V10-2 (2023) 54-65
International Journal of Computers and Information
(IJCI)
Available online at *https://ijci.journals.ekb.eg/*

Machine Learning Algorithms for Enhancing Emotion Recognition from EEG Signals

Aseel Mahmoud[1] , *Khalid Amin*[2] and *Mina Ibrahim*[3]
Department of Information Technology
Faculty of Computers and Information, Menoufia University, Egypt
Email: aseel.mahmoud13@gmail.com[1], kh.amin.0.0@gmail.com[2],
Mina.ibrahim@ci.menofia.edu.eg[3]

Abstract

Emotion recognition through electroencephalography (EEG) signals is an important aspect of human-computer interaction that poses a significant research challenge. Most of the current approaches utilize up to 18 channels from 32 available channels for extracting emotions features. Moreover, they only use valence and arousal model to classify emotions. Therefore, the current approaches are unable to detect emotions accurately. In this paper, a framework that utilizes a three-dimensional model incorporating arousal, valence, and dominance for identifying emotions is proposed. Our framework can define any number of emotions, even in the absence of discrete emotions labels. The electroencephalography signals from DEAP database are utilized for emotion detection. The effectiveness of three classification techniques is examined, namely Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Multilayer Perceptron (MLP). The classification accuracies for valence, arousal, and dominance are 90.19%, 91.91%, and 89.86%, respectively. Our results demonstrate that EEG data's time-domain statistical and power features can effectively classify different emotional states. Furthermore, our framework enables accurate identification of identical emotions that cannot be distinguished by a two-dimensional model.

Keywords: Emotion recognition; Machine learning; Classification; Valence-Arousal-Dominance model

1. Introduction

Emotions are intense mental states that arise automatically in the nervous system, rather than through conscious effort **[1]**. They play a pivotal role in interaction between human and computer, necessitating the building of modern systems that can better understand and react to human feelings **[2]**. To this end, various algorithms have been developed that leverage physiological signals, facial expressions **[3]**, tone of voice **[4]**, and even electrical signals from the brain **[5]** to detect emotions. These advancements in technology enable computers to better recognize and respond to human emotions, improving their ability to interact with people in a more natural and intuitive manner.

The brain is an enigmatic and remarkable organ responsible for all thoughts, beliefs, memories, behaviors, and moods. It acts as the control center for the entire body, making it a vital component of human existence. Emotions are thought to be connected to the activity in specific areas of the brain that influence our attention, and their effects can be detected in the brain's electrical signals, as measured by EEG **[6]**. The use of EEG emotion recognition has the potential to enhance the human-computer interface and can be applied across various industries, including marketing, gaming, and E-learning. Several studies have attempted to differentiate between emotions using electrical brain signals, each utilizing a different approach to elicit emotions **[7] [8]**. These studies typically vary in emotions that can be detected, the quantity of participants involved in the study to form a database, the extracted features for classification, and the number and location of electrodes placed on the participant's head **[6]**.

An emotional model can be either discrete or dimensional. Emotions are defined as a collection of fundamental emotions in the discrete model **[9]**. On the other hand, the dimensional model views emotions as a combination of multiple dimensions, including arousal and valence **[10]**. Valence represents the degree to which an emotion is pleasurable, ranging from undesirable to very pleasant, where undesirable refers to low valence and very pleasant refers to high valence. Arousal measures the level of emotional intensity, ranging from complete calm to very excited, where complete calm refers to low arousal and very excited refers to high arousal. Some models that rely solely on the ratio of arousal and valence may not be able to distinguish between sensations that share similar arousal and valence values. For example, research suggests that feelings of fear and anger share high levels of attention and negative valence **[11] [12]**. To address this issue, a third dimension called dominance has been added to the two-dimensional model. This addition creates a complete model where a group of emotions cannot be similar in all three dimensions **[13]**. Dominance can be defined as the level of control or power exerted by the emotion **[11]**, providing a comprehensive representation of emotions.

Although several earlier studies have suggested some positive results utilizing machine learning models, the effectiveness of emotion recognition is highly dependent on each subject's impulsive responses to the stimuli. Moreover, the number and locations of the channels that were chosen is very critical. Additionally, EEG signals are not universal, and distinct EEG channels may be used to analyze the same emotion in two different people. Therefore, it is essential to strive for a system that has the capability of accurately and unerringly differentiating emotions. Modifying the data without destroying crucial channel properties is also challenging.

In this paper, emotion recognition framework that utilizes Valence, Arousal, and Dominance (VAD) dimensions to define the emotional states is proposed. The proposed framework starts with a data acquisition process from a standard dataset called DEAP dataset. After that, these data are preprocessed, and DWT divides each time window into its component frequency bands. Then, each sub-bands spectral features are extracted. Finally, the classification algorithms are utilized to classify emotions. Three classification algorithms are used, which are SVM, KNN, and MLP. The experimental results showed that the SVM algorithm achieves the highest accuracy with 90.19% for valence, 91.91% for arousal, and 89.86% for dominance.

The paper is structured as follows: Section 2 presents related works. Section 3 describes the proposed methods. The results and discussion are presented in Section 4. Finally, Section 5 concludes the paper and presents future work.

2. **Related Work**

Many investigators are interested in the topic of emotion recognition, and there are multiple studies on it, as it is not limited to electrical brain signals only. In our research, our attention will focus on recognizing feelings through EEG signals. There is a close relationship between the electrical brain signals and the human feeling [14]. With exception of the known concept that heart is the state of feelings, the human brain is where emotions first develop, and the EEG has the potential to track and evaluate a person's present level of mental activity [15].

Wu et al. [16] proposed an algorithm that used $Fp1$ and $Fp2$ channels of frontal EEG signals based on the idea of the frontal brain. Spatial and frequency features are presented. A GBDT classifier was assessed and chosen, as well as the method was then applied on the DEAP database in order to determine how effective it was. The mean classification accuracy reached 75.18%, and the maximum accuracy reached 76.34%. There are few studies on emotion recognition which depend on the principle of frontal two channels, and the results from those studies are inconsistent and pessimistic because of the lack of data according to limited number of channels.

Mohammadi et al. [17] improved the results which were applied on the DEAP database. EEG signals have been split to relevant bands based on frequency using a transformation method called discrete wavelet transform, and subsequently a number of features have been retrieved. SVM and KNN classifiers were used. The classification accuracy results were 84.05% for valence and 86.75% for arousal over the recorded 10-channel EEG. The disadvantage of utilizing a 2D model instead of a 3D model is that it can result in inaccurate predictions of emotions in real-time applications. Tong et al. [18] proposed EEG features for emotion recognition include the multi feature-fusion in a time domain and the composite features based on wavelet feature and information entropy. The DEAP dataset was used to apply the proposal. In terms of valence arousal, the average accuracy was 72.03% and 71.7% respectively. They chose the channels using the relief method. To cut down on the amount of data, they utilized 6 and 13 channels. From my point of view, they worked efficiently on decreasing time which led to lose a significant amount of data, as a result, the accuracy decreased.

Nawaz et al. [19] used VAD model to identify the emotions elicited by selected music videos. DEAP dataset was used in data acquisition phase. The authors proposed an algorithm that used only 14 channels. A number of features were elicited from EEG signals such as (statistical, entropy and power). SVM, KNN, and decision tree algorithms were utilized to validate features. For valence, arousal, and dominance, the classification accuracies were 77.62%, 78.96%, and 77.60% respectively. The results were not enough to be applied in emotion recognition applications.

Gao et al. [20] proposed a novel multi-feature fusion network composed of spatial and temporal neural network structures capable of learning discriminative spatial temporal emotional information and recognizing emotion. They used 32 channels of 10-20 system. Additionally, the time domain (e.g., sample and differential entropy) and frequency domain (e.g., Hjorth and power spectral density) features were used. The SVM classifier was used to classify features, and the average emotion recognition accuracies in the DEAP database's valence and arousal classification tasks were 80.52% and 75.22%, respectively. 2D emotion model still can’t determine the emotions obviously.

Bazgir et al. [21] used valence-arousal model to classify four emotional states based on high or low. DEAP dataset was used to apply the framework. They used only these channels F3-F4, F7-F8, FC1-FC2, FC5- FC6, and FP1- FP2 based on 10-20 system. Wavelet transform was applied to EEG signals during the feature extraction phase, along with energy and entropy for the extracted frequency bands. Additionally, they used the PCA algorithm for feature selection. The classification algorithms used in the proposal were SVM, KNN, and ANN. The accuracy for valence was 91.1% and for arousal was 91.3%. The authors only utilized 2D model and employed the frontal 10 channels. Therefore, they were only able to classify emotions as either negative or positive.

From the discussed related works, the previous approaches suffer from the following limitations:

1. Most of the previous approaches only utilize 18 channels from 32 channels.
2. Most of the previous approaches only use two-dimensional model.
3. The obtained accuracy is not good enough to be used in critical applications.

Through the utilizing all 32 channels using the 10-20 method and extracting the most significant characteristics of each channel, the first issue is solved, which significantly improved the results. In order to solve the issue of improper classification of feelings and learn how to overcome it in terms of effectively identifying emotions without interaction, the two-dimensional model is changed to a three-dimensional one. The achieved accuracy is insufficient for crucial applications, which is why our proposal focuses on improving accuracy as much as possible.

In the next section, the proposed framework to tackle these limitations will be discussed.

3. Proposed Work

In this section, the proposed framework will be discussed in detail. Fig.1 shows the main processes in the proposed framework, which are: 1) Data acquisition, which presents the acquisition of EEG data through using DEAP dataset; 2) Data preprocessing, which includes required processing to adapt data for next processes; 3) Feature extraction, which is used to identify prominent characteristics from the pre-processed data; and finally 4) classification, which is used to classify the emotions.

3.1. EEG data acquisition

Instead of using multiple relatively small datasets, DEAP EEG dataset is used for allowing us to qualitatively assess the derived features. The DEAP EEG dataset is pre-processed publicly available dataset, which includes the EEG and peripheral signals of 32 human individuals [22]. For emotional stimuli, 40 video clips were used, and their duration does not exceed one minute. After watching each clip, they were rated from 32 people on the three criteria arousal, dominance, and valence. The level of liking and intimacy using a questionnaire ranging from 1 to 9.

The labels of VAD dimensions are based on the self-assessment rates of the employed subjects. Valence dimension, which ranges from miserable to happy, is a measure of pleasantness. From the least thrilled to the most excited, arousal dimension rates the intensity of the emotions. From complete control to complete lack of control, dominance dimension quantifies the level of emotional regulation.

3.2. Data preprocessing

In line with what was discussed in the previous section, each of the 32 subjects had 40 EEG recordings, each measuring one minute. As a result of mood swings between the videos, the first 20 seconds of the video were cropped out since it might have been influenced by the previous video. As a result, the sensations that were recognized are erroneous. While employing electrical brain signals to determine feelings, the width of the temporal window is a crucial decision that must be made very carefully. There will be 40 seconds remaining in the time clip once the 20 seconds have been subtracted. Then divided it into time windows of 10 seconds each, and since there are four windows in each clip, the total number of video segments for each person is 40 * 4, equal 160. The EEG signals are first inserted, and for the 32 individuals, 32 channels are then extracted from the selected EEG channels that are positioned in Fig.2 according to 10-20 international standards [22]. Each channel receives 40 videos in return, with processing continuing up to channel 32 beginning from the first channel.

3.3 Feature extraction

The primary goal of feature extraction stage is to detect significant attributes from the employed data which can accurately associate EEG segments with their corresponding emotional states. To assess the performance of different aspects in classifying emotional states, multiple features were extracted from the data. For EEG-based emotion recognition, all of the retrieved features, which are described in more detail below, have been widely employed [23].

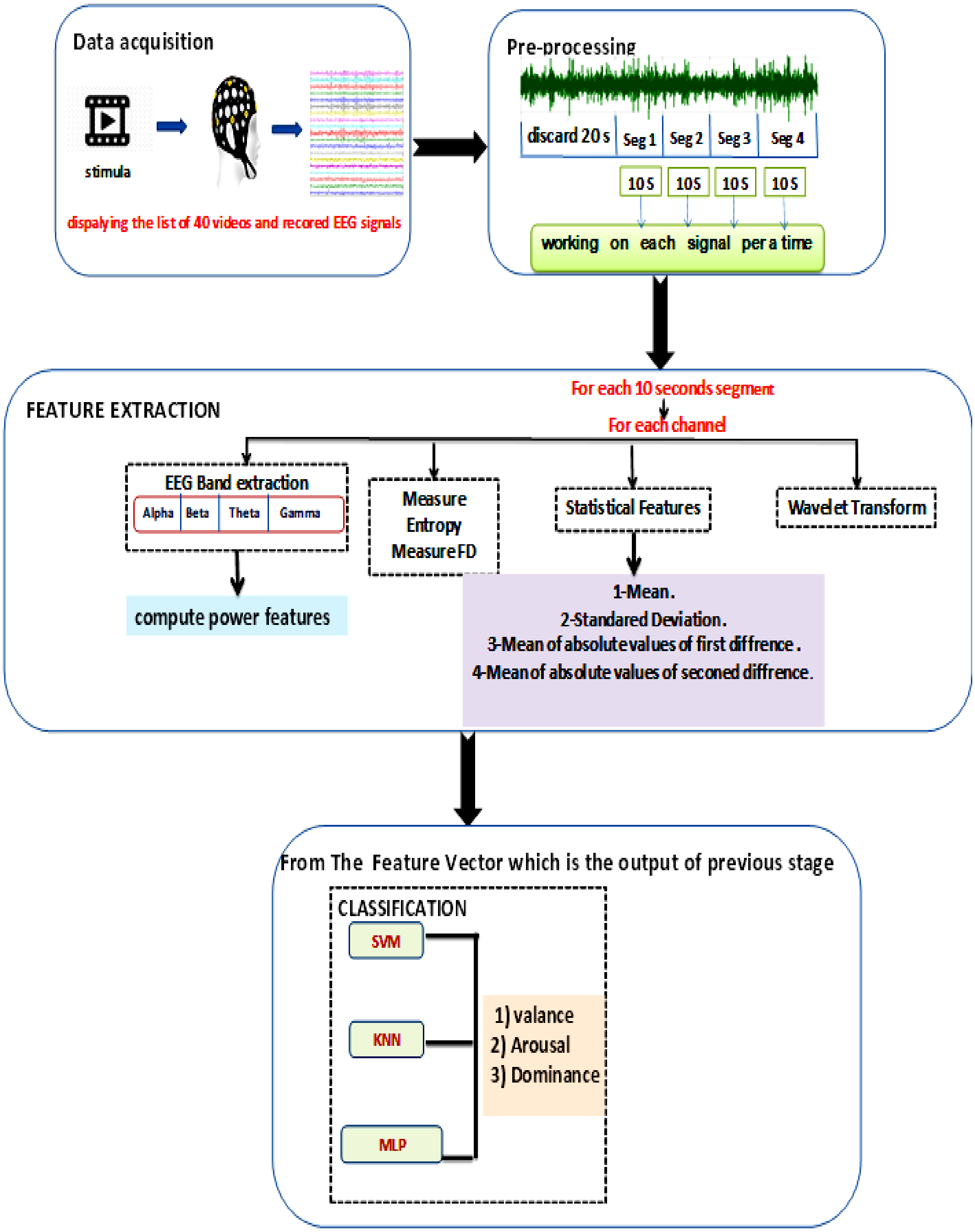

Fig. 1 The proposed framework

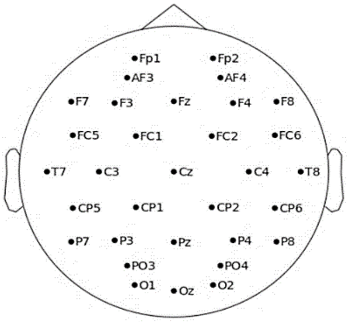

Fig. 2 The 10-20 electrode distribution (Black circles)

3.3.1. Power features

The EEG signal is consisting of frequency bands such as theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and gamma band (above 30 Hz). Once pre-processing is complete, Welch technique [19] is used to calculate the power spectrum feature. This involves dividing the time-series signal into sequential blocks and constructing a period gram for each block, after which the Welch technique is used to estimate the power spectra. The blocks are then averaged to calculate the mean. Fig.3 demonstrates the power spectra plot for 32 channels.

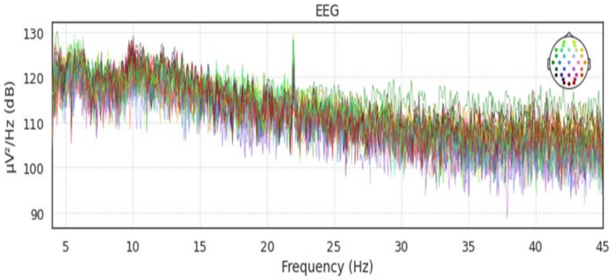

Fig. 3 The power spectral density (Feature) for 32 channels

Equation (1) shows how to split the time-series signal into N blocks.

$$x = \{x_1, x_2, \dots, x_N\} \quad (1)$$

Equation (2) shows how to create each block's period gram (P_{x_n}) [19]. The L represents the total number of points in the n^{th} block.

$$P_{x_n} = \frac{1}{L} |fftx_n|^2 \quad (2)$$

Equation (3) shows how to calculate the power spectral density ($\hat{s}$) estimated by Welch technique [19]. The k represents the total number of blocks in signal after Welch estimated.

$$\hat{s} \triangleq \frac{1}{k} \sum_{m=0}^{k-1} p_{x_n}(m) \quad (3)$$

The energy feature was computed for the alpha, beta, theta and gamma bands and the delta band is not calculated (slow and fast delta waves) which were detected in the range (0–2 Hz and 2–4 Hz respectively). Delta band detect deep sleep status which is not vital in our research that’s why the DEAP preprocessed signals filtered between 4 and 45 Hz. The power of four bands is calculated on each channel to create a power feature vector of 128 dimensions (4 (bands) * 32 (channels) = 128). Effective EEG power features have been discovered for recognizing emotional changes following music stimulation [24].

3.3.2. Entropy

The level of uncertainty or unpredictability in the pattern, which is also roughly equivalent to the quantity of data contained in the signal, can be measured statistically by applying the notion of entropy to time series [19] like electroencephalography as shown in Equation (4). Where psd is the normalized power spectral density and f_n is the half of the sampling frequency based on Nyquist theory.

$$E = \sum_{f=0}^{f_n} psd(f) \log_2 psd(f) \tag{4}$$

3.3.3. Statistical features

A sequence of signal statistics from time-series EEG data is used to recognize emotion [1]. Two statistical features (Mean and Standard deviation) are taken from [25] and applied in the current investigation. The time-series EEG signals have been described using Mean (x) defined in Equation (5) and Standard deviation (σ_x) defined in Equation (6). Here, N denotes the total number of samples in the EEG signal and $X(n)$ represents the EEG signal that has been normalized to have zero mean and unit variance.

$$\text{Mean } (x) = \frac{1}{N} \sum_{n=1}^{N} X(n) \tag{5}$$

$$\sigma_x = \sqrt{\frac{1}{N} \sum_{n=1}^{N} \frac{1}{N} (X(n) - \mu_{x^1})^2} \tag{6}$$

3.3.4. Wavelet feature

One of the strongest characteristics of EEG signals used in emotional recognition is the wavelet feature in the time-frequency domain [26]. Each channel is analyzed in the alpha, beta, theta, and gamma range. As mentioned before, the DEAP dataset was filtered in the 4-45 Hz range, excluding the delta band because it detects deep sleep, which appears to lack active emotions. Sampling rate of 128 samples/sec is used and implemented a four-level approach to extract the EEG band coefficients as follows:

1. D1 for Gamma band (32–64 Hz).
2. D2 for Beta band (16–32 Hz)
3. D3 for Alpha band (8-16 Hz)
4. D4 for Theta band (4-8 Hz)

The wavelet energy E_j is calculated for each of the 32 channels of four bands, as shown in Equation (7), the result is 128-dimensional (4 bands *32 channels) wavelet energy feature vector (E_j). The D_j is for the detail coefficients of the wavelet decomposition's j^{th} level, which refers to EEG band j, and k refers to the total number of wavelet coefficients.

$$E_j = \sum_{k=1}^{N} (D_{j(k)^2}) \tag{7}$$

3.4. Classification

In classification stage large number of classifiers are tested such as, Naïve Bayes, Decision Tree, Regression, KNN (with different k number), SVM (with RBF, linear and polynomial kernels) and MLP. Based on the encouraging empirical results of SVM, KNN, and MLP classifiers in the EEG emotion identification [26], each classifier applied separately to get the higher result depending on the accuracy metric. The SVM classifier is used to assess the RBF, linear, and poly kernels [27]. The RBF kernel is utilized because by comparing the results of different kernels such as linear, polynomial and RBF, the better accuracy result is with RBF kernel in our work. Between the two distinct classes, SVM creates a separate hyper plane and aims to optimize the separation between each class and the separated hyper plane.

On the other hand, the KNN algorithm is very reliable in classifying EEG data [28]. The KNN methodology depends on searching for a specific number of k samples that are closest to the prediction sample. Each sample selects for the closest category. The prediction sample is appointed to the most voted category, as for the MLP classifier, it also achieved high results.

4. Results and Discussion

In our experiments, the DEAP dataset's 32 subjects is used in data acquisition stage, which is discussed in Section 3. The data from each subject was clipped into a 10-second time window in which each should be categorized into one labeled emotion. The data is divided based on: valence, arousal, and dominance. 32 EEG electrodes (channels) are taken into account. These channels' positions according to the 10-20 System is FP1,FC1,P7, $FC5$, $F4$, $F7$, $AF3$, $CP6$, $T7$, $C3$, $FC6$, $P4$, $Fp2$, $F8$, $P3$, $CP5$, $O1$, $F3$, $P8$, $CP2$, $CP1$, $PO4$, $O2$, Pz, Oz, $T8$, FC2, Fz, $AF4$, $PO3$, Cz and $C4$. In order to validate and test the proposed framework, three classification methods are adopted, which are SVM, KNN, and MLP.

Each participant has 160 (40 videos * 4 clips = 160) observations total. Cross-validation approach with $k = 5$ is used to split the subject's data into training and testing [29]. The data is separated into five equal pieces using this 5-fold cross validation approach. The Algorithm is trained for each individual using the initial four segments consisting of 128 observations, while the fifth segment, consisting of 32 observations, is utilized as test data. This process is iterated five times iteratively, with the first four segments being utilized for training and the final segment being utilized for testing on each occasion.

To evaluate the effectiveness of the applied approaches, the classification accuracy metric is used. The classification accuracy is calculated using Equation (8) where TP is True positive; FP is False positive; TN is True negative; FN is False negative.

$$classification\ accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{8}$$

Fig.4 displays the labels for 32 participants' responses to 40 videos. This graph shows the rate from 1 to 9 over the y-axis and the videos over the x-axis. Blue, red, and orange colors can be used to distinguish the three categories of VAD model.

The classification performance results of SVM, KNN, and MLP algorithms through different frequency bands are presented and compared in Table 1. From the table, the following are found:

- For valence dimension, our experiment investigates the correlation for electing valence emotions with theta and gamma bands, which means that the valence emotions reflect their impact in this region of brain. For example, the SVM achieves 78.23% and 80.2% for alpha and beta respectively, while it achieves 91.19% and 90.65% for theta and gamma respectively.
- For arousal dimension, our experiment investigates the correlation for electing arousal emotions with alpha, beta and gamma bands which means that the arousal emotions reflect their impact in this region of brain. For example, the KNN achieves 80.45%, 89.88% and 88.14% for alpha, beta and gamma respectively, while it achieves 74.1% for theta.
- For dominance dimension, gamma band and dominance levels are positively correlated. It means that electing dominance emotions will result in an increase the gamma band. For example, the MLP achieves 79.48% for gamma, while it achieves 59% for alpha, 57.98% beta and 78.23% for theta.

- The best performance is given by SVM classifier which exhibits an average accuracy value of 91% (the average accuracy is computed by calculating the summation of VAD and divide it on 3)

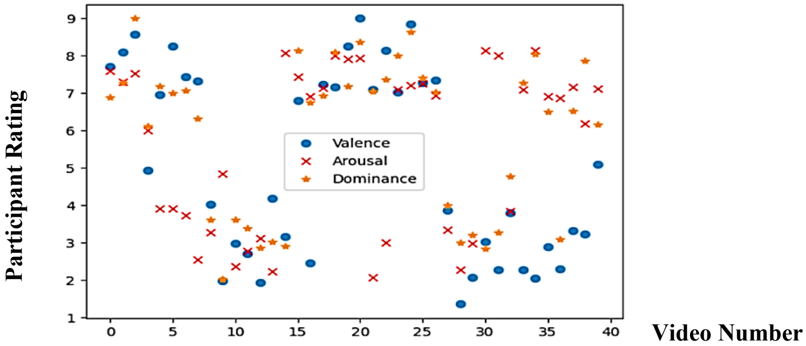

Fig. 4 The labels for 32 participants' responses to 40 videos.

Table 1 The classification performance results of SVM, KNN, and MLP through different frequency bands.

Classifier	Models	Frequency Bands			
		Alpha	Beta	Theta	Gamma
SVM	Valence	78.23	80.2	91.19	90.65
	Arousal	90.95	88.65	68.21	91.91
	Dominance	70.54	75.8	83.25	89.86
KNN	Valence	74.23	85.47	90.02	90.02
	Arousal	80.45	89.88	74.1	88.14
	Dominance	71.2	78.1	89.54	89.0
MLP	Valence	59.0	60.2	78.58	78.99
	Arousal	62.99	60.32	60.0	63.99
	Dominance	59.0	57.98	78.23	79.48

When compared to prior research that validated their techniques using the DEAP dataset, the classification rate in the current study is compelling. Our approach was compared to recent studies on emotion recognition that also employed the DEAP dataset for this purpose, and an overall comparison is presented in Table 2. Wu et al. [16] proposed an algorithm that only used $Fp1$ and $Fp2$ of frontal EEG signals. A GBDT classifier was assessed and chosen. The maximum accuracy reached 76.34%. Mohammadi et al. [17] used SVM and KNN classifiers. Results showed classification accuracy of 86.75% for arousal level and 84.05% for valence. Tong et al. [18] proposed EEG features for emotion recognition include the multi feature-fusion in a time domain and the composite features. In terms of valence arousal, the average accuracy was 72.03% and 71.7% respectively. They reduced the selected number of channels to 6 and 13 channels. Nawaz et al. [19] used a VAD model of emotion to recognize the emotions elicited by music videos. SVM, KNN, and decision tree classifiers are used to validate features. For valence, arousal, and dominance, the classification accuracies were 77.62%, 78.96%, and 77.60% respectively. In another study Gao et al. [20] proposed a novel multi-feature fusion network composed of spatial and temporal neural network structures capable of learning discriminative spatial temporal emotional information and recognizing emotion. The SVM plus CNN classifier were used to classify features, and the average emotion recognition accuracies in the DEAP database's valence and arousal classification tasks were 80.52% and 75.22% respectively. Omid Bazgir et al. [21] used 2D model of emotion to recognize the four emotional states which are: high arousal high valence, high arousal low valence, low arousal high valence, and low arousal low valence. DWT, entropy and energy were used as a features, then PCA applied to reduce the number of features. SVM, KNN, and ANN were the algorithms used in classification. Results were 91.1% for valence and 91.3% for arousal. In this work, for each EEG segment (10 seconds) alpha, beta, theta, and gamma bands are extracted which achieve best accuracy of 90.19% for arousal, and 91.91% for valence. Our proposed method achieved average accuracy of 90.19%, 91.91%, and 89.86% for valance, arousal and dominance respectively.

Table 2 Comparison of the precision of several research projects utilizing DEAP data

Reference	Model	No. Channels	Features	Classifier	Best Accuracy
[16]	Aalence Arousal	2	FFT Wavelet transform	GBDT	Valence= 76.34% Arousal= 75.18%
[17]	Valence Arousal	10	DWT	SVM KNN	The best accuracy is achieved by KNN Valence= 84.05% Arousal= 86.75%
[18]	Valence Arousal	13	DWT Statistical	SVM	Valence= 70.33% Arousal= 71.7%
[19]	Valence Arousal Dominance	14	Power features Statistical Fractal dimension Wavelet feature	DT SVM KNN	The best accuracy is achieved by SVM Valence= 77.62% Arousal= 78.96% Dominance = 77.60%
[20]	Valence Arousal	32	PSD SE DE Hjorth	SVM+ CNN	Valence= 80.25% Arousal= 75.22%
[21]	Valence Arousal	10	DWT Statistical	SVM KNN ANN	Valence= 91.1% Arousal= 91.3%
Proposed Framework	Valence Arousal Dominance	32	Power features Statistical Wavelet feature	SVM KNN MLP	The best accuracy is achieved by SVM Valence= 90.19% Arousal= 91.91% Dominance = 89.86%

5. Conclusion

The number and categorization of fundamental human emotions remains a topic of debate among researchers. Different scholars have defined distinct groups of primary emotions according to their own interpretation, leading to variations in the total number of emotions. For instance, some researchers have categorized anger, fear, helplessness, and contempt as four of the fourteen primary emotions, while others have identified around twenty-seven additional emotions. Various systems have been proposed for extracting information from brain signals, each with their own procedures, and the performance of these systems is assessed based on their ability to accurately classify emotional states in a given database. This paper introduces an EEG-based framework for recognizing emotions. The EEG signal is analyzed to derive a diverse set of properties such as power spectral density (power features), entropy, statistical, and wavelet features. Three different machine learning classifiers, SVM, KNN, and MLP, were applied by comparing the feature extraction techniques. The SVM algorithm was utilized to achieve classification accuracies of 90.19%, 91.91%, and 89.86% for VAD respectively. To fully reap the benefits of emotion recognition in real-world settings, it is crucial to ensure higher levels of precision and real-time compatibility. This necessitates further research and exploration of deep network optimization, significant feature extraction, and channel reduction in the future work.

References

[1] R. Jenke, A. Peer, and M. Buss, "Feature extraction and selection for emotion recognition from EEG," *IEEE Transactions on Affective Computing*, vol. 5, no. 3, pp. 327-339, 2014.

[2] K.Schaaff and T.Schultz, "Towards emotion recognition from electroencephalographic signals," *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, pp. 1-6, 2009.

[3] M.Soleymani, S.Asghari-Esfeden, Y.Fu, and M.Pantic, "Analysis of EEG Signals and Facial Expressions for Continuous Emotion Detection," *IEEE Transactions on Affective Computing*, vol. 7, pp. 17-28, 2016.

[4] N.Ho, H.Yang, S.Kim, and G.Lee, "Multimodal Approach of Speech Emotion Recognition Using Multi-Level Multi-Head Fusion Attention-Based Recurrent Neural Network," *IEEE Access*, vol. 8, pp. 61672-61686, 2020.

[5] J.Brosschot and J.Thayer, "Heart rate response is longer after negative emotions than after positive emotions," *International Journal of Psychophysiology*, vol. 50, pp. 181-187, 2003.

[6] A.Al-Nafjan, M.Ibrahim Hosny, Y.Al-Ohali, and A.Al-Wabil, "Review and Classification of Emotion Recognition Based on EEG Brain-Computer Interface System Research: A Systematic Review," *Applied Sciences*, vol. 7, p. 1239, 2017.

[7] M.Black and Y.Yacoob, "Recognizing Facial Expressions in Image Sequences Using Local Parameterized Models of Image Motion," *International Journal of Computer Vision*, vol. 25, pp. 23-48, October 1997.

[8] Z.Wang, J.Zhang, Y.He, and J.Zhang, "EEG emotion recognition using multichannel weighted multiscale permutation entropy," *Applied Intelligence*, vol. 52, pp. 12064–12076, 2022.

[9] J.Currie, "Music After All," *Journal of the American Musicological Society*, vol. 62, pp. 145-203, April 2009.

[10] A.Mehrabian, "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in Temperament," *Current Psychology*, vol. 14, pp. 261-292, 1996.

[11] Y.Liu and O.Sourina, "EEG-based Dominance Level Recognition for Emotion-Enabled Interaction," *2012 IEEE International Conference on Multimedia and Expo*, pp. 1039-1044, 2012.

[12] J.Li et al., "Cross-subject EEG emotion recognition combined with connectivity features and meta-transfer learning," *Computers in Biology and Medicine*, vol. 145, p. 105519, 2022.

[13] H.Joong Yoon and S.Youb Chung, "EEG-based emotion estimation using Bayesian weighted-log-posterior function and perceptron convergence algorithm," *Computers in Biology and Medicine*, vol. 43, pp. 2230-2237, 2013.

[14] A.Mehmood Bhatti, M.Majid, S.Anwar, and B.Khan, "Human emotion recognition and analysis in response to audio music using brain signals," *Comput. Hum. Behav.*, vol. 65, pp. 267-275, 2016.

[15] O.Pollatos, W.Kirsch, and R.Schandry, "On the relationship between interoceptive awareness, emotional experience, and brain processes," *Cognitive Brain Research*, vol. 25, pp. 948-962, 2005.

[16] S.Wu, X.Xu, L.Shu, and B.Hu, "Estimation of valence of emotion using two frontal EEG channels," in *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2017.

[17] Z.Mohammadi, J.Frounchi, and M.Amiri, "Wavelet-based emotion recognition system using EEG signal," *Neural Computing and Applications*, vol. 28, pp. 1985-1990, 2017.

[18] L.Tong, J.Zhao, and W.Fu, "Emotion Recognition and Channel Selection Based on EEG Signal," in *2018 11th International Conference on Intelligent Computation Technology and Automation (ICICTA)*, pp. 101-105, 2018.

[19] R.Nawaz, K.Hwa Cheah, H.Nisar, and V.Yap, *Biocybernetics and Biomedical Engineering*, vol. 40, pp. 910-926, 2020.

[20] Q.Gao, Y.Yang, Q.Kang, Z.Tian, and Y.Song, "EEG-based Emotion Recognition with Feature Fusion Networks," *International Journal of Machine Learning and Cybernetics*, vol. 13, February 2022.

[21] O.Bazgir, Z.Mohammadi, and S.Habibi, Emotion Recognition with Machine Learning Using EEG Signals, March 2019.

[22] S.Koelstra et al., "DEAP: A Database for Emotion Analysis ;Using Physiological Signals," *IEEE Transactions on Affective Computing*, vol. 3, pp. 18-31, 2012.

[23] R.W. Picard, E. Vyzas, and J.Healey, "Toward Machine Emotional Intelligence: Analysis of Affective Physiological State," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, pp. 1175-1191, 2001.

[24] R.Nawaz, H.Nisar, and Y.Voon, "The Effect of Music on Human Brain; Frequency Domain and Time Series Analysis Using Electroencephalogram," *IEEE Access*, vol. 6, pp. 45191-45205, 2018.

[25] Z.Lan, O.Sourina, L.Wang, and Y.Liu, "Real-time EEG-based emotion monitoring using stable features," *The Visual Computer*, vol. 32, pp. 347-358, 2015.

[26] X.Wang, D.Nie, and B.Lu, "Emotional state classification from EEG data using machine learning approach," *Neurocomputing*, vol. 129, pp. 94-106, 2014.

[27] C.Chang and C.Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, pp. 27:1-27:27, 2011.

[28] A.Adebola Yusuf, S.Wijaya, and P.Prajitno, "EEG-based human emotion recognition using k-NN machine learning," *PROCEEDINGS OF THE 4TH INTERNATIONAL SYMPOSIUM ON CURRENT PROGRESS IN MATHEMATICS AND SCIENCES (ISCPMS2018)*, 2019.

[29] P.Refaeilzadeh, L.Tang, and H.Liu, "Cross-Validation," in *Encyclopedia of Database Systems*, L. I. U. LING and Ö. Z. S. U. M. TAMER, Eds. Boston, MA: Springer US, 2009, pp. 532–538.