EEG BASED MULTIMODAL EMOTION RECOGNITION ESPOUSED DEEP KRONECKER NEURAL NETWORK

G. Kalyana Chakravarthy[1*], Dr M Suchithra[2]

Abstract

Emotion recognition plays numerous important roles in the life of individuals in the context of artificial intelligence technology. The majority of existing emotion recognition techniques performs poorly in practical applications, preventing their advancement. Hence, put forth a Deep Kronecker Neural Network based multi-modal expression EEG interaction technique to address this issue. First, Hexadecimal Local Adaptive Binary Pattern (HLABP) is used as an objective way of feature extraction. Depending on facial expression, the features are selectedwith the help of Weibull Distributive Generalized Multidimensional Scaling (WDGMS), the solution vector coefficients are scrutinized to scale the facial expression type of test samples. Finally, Deep Kronecker Neural Network (DKNN)completesthe classification task. Then, the proposed method is simulated utilizing MATLAB under several performance metrics, like F1 score, accuracy, error rate, average running time. The proposed technique attains 23.34%, 16.64% higher accuracy and 34.61%, 41.23% lower average running time when comparing to the existing methods, such as Expression-EEG Interaction Multiple Modal Emotion Recognition utilizing Deep Automatic Encoder (EEG-MER-DAE) and EEG-based emotion recognition utilizing deep convolutional neural networks (EEG-DNN) respectively.

Keywords: EEG, Multimodal Emotion Recognition, Hexadecimal Local Adaptive Binary Pattern, Weibull Distributive Generalized Multidimensional Scaling, Deep Kronecker Neural Network.

[1*]Research Scholar, Department of CSE, SRM Institute of Science and Technology, Kattankulathur -Tamil Nadu, Email: kalyan5803@gmail.com
[2]Associate Professor, Department of CT, SRM Institute of Science and Technology, Kattankulathur -Tamil Nadu Email: suchithm@srmist.edu.in

***Corresponding Author:** G. Kalyana Chakravarthy
[*]Research Scholar, Department of CSE, SRM Institute of Science and Technology, Kattankulathur -Tamil Nadu, Email: kalyan5803@gmail.com

DOI: - 10.31838/ecb/2023.12.si5.089

1 INTRODUCTION

Emotion is a process which is related with expressions and works in conscious and unconscious situation in the human beings [1]. By using expressions the communications are done between peoples. Some emotions are sad, happy, angry, fear etc. In human-computer interaction more researches are done behalf of emotion recognition [2]. Human-computer interaction system is dynamic and complex therefore system consists of emotional connections. By using EEG signals, emotion recognition was characterized in the brain. So, nowadays EEG signal plays an important role in research [3]. Nowadays multimedia and human-computer interaction technology is highly upgraded so the emotions are automatically recognized [4]. By using emotion recognition, emotions will be adjusted through the recommendation of the third party. Where, Emotional AI is used to process the human emotion in various places of interest such as health care, entertainment, education etc. By using AI research in robotic field is increased simultaneously. In modern world all the Multi National Company's such as Microsoft, Google, and Samsung were invested Trillions of dollars in Emotion recognition techniques [5]. To execute this technique more domain knowledge is needed and it will take more time. As a result emotion recognition detects the user's emotion and responses to the user by using multimedia content.

In various ways people communicate their feelings, such as expressive speech, facial gestures, body languages and so on. Thus, emotional signals from various modalities are utilized for predicting the emotional state of subject. But, the existing techniques not judge the emotions of people easily. Hence, some solutions need to be put forward to fix these problems has motivated to do research in this area.

The rest of this manuscript is deliberated as: the recent studies arerevealed in section 2, the proposed system is clarified in section 3, the results with discussion are demonstratedin section 4, finally, section 5 concluding this manuscript.

2 LITERATURE SURVEY

Several research works were previously presented in the literature based upon EEG based emotion recognition; a few works are reviewed here,

In 2020, Zhang, H., [6] have suggested collaborative multimodal emotion recognition depending on expression-EEG utilizing deep autoencoder. It provides high F1 score and low precision. In 2021, Ozdemir, M.A., et.al., [7] have presented EEG-based emotion recognition utilizing deep convolutional neural networks. It provides maximum accuracy and minimum F1 score. In 2021, Fang, Y., et.al., [8] have suggested multiple feature input deep forest for EEG-base emotion recognition. It provides higher accuracy and lower precision. In 2020, Farashi, S. [9] have presented EEG basally emotion recognition utilizing minimal spanning tree. It provides maximum precision and lower F1 score.

3 PROPOSED METHODOLOGY

The ability of EEG signals to represent changes in human brain stage makes emotion recognition-depending upon EEG signals. For emotion recognition, a facial expression signal is included in addition to EEG signals as an external physiological characterisation signal. Deeply examines the EEG and facial expression signals capacity to characterise various emotions, and fusing EEG and facial expression signals using Deep Kronecker Neural Network to create multiple mode emotion recognition fusing with internal neural models as well as external sub-consciousness activities.Meanwhile, the steadiness of emotions expression capacity for EEG and facial expression signals over time. The emotion recognition accuracyisenhanced after fusion the signals of EEG and facial expression. Hexadecimal Local Adaptive Binary Pattern (HLABP) is used as an objective way of feature extraction. Facial expression basis, the features are selected through Weibull Distributive Generalized Multidimensional Scaling (WDGMS), the solution vector coefficients examined to define facial expression type of test samples. At last, Deep Kronecker Neural Network (DKNN) completesthe classification task.Figure 1 depicts the overall block diagram of EEG-MER-DKNN.

3.1 Signal Acquisition

The input signals are taken via publicly available SEED dataset. Fifteen test participants/subjects (7 males,8females) with 23.3 mean age and 2.4 standard deviation contributed in the experiments. EEG dataset has signals derived from the subjects while the subjects are presentation emotional video tapes. Participants requested to make the trials for 3 sessions each to assess neurological signals and significant reactions.This dataset has data from 45 diverse experiment sessions.

3.2 EEG Signal Feature Selection utilizing WDGMS

Thefeatures in EEG signal are selected using WDGMS. WDGMS base EEG signal feature selection achieves objectivity of feature selection and higher classification accuracy.

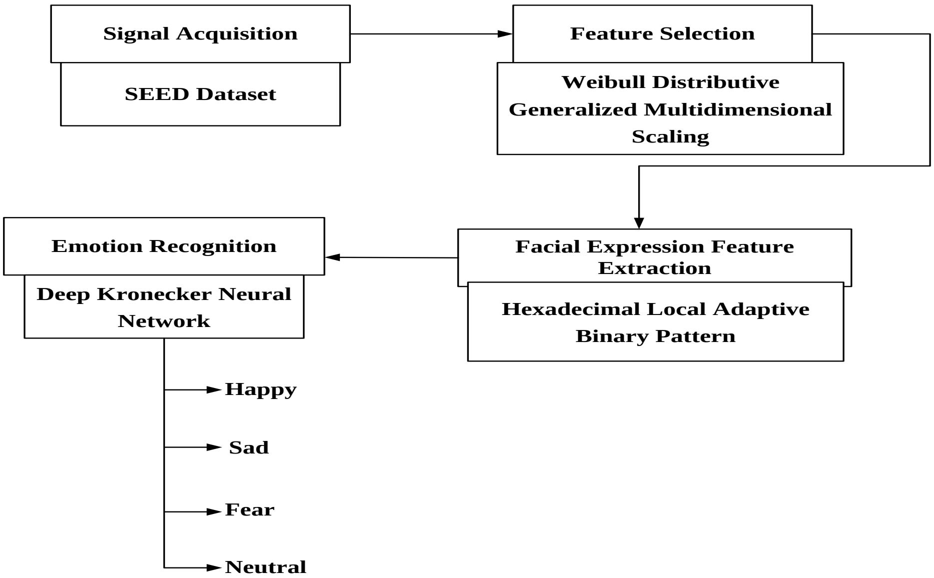

Figure 1: Overall block diagram of proposed EEG-MER-DKNN methodology

EEG signal is non-stationary, non-linear, higher dimension weak physiological signal. The feature selection process utilising WDGMS provides three benefits: simplicity of understanding, no necessary prior knowledge, high performance in selection. Consequently, WDGMS can increase classification accuracy as well as objectively choose features from high-dimensional and complex EEG signals. The proposed weibull distributive generalized multidimensional scaling is employed to vital feature selection. WDGMS select the significant features of higher dimensional data space. Weibull distributive generalized multidimensional scaling is a dimensionality lessening strategy for significant transformation. Weibull distributive generalized multidimensional scaling uses a number of features $a_1, a_2, a_3, ..., a_n$ are collected from the data. Afterward, structuring the feature matrix in eqn (1),

$$F = \begin{bmatrix} a_1 & a_2 & \dots & a_n \\ a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad (1)$$

Let F implies feature matrix. Weibull distribution is employedafter structuring the feature matrix, that is depicted in eqn (2),

$$D = \frac{\rho}{\varphi}\left(\frac{a_i - o_j}{\varphi}\right)^{\rho-1} \exp\left(-\frac{a_i - o_j}{\varphi}\right)^{\rho} \quad (2)$$

Where, ρ denotes shape parameter, o_j denotes features and objectives (attack identification) in the data, φ specifies scale parameter $\varphi = 1$, a_i refers attributes. It returns 0 to 1 value of output. The distribution outcomes basis the significant features are recognized to lessen the complexity of expression recognization and are exhibited in eqn (3),

$$Y = \begin{cases} D > T; & \textit{more significant feature} \\ D < T; & \textit{Less significant feature} \end{cases} \quad (3)$$

Here Y implies output of scaling, D implies output of weibull distribution, T implies threshold. Algorithmic procedure of weibull distributive generalized multidimensional scaling is specified in eqn (4),

$$f_R^{(r)} = xyr^{y-1}\exp\left(-xr^b\right), \; r > 0 \quad (4)$$

Where $x > 0$ denotes distributed scale parameter, $0 < y \le 2$ implies shape parameters. To $y = 1$, Weibull coincides to the exponential distribution. To $y = 2$, Weibull coincides to the Rayleigh distribution. Multiple variate Weibull distribution is revealed in eqn (5),

$$f_R^{(r)} = \frac{2(-1)}{\Gamma(L)}\sum_{n=1}^{L} D_n \frac{x^n}{k!} r^{ny-1}\exp\left(-xr^y\right) \quad (5)$$

Where,

$$D_n = \sum_{m=1}^{n}(-1)^m \binom{n}{m} \frac{\Gamma\left(1+ {my}/{2}\right)}{\Gamma\left(1+ {my}/{2} - L\right)} \quad (6)$$

From equation (6), $\Gamma\left(1+ {my}/{2}\right)$ and $\Gamma\left(1+ {my}/{2} - L\right)$ indicates the gamma function. To make simpler 2 parameters weibull distribution examination, consider normalized distribution manner, wherein scale parameter is limited for enforcing $E\left[r^2\right]=\sqrt{L}$. Such normalization is reached through concerning the scale parameter to y shape parameter exhibited in eqn (7),

$$x=\left[\Gamma\left(\frac{2}{y}+1\right)\right]^{y/2} \quad (7)$$

Eqn (7)depicts the associated feature selected. Weibull distribution is used to create the feature matrix. Weibull distribution is used to calculate the degree of similarity between the feature and the objective function. The more important traits are determined using the similarity measure. The feature is supposed to be appropriate feature when the similarity is higher, or, the feature is an inappropriate feature. It lessens time complexity of emotion recognition.

3.3 Facial Expression Feature Extraction Using Hexadecimal Local Adaptive Binary Pattern

A crucial component of expressing emotions is facial expression. Human faces and facial expressions can reveal a person's many emotional states. Recognize these states by using HLABP. In this phase, the selected features of EEG signal are delivered to a HLABP for feature extraction. It also utilizes two variable patterns is named central and linear symmetrical patterns. The mathematic representations of signum and ternary functions are expressed in equation (8) as follows,

$$B^{signum}(a)=Signum(a,b)=\begin{cases}0, & a-b<0\\ 1, & a-b\geq 0\end{cases} \quad (8)$$

Where, B^{signum} indicates the bit extracted from the signum function, $Signum(a,b)$ expresses the signum function. Then the mathematical notations of ternary functions are expressed in equation (9) as follows,

$$T(a,b)=\begin{cases}-1, & a-b<-tsd\\ 0, & -tsd\leq a-b\leq tsd\\ 1, & a-b>tsd\end{cases} \quad (9)$$

From equation (9), $T(a,b)$ indicates the ternary function and tsd denotes the threshold value. Then the lower and upper bits extracted as ternary function are expressed in equations (10) and (11) as follows,

$$B_L^T(a)=\begin{cases}0, & T(a,b)>-1\\ 1, & T(a,b)=-1\end{cases} \quad (10)$$

$$B_U^T(a)=\begin{cases}0, & T(a,b)<1\\ 1, & T(a,b)=1\end{cases} \quad (11)$$

The features extracted using HLABP is converted to decimal values with equation (12) as follows,

$$D_{V|}=\sum_{a=1}^{8}B(a)\times 2^{8-a} \quad (12)$$

From this time domain, various features are extracted directly EEG signals without any transformation; thus, it consists of less computational cost are simple to execute. By this more effective feature were extracted using Hexadecimal Local Adaptive Binary Pattern (HLABP). Then the extracted features are move toward the emotion recognition stage.

3.4 Multimodal emotion recognition depending on Deep Kronecker Neural Network

The classification process is crucial in identifying emotion recognition based on facial expression. In order to create a clear and accurate mode that is applied to forecast emotions through real-time, the best classification algorithm should be chosen. The performance and accuracy of multimodal emotion recognition is determined by this. A dynamic model for emotion classification is presented in this paper. The Flexible activation Functions with Deep Kronecker Neural Network is a dynamic model (DKNN) is used. The DKNN classifier is chosen to find the emotion recognition like sad, fear, happy and neutral. With average performance that is three times better than existing approaches, DKNN's capability to construct multiple levels of tree over the massive training data set using only two scans enables it to be tailored to the individual properties of emotions as well as address the issues with facial expression. DKNN also has the capacity for updating decision tree with reference to the dynamic addition via the emotion recognition, which is an additional significant issue in EEG signal. DKNN uses little run-time resources and doesn't need any storage to store temporary data. In classification unit, DKNN is used for categorization the emotions as sad, fear, happy and neutral.

A depth-based feed-forward neural network an input layer, $C-1$ hidden layers, and an output layer make up the numerous layers that determine the function C. There are N_k amount of neurons in the k^{th} hidden layer, $z^{k-1}\in R^{N_{k-1}}$ output from the emotion preceding layer, here an affine modification is exhibited in equation (13),

$$L_k(z^{k-1})\overset{\Delta}{=}T^k z^{k-1}+y^k \quad (13)$$

$T^k\in R^{N_k\times N_{k-1}}$ Signifies weight matrix in this instance, $y^k\in R^{N_k}$ signifies bias vector connected to the k^{th} layer. Each part of the changed vector is given a nonlinear activation function $\phi_1(.)$ before used as an input by the following layer. After output layer, the activation function implicates identity function. In this, the representation of final neural network is expressed in equation (14),

$$u^{EE}(z) = (L_C \circ \phi_1 \circ L_{C-1} \circ ... \circ \phi_1 \circ L_k)(z) \quad (14)$$

Where the composition operator $\circ$ is expressed by the operator, let $\Theta_{EE} = \{T^k, y^l\}_{k=1}^{C}$ represent the parameters that can be trained for the network. Let's review the various v norms for the vector $v = [v_1,..,v_n]^W \in R^n$ is expressed in equation (15)

$$\| v \|_1 = \sum_{i=1}^{n} | v_i |, \quad \| v \|_2 = \sum_{i=1}^{n} | v^2{}_i |, \quad \| v \|_\infty = \max_{1 \le i \le n} | v_i |$$
(15)

Let $\sigma_{\min}(M)$ be the n^{th} biggest singular value of a matrix $M \in R^{m\times n}$ where $m \ge n$. Furthermore, the definitions of spectral as well as Frobenius norm is described in equation (16)

$$\| M \| = \max_{\|a\|=1} \| Ma \| \quad \| M \|_E^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} M^2{}_{ij} \quad (16)$$

Here M_{ij} implies (i, j)-component of M. Consider $I_{r\times w}$ implies matrix of $r \times w$ size those entries are 1s.

Consider K denotes fixed positive integer. FNN's parameters $\Theta_{EE} = \{T^k, y^w\}_{k=1}^{C}$; k^{th} block weight matrix along block bias vector in equation (17)

$$1_{l\times l} \otimes T^k = \begin{bmatrix} T^k & ... & T^k \\ ... & ... & ... \\ T^k & ... & T^k \end{bmatrix} \in R^{N_k \times N_{k-1} K}, \quad 1_{K\times 1} \otimes y^1 = \begin{bmatrix} y^k \\ ... \\ y^k \end{bmatrix} \in R^{N_k K}$$
(17)

Let $\otimes$ denotes Kronecker product. To determine a block activation function $\vec{\phi}$, block-wise is applied, i.e. $z_j \in R^n$ for $1 \le j \le K$, let $z = [z_1 ... z_k]^W \in R^{nl}$,it is described in equation (18),

$$\vec{\phi}(z) = \begin{bmatrix} \phi_1(z_1) \\ ... \\ \phi_l(z_l) \end{bmatrix} \quad (18)$$

Here ϕ_j is activation functions applied element-wise. Any network packet is regarded as emotions if it matches the splitting decision requirement. When multimodal emotions are identify aberrant activity, it undo all of the modifications that were made to the contract. Because of this, an opportunity to avoid suffering an irreparable loss due to the vulnerability may cause.

4. RESULT WITH DISCUSSION

This section describes the experimental result of proposed EEG-MER-DKNN method. The proposed technique is simulated utilizing MathWorks Inc, MATLAB® version 9.7.0.1190202 (R2019b) under performance metrics. The acquired outcomes ofEEG-MER-DKNN are analyzed with the existing methods, such as EEG-MER-DAE [21], EEG-DNN [22].

4.1Performance measures

The performance metrics, such as F1 score, accuracy, error rate, average running timeis assessed. The following confusion matrix is deemed to scale the performance metrics.

4.1.1 F1 Score

This is determined by equation (19),

$$F1score = \frac{h}{\left(h + \frac{1}{2}[i + j]\right)} \quad (19)$$

Where h indicates true positive, i indicates true negative, j indicates false positive.

4.1.2 Accuracy

It is measured by following equation (20),

$$A = \frac{h + k}{h + i + j + k} \quad (20)$$

Here, k indicates false negative.

4.1.3Error rate

This is scaled by equation (21),

$$E = 100 - accuracy \quad (21)$$

Where E is represented as error rate.

PERFORMANCE ANALYSIS

Table 1-4tabulates the performance of proposed EEG-MER-DKNN method. The performance metrics is analyzed. Then the performance is analyzed with existing EEG-MER-LIBSVM and EEG-MER-CNN models respectively.

Table 1: Comparison of accuracy analysis

Techniques	Accuracy (%)			
	Happy	Sad	Fear	Neutral
EEG-MER-LIBSVM	74.56	88.56	89.77	82.45
EEG-MER-CNN	82.45	76.33	79.67	85.88
EEG-MER-DKNN(proposed)	96.63	97.51	98.89	96.67

Table 1 demonstrates the Comparison of accuracy analysis. Here, the proposed EEG-MER-DKNN method provides 25.44% and 32.46% higher accuracy for happy; 22.45% and 21.77% higher accuracy for sad; 44.23% and 23.22% higher accuracy for fear; 44.23% and 44.23% higher accuracy for Neutral compared with existing methods, such as EEG-MER-LIBSVM and EEG-MER-CNN respectively.

Table 2: Comparison of F1 score analysis

Techniques	F1 score (%)			
	Happy	Sad	Fear	Neutral
EEG-MER-LIBSVM	74.56	92.45	83.45	87.45
EEG-MER-CNN	82.34	78.5	76.56	78.55
EEG-MER-DKNN(proposed)	97.78	98.85	99.71	98.86

Table 2demonstrates the Comparison of F1 score analysis. Here, the proposed EEG-MER-DKNN method provides 34.56%, and 32.46% higher F1 score for happy; 37.61%, and 21.77% higher F1 score for sad; 44.23%, and 44.23% higher F1 score for fear; 44.23%, and 44.23% higher F1 score for Neutral compared with existing methods, such as EEG-MER-LIBSVM and EEG-MER-CNN respectively.

Table 3: Comparison of Average Running Time

Techniques	Average Running Time (s)			
	Happy	Sad	Fear	Neutral
EEG-MER-LIBSVM	5.55	6.66	7.56	7.45
EEG-MER-CNN	8.9	6.66	7.45	9.7
proposed	2.36	3.96	4.17	5.61

Table 3 demonstrates the Comparison of Average Running Time. Here, the proposed EEG-MER-DKNN method provides 34.56%, and 32.46% lower Average Running Time for happy; 37.61%, and 21.77% lower Average Running Time for sad; 44.23%, and 44.23% lower Average Running Time for fear; 44.23%, and 44.23% lower Average Running Time for Neutral compared with existing methods, such as EEG-MER-LIBSVM and EEG-MER-CNN respectively.

Table 4: Comparison of Error rate

Techniques	Error rate (%)			
	Happy	Sad	Fear	Neutral
EEG-MER-LIBSVM	7.5	8.45	9.67	4.56
EEG-MER-CNN	6.45	4.3	3.4	5.7
proposed	2.30	1.35	1.67	2.11

Table 4 demonstrates the Comparison of Error rate. Here, the proposed EEG-MER-DKNN method provides 34.56%, and 32.46% lower Error rate for happy; 37.61%, and 21.77% lower Error rate for sad; 44.23%, and 44.23% lower Error rate for Fear; 44.23%, and 44.23% lower Error rate for Neutral compared with existing methods, such as EEG-MER-LIBSVM and EEG-MER-CNN respectively.

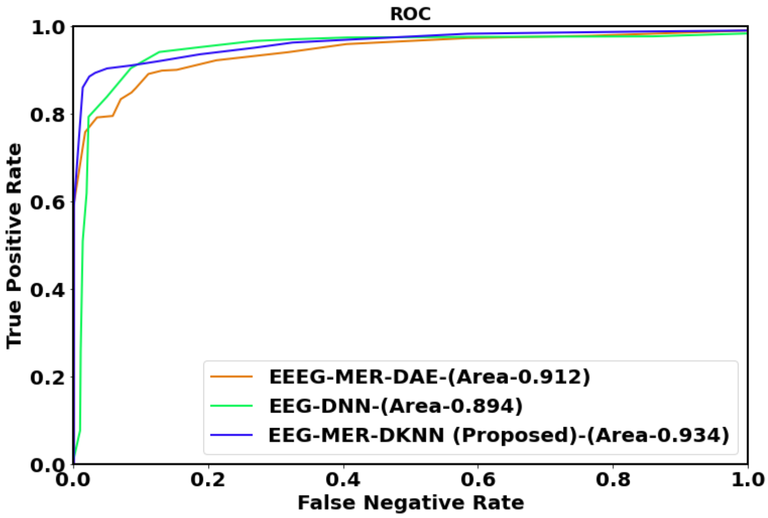

Figure 2: Analysis of RoC

Figure 2 depicts the analysis of RoC. Here, the proposed EEG-MER-DKNN method attains 2.292% and 3.915% higher AUC compared with existing methods such as EEG-MER-LIBSVM and EEG-MER-CNN respectively.

5. CONCLUSION

Here, EEG Based Multimodal Emotion Recognition Espoused Deep Kronecker Neural Network (EEG-MER-DKNN) is successfully implemented. Here the proposed method is simulated utilizing MATLAB under several performance metrics. The proposed technique attains11.25% and 4.25% higher F1-Score and 22.47% and 31.58% lesser error ratewhen comparing to the existing EEEG-MER-DAE and EEG-DNNmodels respectively.

REFERENCES

1. Abdullah, S.M.S.A., Ameen, S.Y.A., Sadeeq, M.A. and Zeebaree, S., 2021. Multimodal emotion recognition using deep learning. *Journal of Applied Science and Technology Trends*, *2*(02), pp.52-58.
2. Koromilas, P. and Giannakopoulos, T., 2021. Deep multimodal emotion recognition on human speech: A review. *Applied Sciences*, *11*(17), p.7962.
3. Siddiqui, M.F.H., Dhakal, P., Yang, X. and Javaid, A.Y., 2022. A Survey on Databases for Multimodal Emotion Recognition and an Introduction to the VIRI (Visible and InfraRed Image) Database. *Multimodal Technologies and Interaction*, *6*(6), p.47.
4. Liu, W., Qiu, J.L., Zheng, W.L. and Lu, B.L., 2021. Comparing recognition performance and robustness of multimodal deep learning models for multimodal emotion recognition. *IEEE Transactions on Cognitive and Developmental Systems*, *14*(2), pp.715-729.
5. Xie, B., Sidulova, M. and Park, C.H., 2021. Robust multimodal emotion recognition from conversation with transformer-based crossmodality fusion. *Sensors*, *21*(14), p.4913.
6. Zhang, H., 2020. Expression-EEG based collaborative multimodal emotion recognition using deep autoencoder. *IEEE Access*, *8*, pp.164130-164143.
7. Ozdemir, M.A., Degirmenci, M., Izci, E. and Akan, A., 2021. EEG-based emotion recognition with deep convolutional neural networks. *Biomedical Engineering/ Biome dizinischeTechnik*, *66*(1), pp.43-57.
8. Fang, Y., Yang, H., Zhang, X., Liu, H. and Tao, B., 2021. Multi-feature input deep forest for EEG-based emotion recognition. *Frontiers in neurorobotics*, *14*, p.617531.
9. Farashi, S. and Khosrowabadi, R., 2020. EEG based emotion recognition using minimum spanning tree. *Physical and Engineering Sciences in Medicine*, *43*(3), pp.985-996.