window are calculated, so each window can get 2 feature points. Each channel will provide 30 feature points (15 windows) with a total of 1280 lines ($32 \text{ channels} \times 40 \text{ trails}$), as shown in Fig. 5.

![Fig. 1 The way KNN performs discrimination](image1)

![Fig. 2 Original data matrix in this research](image2)
**Fig. 3** Data matrix after windowed for each participant

**Fig. 4** Matrix feature extraction process
4.2 Results

The experimental environment of this report is simulated by platform PyCharm under the operating system Windows 10. The CPU used is Intel (R) Core (TM) i7-6700HQ. From the experimental processes in Sect. 4.1, all the experiment was processed through several parameters such as sliding window, overlapping, wavelet type, decomposition level, frequency band, and of KNN parameters achieved the highest accuracy from selected DWT features. To expedite the process, we conducted window and overlap experiments on the EEG data of the first subject, and the following accuracy rates belong to the first subject. Please note that all overlaps are half the size of the experimental sliding windows, and other parameters are adopted by the original author [31]: band-pass filtering is not considered; the features are energy and entropy; wavelet DB1 and decomposition level 4 are applied; k = 3 for both arousal and valence. We can see that the optimal average recognition accuracy for the first participant is 88.4%, when the 4S windows and without overlapping is tested. Here, Fig. 6 makes it easier to determine which set of parameters can achieve the highest accuracy:

In Fig. 6, if the 4S window is selected and overlapping is not considered, the accuracy of valence can reach 89.5%, that of arousal can reach 87.2%, and the average accuracy can reach 88.4%. In this experiment, overlap is a considered factor for each window size: blue bar means overlap is not considered, red bar means overlap is considered, and the overlap size is half of the selected window size; x-axis represents different window selection under the corresponding emotional tag Arousal or Valence; average accuracy is used as the evaluation standard of the final accuracy. From single subject analysis, 88.4% represent the final accuracy in combine arousal and valence domains.

Next, this research considered how the situation will be changed when all features of 32 participants join into the data, and what corresponding optimization needs to be done. After determining which window size to choose to use and whether to use overlap, the same method is applied to all the data of 32 subjects and did the following tests. All the features extracted by using DWT from all windows within each frequency band, and KNN model with KDTree is trained as well. In order to determine the optimal fitting parameter of wavelet transform, this work tested the wavelet DB1-DB4, and decomposition level 1-5. Finally, DB4 and decomposition level 4 can best fit the original signal, as shown in Fig. 7.

Figure 7 shows all results of both Arousal/Valence accuracy under all frequency bands, wavelets, decomposition levels, and KNN values, and found Level4-DB4 performed best under the Beta frequency band. The x-axis presents different wavelets from DB1-DB5 and decomposition level 1-4; y-axis is the accuracy corresponding to emotional tags under the selection of wavelets-levels; for each emotional tag, the accuracy bar on the left represents the accuracy through the Alpha frequency band, the bar in the middle represents the Beta frequency band, and the frequency band on the right is the Gamma frequency band. After getting the best settings: Beta band, wavelet DB4 and decomposition level 4, we tested each value for K-values under best settings. Because the K-value is usually lower than 15, a large K-value will lead to the inaccuracy of the result. Here we choose to test the K-value from 1 to 15 and used 10-fold cross-validation to validate the obtained accuracy. The sorted results are shown in Fig. 8. Here, N = 10 for valence and N = 11 for arousal, achieve the highest average accuracy of 86.4%.
5 Discussion

5.1 Comparison with the original author’s method

To sum up, this research improved the accuracy from 74.6 to 86.4% based on the original author’s method and made a series of improvements regarding signal preprocessing, feature extraction, classification, and validation process, as shown in Table 3.

Since significant improvements have been made in this research, the reason for each change has been provided in Sect. 4. Compared with the original method, the influence of baseline signal is taken into account; in the preprocessing, band-pass filter and min–max normalization are applied to the original dataset; in the feature extraction and classification process, different parameters of DWT and KNN were tested and optimized; in the validation process, this study used 10-fold cross-validation to measure the

---

**Fig. 6** Single subject trail: window and overlapping size

**Fig. 7** Different wavelet and decomposition level test result
accuracy of the results and increased the accuracy rate from 74.6 to 86.4%.

5.2 Comparison with state-of-art

The purpose of this study is to extract the features and identify emotions from the EEG signal of 32 video participants based on the DEAP dataset. The arousal-valence emotion model is adopted, and DWT and KNN models are used to calculate the features and classify the emotions, respectively. Besides, the Butter filter and min–max normalization is applied to the original signal for optimizing and denoising. Some parameters like window size, overlapping, as well as the KNN model parameters are both adjusted to optimize the recognition accuracy. The validation method used is 10-fold cross-validation, and the final recognition accuracy can reach 86.4%. According to the research [32], when the amount of data suddenly increases, the accuracy rate will often decrease. However, our research considers the data of all 32 participants. Although this is a challenge to our final accuracy, this method still achieves a high accuracy rate by processing the original data and adjusting the parameters. In my review of all the literature, our method can achieve a high recognition rate, as shown in Table 4.

5.3 Comparison with the original author’s method

To sum up, this research improved the accuracy from 74.6 to 86.4% based on the original author’s method and made a series of improvements regarding signal preprocessing, feature extraction, classification, and validation process, as shown in Table 4.

The innovation of this work is that it does not use complex and tedious hierarchical structure or feature extraction methods like other studies. On the contrary, we optimize some lightweight frameworks. Therefore, we can achieve high data processing speed and recognition speed and can achieve the same or even surpass their accuracy rate with other methods. This will be in the future of big data processing and recognition. The direction of development is worth advocating.

5.4 Research limitations

The main problem of this study is that the calculation costs, the cost of the model is reduced by omitting features with low entropy or correlation coefficient and selecting features with high value. Another problem is that since the input dataset has a large number of attribute values, the input streams are also reduced to achieve good performance with
short processing times, making the model useful for applications in real time and smart health care.

6 Conclusion and future work

Emotion recognition is a new research field. The research on emotion-based on EEG is an important research topic in the field of emotion recognition and smart healthcare models. Based on the DEAP dataset, this paper introduces the basic concepts of EEG and the theory of emotion recognition, and focuses on the following aspects:

1. In order to denoise the original signal from DEAP dataset, the signal optimization methods based on Butter filter and min–max normalization are used in the processing of original data, the average accuracy of $\beta$ band can achieve the highest among all frequency bands, reaching 86.4%.

2. The main reason for the high accuracy of this research is mainly the optimization of different parameters. To determine the best window size of feature extraction, we divide the original data into different window sizes (from 1 to 4 s, 1 s for interval) and consider the overlapping. The experimental results show that, without considering the overlap, the recognition rate of a single subject is 89.5% and 87.2% when choosing the size of 4s window.

3. In addition, the DWT used in this paper can be used for simple and effective feature extraction without complex hierarchical structure, and KNN model can achieve the highest recognition rate for the classification process. In the process of the wavelet transform, wavelet DB1 to DB4 and decomposition levels 1–5 are tested. The experimental results show that DB4 wavelet and 4-level decomposition are the best combinations for the data fitting of 32 subjects. At the same time, changing the neighborhood parameters of the kNN model, namely, set $N = 10$ for valence and $N = 11$ for Arousal, can achieve the highest accuracy.

Acknowledgements This work has supported by the Xiamen University Malaysia Research Fund (XMUMRF) (Grant No: XMUMRF/2022-C9/IECE/0035). The authors are grateful to the Taif University Researchers Supporting Project Number (TURSP-2020/79), Taif University, Taif, Saudi Arabia for funding this work.

Author contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by ZZ, MAA. Research road map and framework were developed by RMM and KS. Data preparation and analysis was done by DN and MS. The first draft of the manuscript was written by ZZ and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This work has supported by the Xiamen University Malaysia Research Fund (XMUMRF) (Grant No: XMUMRF/2022-C9/IECE/0035). This work was supported by the Taif University Researchers Supporting Project Number (TURSP-2020/79), Taif University, Taif, Saudi Arabia.

Data availability The datasets analysed during the current study are available in the DEAP Dataset repository, https://www.eecs.qmul.ac.uk/mmv/datasets/deap/”

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose

Ethical approval Ethics approval was not required for this research.

Informed consent All authors signed the consent form.

Research involving human and animal participants This work did not involve humans and animals.

References

1. Albraikan, A., Tobin, D.P., El Saddik, A.: Toward user-independent emotion recognition using physiological signals. IEEE Sens. J. 19(19), 8402–8412 (2019)
2. Hossain, M.S., Muhammad, G., Alamri, A.: Smart healthcare monitoring: a voice pathology detection paradigm for smart cities. Multimed. Syst. 25(5), 565–575 (2019)
3. Hossain, M.S., Muhammad, G.: Deep learning based pathology detection for smart connected healthcare. IEEE Netw. 34(6), 120–125 (2020)
21. Geethanjali, B., Adalarasu, K., Jagannath, M., Guhan Seshadri, N.P.: Music-induced brain functional connectivity using EEG sensors: a study on Indian music. IEEE Sens. J. 19(4), 1499–1507 (2018)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Zijian Zhou graduated from the Xiamen University Malaysia. He earned his bachelor’s degree in computer science and Technology from Xiamen University Malaysia. Mr. Zhou will start his master’s degree in Computer Science at Northeastern University in Boston in 2021.
Muhammad Adeel Asghar was born in Taxila, Pakistan. He received the B.S. degree in electrical engineering from HITEC University, Taxila, Pakistan, in 2013, and the M.S. degree in electrical engineering from HITEC University. He is currently pursuing the Ph.D. degree in telecommunication engineering with the University of Engineering and Technology Taxila, Pakistan. He was the Operation Manager with Orascom Communications, a Field Engineer with Ericsson, Pakistan, and a Design Engineer with RWR PK Ltd. He designed and developed a 32-channel wireless EEG system for research area and public interest. His current research interests include computer vision, machine learning, digital signal processing, and brain-computer interface.

Daniyal Nazir received the B.S. degree in software engineering from Mirpur University of Science and Technology, Mirpur AJK, Pakistan, in 2016. He obtained his M.S. in software engineering from University of Engineering and Technology Taxila, Pakistan. He is currently Lecturer with University of Azad Jammu and Kashmir, Pakistan. His research interests include image processing, machine learning, computer vision, pattern recognition, and neural networks.

Kamran Siddique (M’19) received the Ph.D. degree in computer engineering from Dongguk University, South Korea. He is currently an Associate Professor with Xiamen University Malaysia. His research interests include cybersecurity, machine learning, IoT, and big data processing.

Mohammad Shorfuzzaman is currently an Associate Professor with the Department of Computer Science, College of Computers and Information Technology (CCIT), Taif University, Taif, Saudi Arabia. He is also a member of the Big Data Analytics and Applications (BDAAG) research group, CCIT. His primary research interests include applied artificial intelligence in the areas of computer vision and natural language processing, big data, and cloud computing.

Raja Majid Mehmood (Senior Member IEEE) is an Assistant Professor in the School of Computing and Data Science, Information and Communication Technology Department, Xiamen University Malaysia, Sepang, Malaysia. Dr. Raja is an IEEE Senior Member. He served as a research professor in the Department of Brain and Cognitive Engineering, Korea University, Seoul, South Korea. He was Lecturer in the Software Engineering Department at King Saud University, Saudi Arabia. He obtained his Ph.D. in Computer Engineering from the Division of Computer Science and Engineering, Chonbuk National University, South Korea. He obtained his M.S. in Software Technology from Linnaeus University, Sweden. He received his B.S. in Computer Science from Gomal University, Pakistan. He has authored more than 50 research papers and is supervising or co-supervising several graduate (MS/PhD) students. His main research interests include affective computing, brain computer interfaces, information visualization, image processing, pattern recognition, multi-task scheduling, and software engineering. Dr. Raja has served as a reviewer of various international journals including IEEE Communications Magazine, IEEE Transactions On Affective Computing, Int. J. of Information Technology and Decision Making, IEEE Access, Multimedia Tools and Applications, ACM Transactions on Multimedia Computing Communications and Applications, Journal of King Saud University - Computer and Information Sciences, Biomedical Engineering (BME), IEEE Internet of Things Journal, IEEE Transactions on Systems, Man and Cybernetics and so on.
Authors and Affiliations

Zijian Zhou¹ · Muhammad Adeel Asghar² · Daniyal Nazir³ · Kamran Siddique¹,⁴ · Mohammad Shorfuzzaman⁵ · Raja Majid Mehmood¹

¹ Information and Communication Technology Department, School of Computing and Data Science, Xiamen University Malaysia, Sepang 43900, Malaysia
² Department of Computer Science, National University of Modern Languages, Hamayun Road, Rawalpindi 46300, Punjab, Pakistan
³ Department of Software Engineering, University of Azad Jammu and Kashmir, Muzaffarabad 13100, Pakistan
⁴ Department of Computer Science and Engineering, University of Alaska Anchorage, Anchorage 99508, USA
⁵ Department of Computer Science, College of Computers and Information Technology, Taif University, Taif 21944, Saudi Arabia

Raja Majid Mehmood
rmeex07@ieee.org; rajamajid@xmu.edu.my
Zijian Zhou
CST1609020@xmu.edu.my; zijianzhou2021@163.com
Muhammad Adeel Asghar
adeel.asghar@numl.edu.pk
Daniyal Nazir
daniyalnazir1@gmail.com; daniyal.nazir@ajku.edu.pk
Kamran Siddique
kamran.siddique@xmu.edu.my
Mohammad Shorfuzzaman
m.shorf@tu.edu.sa