EEG Emotion Recognition based on Hierarchy Graph Convolution Network

Fa Zheng
School of Information Science and Engineering
Shandong Normal University
Jinan, China
zfsdnu305@qq.com

Bin Hu*
School of Information Science and Engineering
Shandong Normal University
Jinan, China
binhu@sdnu.edu.cn

Shilin Zhang
School of Information Science and Engineering
Shandong Normal University
Jinan, China
zhang_shilin_sd@163.com

Yalin Li
Key Laboratory of TCM Data Cloud Service in Universities of Shandong
Shandong Management University
Jinan, China
yalinli2012@163.com

Xiangwei Zheng*
School of Information Science and Engineering
Shandong Normal University
Jinan, China
xwzhengcn@163.com
*corresponding authors

2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) | 978-1-6654-0126-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/BIBM52615.2021.9669465

***Abstract*—Emotion recognition has become a research focus in the field of human-computer interaction (HCI). As an excellent physiological signal, electroencephalographic (EEG) is considered to be a favorable tool for emotion recognition. Most traditional methods focus on extracting features in time domain and frequency domain but the adjacent information and asymmetric information from adjacent and asymmetric channels are often ignored. Although several graph neural network (GNN) models are utilized to learn EEG features, most of the emotion recognition studies of GNN ignore the information existing between adjacent electrodes. In this paper, we propose an EEG emotion recognition method based on hierarchy graph convolution network (HGCN) named ERHGCN. Firstly, six different features including power spectral density (PSD), differential entropy (DE), differential asymmetry (DASM), rational asymmetry (RASM), asymmetry (ASM) and differential caudality (DCAU) from five frequency bands are extracted. Secondly, to improve graph convolution network (GCN) shortcoming of only extracting time and frequency features, HGCN is applied to extract deeper spatial feature by treating the longitudinal and transverse adjacent electrode pairs in different ways. Finally, six extracted features are fed into the HGCN model, then all features are integrated by two full connection layers. We conducted extensive experiments on DEAP dataset and experimental results show that the proposed method can obtain 90.56% and 88.79% recognition accuracies for valence and arousal classification tasks.**

***Index Terms*—EEG, Emotion recognition, HGCN, Adjacent feature, Asymmetric feature.**

I. Introduction

Emotion recognition is widely used in many fields, such as rehabilitation treatment, human-computer interaction (HCI) and depression detection [1]. Electroencephalographic (EEG) is a signal that records the surface activity of the cerebral cortex, which is the result of synaptic activation of brain neurons. With its merits of efficiency, portability and non-trauma, EEG has become a more effective tool for recognizing emotions, and has acquired decent results in emotion recognition [2].

In recent years, the response of brain to emotional events has been actively studied. Generally, EEG signals are represented by features, which can be categorized in time domain, frequency domain and time-frequency domain according to different types. Zheng et al. used event-related potential (ERP) components and modified multi-scale sample entropy (MMSE) for emotion recognition [3]. However, the spatial information from adjacent and symmetric channels is often ignored. Song et al. extracted the features from the relationship between electrodes position by graph convolutional neural network (CNN) [4]. Zheng et al. extracted the power spectral density (PSD), differential entropy (DE), differential asymmetry (DASM), rational asymmetry (RASM), asymmetry (ASM) and differential caudality (DCAU) features, the experimental results on SEED dataset indicate that the lateral temporal areas are activated more obviously for positive emotion than negative one in beta and gamma bands [5].

At present, deep learning has been widely used in emotion recognition. Zheng et al. proposed an emotion recognition method based on three-dimensional (3D) feature maps and CNNs, the average accuracy for valence and arousal was 93.61% and 93.04% on dataset DEAP [6]. Yin et al. proposed an EEG emotion recognition method using fusion model of graph convolutional neural networks (GCNN) and Long Short-Term Memory (LSTM), they attained the average classification accuracy of 90.45% and 90.60% for valence and arousal in subject-dependent experiments while 84.81% and 85.27% in subject-independent experiments [7]. However, for the special physiological data of EEG, there are still some differences between electrode channels, which will affect the classification effect. In this paper, we propose an EEG emotion recognition method based on hierarchy graph convolution network (HGC-

Authorized licensed use limited to: Universita degli Studi di Salerno. Downloaded on May 04,2023 at 07:54:35 UTC from IEEE Xplore. Restrictions apply.

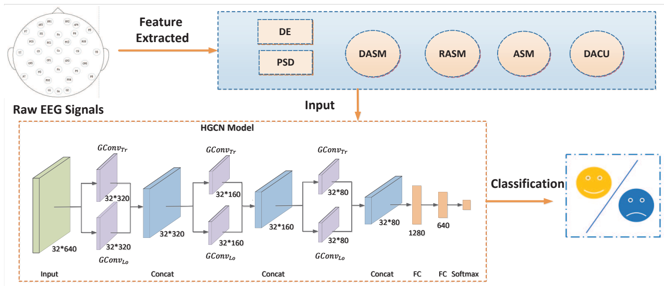

Fig. 1: The framework of the proposed method.

N) named ERHGCN.

II. Materials and Methods

A. Method overview

Fig.1 illustrates the process of the proposed EEG emotion recognition based on HGCN named ERHGCN. The proposed method mainly includes the following steps. Firstly, as we all known, the power of EEG signal in the frequency domain is one of the most popular EEG features for emotion analysis. We extract the following six different features on five bands (delta (1-3Hz), theta (4-7Hz), alpha (8-13Hz), beta (14-30Hz) and gamma (31-50Hz), namely PSD, DE, DASM, RASM, ASM and DCAU. Secondly, the extracted features are taken as the input of HGCN, which is proposed to extract the deeper spatial information by treating the longitudinal and transverse adjacent electrode pairs in different ways. HGCN model is applied to propagate the features between adjacent electrode pairs. Finally, two fully connected layers are used to summarize the features of all electrode pairs, and softmax layer is applied to the classification as a classifier.

B. Feature extraction

Motivated by the neurological discoveries about the regional and asymmetric properties of EEG, we extract adjacent features and asymmetric features in parallel based on the acquired frequency features. Next, we will detail the calculation method of each feature.

PSD is defined as the signal power in the unit frequency band and represents the change of signal power with frequency, that is, the distribution of signal power in the frequency domain. The calculation steps of PSD of discrete time series are as follows: the average power of the power signal $f(t)$ in the time period $t \in [-T/2, T/2]$ can be expressed as:

$$P = \int_{-T/2}^{T/2} f(t)^2 dt \tag{1}$$

If $f(t)$ can be represented by $f_T(t)$ in the time period $t \in [-T/2, T/2]$, and the Fourier transform is $f_T(w) = F[f_T(t)]$, where $F[\,]$ represents the Fourier transform. When T increases, the power of $f_T(w)$ and $|f_T(w)|^2$ increases. When $T \rightarrow +\infty$,$f_T(t) \rightarrow f(t)$, $\frac{|f_T(w)|^2}{2\pi T}$ may be approaching the limit at this time. If the limit exists, its average power can also be expressed in the frequency domain, that is:

$$P = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f^2(t) dt \& = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \lim_{T \rightarrow +\infty} \frac{|f_T(w)|^2}{2\pi T} dw \tag{2}$$

$\frac{|f_T(w)|^2}{2\pi T}$ is defined as power spectrum density of $f(t)$, its expression is as follows:

$$PSD = \lim_{T \rightarrow \infty} \frac{|F_T(w)|^2}{2\pi T} \tag{3}$$

DE is used to measure the complexity of continuous random variables. DE describes the entropy of continuous random variables. It has the meaning of information measurement and plays an important role in EEG feature. The DE feature is defined as follows :

$$\begin{aligned} h(X) = -\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{(x-\mu)^2}{2\sigma^2} \log \frac{1}{\sqrt{2\pi\sigma^2}} \\ \exp \frac{(x-\mu)^2}{2\sigma^2} dx = \frac{1}{2} \log 2\pi e \sigma^2 \end{aligned} \tag{4}$$

where X submits the Gauss distribution $N(;2)$, x is a variable, and π and e are constants.

Previous studies have shown that both symmetrical and adjacent electrodes have features associated with emotion. Therefore, we investigate asymmetry features. We compute DASM and RASM features as the differences and ratios between the DE features of 14 pairs of hemispheric asymmetry electrodes (FP1-FP2, AF3-AF4, F7-F8, F3-F4, FC5-FC6, FC1-FC2, C3-C4, T7-T8, CP1-CP2, CP5-CP6, P3-P4, PO3-PO4, P7-P8, O1-O2). DASM and RASM can be expressed, respectively, as

$$DASM = DE\left(X_{\text{left}}\right) - DE\left(X_{\text{right}}\right) \tag{5}$$

and

$$RASM = DE\left(X_{\text{left}}\right) / DE\left(X_{\text{right}}\right) \tag{6}$$

ASM features are the direct concatenation of DASM and RASM features for comparison. To characterize the spectral-band asymmetry in respect of caudality (in frontal-posterior direction), we define DCAU features as the differences between DE features of 8 pairs of frontal-posterior electrodes (F7-P7, FC5-CP5, F3-P3, FP1-CP1, FZ-PZ, FP2-CP2, F4-P4, FC6-CP6, F8-P8). DCAU is defined as

$$DCAU = DE\left(X_{\text{frontal}} - DE\left(X_{\text{posterior}}\right)\right. \tag{7}$$

The dimensions of PSD, DE, DASM, RASM, ASM and DCAU are 160 (32 electrodes $\times$ 5 bands), 160 (32 electrodes $\times$ 5 bands), 70 (14 electrode pairs $\times$ 5 bands), 70 (14 electrode pairs $\times$ 5 bands), 140 (28 electrode pairs $\times$ 5 bands), and 40 (8 electrode pairs $\times$ 5 bands), respectively. So far, we extracted 640 features in 5 frequency bands and denoted as $\mathbb{R}^{640}$.

Authorized licensed use limited to: Universita degli Studi di Salerno. Downloaded on May 04,2023 at 07:54:35 UTC from IEEE Xplore. Restrictions apply.

C. GCN

As we all known, there is an apparent relationship between these adjacent channels. Emotion is reflected on the EEG by the surrounding adjacent electrodes. Therefore, the key to emotion recognition is to extract features by convolution between adjacent electrodes. We can utilize the relationship between the electrode pairs to construct GCN model. We define raw EEG signals as x and the convolution kernel $g_\theta = \operatorname{diag}(\theta)$ with parameters $\theta \in R^N$ in the Fourier domain. The spectral convolution on the graph is expressed as follows:

$$g_\theta \star x = U g_\theta U^T x \tag{8}$$

where U denotes eigenvector matrix, which is obtained by symmetric normalized Laplacian operator:

$$\begin{aligned} L &= I_N - D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \\ L &= U \Lambda U^T \end{aligned} \tag{9}$$

where denotes eigenvalue matrix and $U^T x$ means the graph Fourier transform of x. It is costly to calculate U. Consequently, the k-order Chebyshev polynomials is used to make an approximation:

$$g_{\theta'}(\Lambda) \approx \sum_{k=0}^{K} \theta'_k T_k(\tilde{\Lambda}) \tag{10}$$

where $\widetilde{\Lambda} = \frac{2}{\lambda_{\max}} \Lambda - I_N$, and $\theta' \in \mathbb{R}^K$ represents the Chebyshev coefficient vector. The Chebyshev polynomial is defined as $T_k(x) = 2x T_{k-1}(x) - T_{k-2}(x)$, in which $T_0(x) = 1$, $T_1(x) = x$. With these approximations, now we have:

$$g_{\theta'} \star x \approx \sum_{k=0}^{K} \theta'_k T_k(\tilde{L}) x \tag{11}$$

where $\tilde{L} = \frac{2}{\lambda_{\max}} L - I_N$, formula (11) shows the Chebyshev k^{th} order approximation, which reduces the parameters of the model and the complexity of the calculation. When k=1, formula (11) can be simplified as:

$$g_{\theta'} \star x \approx \theta'_0 x + \theta'_1 \left(L - I_N\right) x = \theta'_0 x - \theta'_1 D^{-\frac{1}{2}} A D^{-\frac{1}{2}} x \tag{12}$$

Let $\theta = \theta'_0 - \theta'_1$, we can get the following approximation:

$$g_\theta \star x \approx \theta \left(I_N + D^{-\frac{1}{2}} A D^{-\frac{1}{2}}\right) x \tag{13}$$

However, if the operation is used directly in deep neural network model, it will lead to gradient explosion or disappearance. Therefore, a normalization technique is introduced:

$$I_N + D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \rightarrow \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} \tag{14}$$

where $\tilde{A} = A + I_N$ and $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$. $S(A) = I_N + D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$ is defined as the the propagation matrix of A. The GCN model can be obtained:

$$\begin{aligned} H^{(l+1)} &= \sigma\left(\tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2} H^{(l)} W^{(l)}\right) \\ &= \sigma\left(S\left(A^l\right) H^l W^l\right) \end{aligned} \tag{15}$$

The overall GCN model is as follows. Three GCN layers are applied to propagate the features between adjacent electrode pairs. Two fully connected layers are used to summarize the features of all electrode pairs.

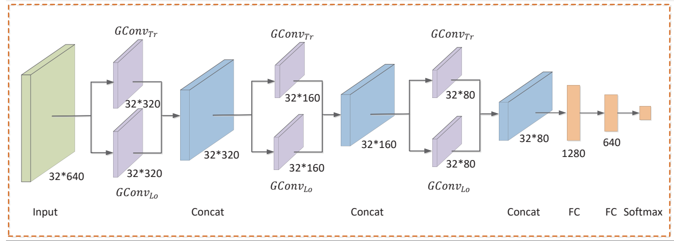

Fig. 2: The HGCN model.

D. HGCN

EEG electrodes are connected to each other by transverse and longitudinal edges, and there are voltage differences between them, and each edge contains different information. It is inappropriate to represent both edges with 1 in the adjacency matrix of A on GCN model. Therefore, we referred study in [8] and improved the GCN model. A natural idea is to represent the transverse and longitudinal edges with two independent adjacency matrixes, which are denoted as A_{Lo} and A_{Tr}.

The overall HGCN model is illustrated in Fig.2. The HGCN consists of seven layers and the specific description of each layer is as follows:

1) Input layer (L1): The input of each sample is a matrix, $N \times M$ (32×640) is the input size of HGCN, where N represents channels, M represents features.

2) Convolutional layer (L2): This layer contains one layer of $GCONV_{Lo}$ and $GCNOV_{Tr}$ and concatenation layer. M are propagated through two branches namely $GCONV_{Lo}$ and $GCNOV_{Tr}$ and concatenation layer is used to connect features. The size of generated propagation matrix is $N \times M(32 \times 320)$.

3) Convolutional layer (L3): This layer is same as L2 layer, The size of generated propagation matrix is $N \times M(32 \times 160)$.

4) Convolutional layer (L4): This layer is same as L3 layer, The size of generated propagation matrix is $N \times M(32 \times 80)$.

5) Fully connected layer (L5): This layer uses features from L4 to generate new features to provide high-dimensional information for the next layer.

6) Fully connected layer (L6): This layer to provide higher-dimensional information for the last layer. The two fully connected layers are applied to summarize the features of all electrode pairs.

7) Output layer (L7): This layer utilizes the calculation results from L6 layer to classify emotional states.

III. EXPERIMENTAL RESULTS AND DISCUSSION

A. Experimental setup

The experiments are conducted on multimodal DEAP dataset. DEAP is a large dataset containing multiple physiological signals and emotional assessments [9]. In our experiments, only EEG signals are used. At the meanwhile, we define the label of DEAP EEG data as follows: arousal/valence with self-score more than 5 is high arousal/positive valence, otherwise it is low arousal/negative valence. To verify the generality of the

Authorized licensed use limited to: Universita degli Studi di Salerno. Downloaded on May 04,2023 at 07:54:35 UTC from IEEE Xplore. Restrictions apply.

method, we randomly set 70% of the features as the training set and remaining 30% as the test set, and use the ten-fold cross-validation method to obtain the classification accuracy when classifying the features of each subject. The mean of them is taken as the result of the subject and the average performance of all subjects is considered to be the final results of the method.

B. Classification accuracy based on GCN

The average accuracies of valence and arousal of 6 features from separate and total bands based on GCN are summarized in Table 1. From Table 1, we can find that the recognition accuracies of all features on β band and γ band is much higher than other bands, especially on γ band. Study in [10] has demonstrated that β (14-30Hz) and γ (31-50Hz) bands contain more emotional information than other bands and they are commonly used in EEG researches, it is consistent with our study. The DE features outperform the other features significantly. In addition, the best accuracies of valence and arousal reached 84.58% and 83.94% of all features from total bands. It is indicated that the all features are benefit to emotion recognition.

C. Classification accuracy based on HGCN

The average accuracies of valence and arousal of 6 features from separate and total bands based on HGCN are summarized in Table 2. From Table 2, we can find that the performance of β band and γ band is better than other bands. It is worth mentioning that the performance of classification accuracy has been significantly improved in any band of any feature both valence and arousal based on HGCN.

In order to further illustrate the superiority of HGCN, we compared the best classification accuracy of the six features, as shown in Fig.3 and Fig.4. From Fig.3, we can clearly know that in the valence dimension, the best classification accuracy of each feature has been improved to varying degrees. Among them, the classification accuracy of DASM, RASM, ASM, and DCAU features has been improved by 15.55%, 12.47%, 12.81% and 11.18%, respectively. From Fig.4, similarly, in the arousal dimension, the classification accuracy of DASM, RASM, ASM and DCAU features has been improved by 14.57%, 13.81%, 14.63% and 12.14%, respectively. For adjacent and asymmetric features, compared with GCN model, HGCN model fully considers the relationship between electrodes and obtains richer spatial feature information, thus greatly improving the classification accuracy.

The best classification accuracies of valence and arousal achieved 90.56% and 88.79%, respectively based on HGCN. It is demonstrated that the HGCN model with transverse and longitudinal adjacent electrode pairs can better summarize all the features, thus achieving better classification performance. In addition, the improvement of classification performance of PSD and DE features is less than that of other features, it also reflects the stability and significance of PSD and DE features in EEG emotion recognition.

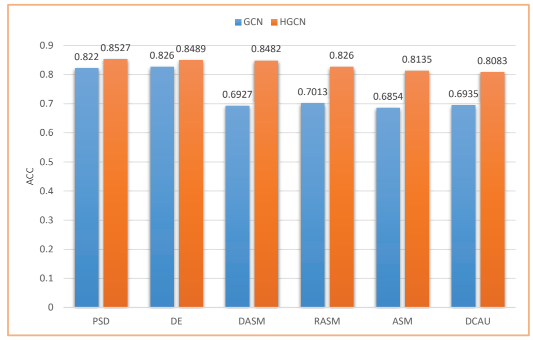

Fig. 3: Comparison against the best classification accuracy of valence based on GCN and HGCN.

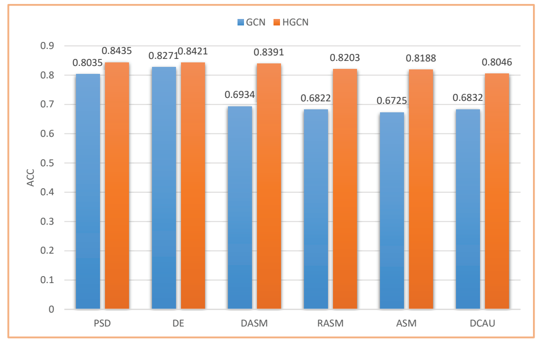

Fig. 4: Comparison against the best classification accuracy of arousal based on GCN and HGCN.

D. Comparison of ERHGCN method with other studies

To evaluate the performance of ERHGCN on DEAP dataset, our method is compared with state-of-the-art researches, as shown in Table 3. Experimental results show that our method is the most effective compared with the other methods. Qiu et al. proposed correlated attention networks (CAN), which transformed each mode to obtain higher mutual information and strengthen the similarity between different modes. The classification accuracies of valence and arousal can achieve 86.45% and 84.79%, which is lower 4.11% and 4% than ours, respectively [11]. Zhang et al. incorporated transfer learning framework and proposed an individual similarity guided transfer modeling method for EEG-based emotion recognition, which achieved classification accuracies of 66.1% and 66.7% on arousal and valence dimensions, respectively, which is lower 24.46% and 22.09% than ours [12]. Tang et al. employed DE feature to LSTM classifier, and obtained the classification accuracies 83.82% and 83.23% in valence and arousal, respectively, which is lower 6.74 % and 5.56% than ours [13]. Our ERHGCN method uses EEG data and can extract multidomain features including frequency, adjacent and asymmetric features, which can provide complementary information for emotion recognition. Therefore, the performance of classification is greatly improved.

Authorized licensed use limited to: Universita degli Studi di Salerno. Downloaded on May 04,2023 at 07:54:35 UTC from IEEE Xplore. Restrictions apply.

TABLE I: Comparison against the average accuracy of valence and arousal of 6 features from separate and total bands based on GCN

Features	δ band	θ band	α band	β band	γ band	total bands
PSD	69.24/68.36	65.33/62.47	66.44/63.87	79.85/78.36	82.20/80.35	80.37/79.26
DE	71.08/70.43	67.39/63.56	68.30/64.27	80.46/80.11	82.60/82.71	82.27/82.68
DASM	65.27/63.45	62.55/61.03	62.39/61.92	67.70/66.72	68.05/68.35	69.27/69.34
RASM	63.74/62.55	60.08/60.39	61.83/62.54	66.92/65.58	68.29/67.91	70.13/68.22
ASM	62.04/61.32	61.46/62.51	61.25/62.01	64.50/63.86	67.37/66.56	68.54/67.25
DCAU	62.55/62.31	62.27/61.59	62.36/63.41	65.44/66.23	66.28/67.41	69.35/68.32
Total	73.28/72.09	72.54/69.38	70.55/68.32	82.36/81.29	85.46/84.30	84.58/83.94

TABLE II: Comparison against the average accuracy of valence and arousal of 6 features from separate and total bands based on HGCN

Features	δ band	θ band	α band	β band	γ band	total bands
PSD	73.37/72.52	74.56/72.48	75.66/75.02	83.25/82.41	85.27/84.35	83.46/82.78
DE	76.44/75.29	78.20/75.46	79.01/78.37	85.27/85.05	84.89/83.93	84.55/84.21
DASM	73.45/71.39	76.02/73.28	77.36/75.28	83.51/83.40	82.26/81.53	84.82/83.91
RASM	72.85/71.27	75.46/72.05	75.20/74.33	81.26/82.05	81.39/80.98	82.86/82.03
ASM	70.36/70.58	72.39/72.21	71.33/72.31	79.86/79.42	80.25/81.03	81.35/81.88
DCAU	70.45/69.92	71.38/72.02	71.35/71.43	78.93/79.26	79.45/80.04	80.83/80.46
Total	77.46/78.05	81.03/78.56	81.86/80.72	87.50/87.42	88.02/86.01	90.56/88.79

TABLE III: Comparison of ERHGCN method with other studies.

method	Accuracy (%)	
	Valence	Arousal
CAN [11]	86.45	84.79
TrAdaboost [12]	66.1	66.7
Bimodal-LSTM [13]	83.82	83.23
ERHGCN	90.56	88.79

IV. CONCLUSION

In this paper, we propose an EEG emotion recognition method based on hierarchy graph convolution network. In view of complex EEG feature information, we fully consider multiple dimensions and extract frequency domain feature, PSD and DE features, and spatial features, as DASM, RASM ASM and DCAU. Meanwhile, based on GCN, we consider the relationship between adjacent electrodes and calculate the information between transverse and longitudinal electrodes respectively. In our future work, we will consider integrating multimodal methods into our model to improve performance. Moreover, how to use fewer EEG channels is worth studying.

ACKNOWLEDGMENTS

We are grateful for the support of the Natural Science Foundation of Shandong Province (No. ZR2020LZH008, ZR2020QF112, ZR2019MF071) and the Taishan Scholars Program of Shandong Province, China.

REFERENCES

[1] J. H. Zhang, Y. Zhong, C. Peng, and A. Stefano, "Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review - sciencedirect," *Information Fusion*, vol. 59, pp. 103–126, 2020.

[2] M. R. Islam, M. A. Moni, M. M. Islam, M. Rashed-Al-Mahfuz, and P. Lio, "Emotion recognition from eeg signal focusing on deep learning and shallow learning techniques," *IEEE Access*, vol. 9, pp. 94 601–94 624, 2021.

[3] X. Zheng, M. Zhang, T. Li, C. Ji, and B. Hu, "A novel consciousness emotion recognition method using erp components and mmse," *Journal of Neural Engineering*, vol. 18, no. 4, p. 046001, 2021.

[4] T. Song, W. Zheng, P. Song, and Z. Cui, "Eeg emotion recognition using dynamical graph convolutional neural networks," *IEEE Transactions on Affective Computing*, pp. 532–541, 2018.

[5] W.-L. Zheng, J.-Y. Zhu, and B.-L. Lu, "Identifying stable patterns over time for emotion recognition from eeg," *https://arxiv.org/abs/1601.02197v1*, 2016.

[6] X. Zheng, X. Yu, Y. Yin, T. Li, and X. Yan, "Three-dimensional feature maps and convolutional neural network-basedemotion recognition," *International Journal of Intelligent Systems*, pp. 6312–6336, 2021.

[7] Y. Yin, X. Zheng, B. Hu, Y. Zhang, and X. Cui, "Eeg emotion recognition using fusion model of graph convolutional neural networks and lstm," *Applied Soft Computing*, vol. 100, p. 106954, 2021.

[8] D. Zeng, K. Huang, C. Xu, H. Shen, and Z. Chen, "Hierarchy graph convolution network and tree classification for epileptic detection on electroencephalography signals," *IEEE Transactions on Cognitive and Developmental Systems*, vol. PP, no. 99, pp. 10–22.

[9] S. Koelstra, "Deap: A database for emotion analysis ;using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012.

[10] N. Zhuang, Z. Ying, L. Tong, Z. Chi, H. Zhang, and B. Yan, "Emotion recognition from eeg signals using multidimensional information in emd domain," *BioMed research international*, vol. 2017, pp. 1–9, 2017.

[11] J. L. Qiu, X. Y. Li, and K. Hu, "Correlated attention networks for multimodal emotion recognition," in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2018, pp. 2656–2660.

[12] X. Zhang, W. Liang, T. Ding, J. Pan, and J. Gao, "Individual similarity guided transfer modeling for eeg-based emotion recognition," in *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2019, pp. 1156–1161.

[13] H. Tang, W. Liu, W. L. Zheng, and B. L. Lu, "Multimodal emotion recognition using deep neural networks," *Neural Information Processing*, pp. 811–819, 2017.

Authorized licensed use limited to: Universita degli Studi di Salerno. Downloaded on May 04,2023 at 07:54:35 UTC from IEEE Xplore. Restrictions apply.