

A Multilevel Feature-based Emotion Classification Technique using Electroencephalogram Signals

Kabita Devi^{id}, Bhabesh Deka^{id}, and Debarun Chakraborty^{id}

Department of Electronics and Communication Engineering, Tezpur University, Assam, India, 784028

Corresponding author: Bhabesh Deka (e-mail: bdeka@tezu.ernet.in), **Author(s) Email:** Kabita Devi (e-mail: kdevi12@tezu.ernet.in), Debarun Chakraborty (e-mail: dkdeb20@tezu.ernet.in)

Abstract Emotion dysfunction can be an indication of mental health conditions like depression, stress, anxiety, etc. Therefore, appropriate recognition of human emotion is crucial for early monitoring of mental well-being. This study aims to develop a robust multilevel feature concatenation framework that effectively integrates handcrafted and deep learning (DL) features to improve electroencephalogram (EEG)-based emotion classification. The proposed framework is evaluated on two publicly available benchmark EEG datasets: DEAP and DREAMER. Initially, handcrafted features extracted from the statistical, time, frequency, and nonlinear domains are categorized into low-level features (LLF) and mid-level features (MLF). To select the most relevant handcrafted features, statistical analysis, minimum Redundancy Maximum Relevance (mRMR) and the Pearson correlation coefficient (PCC) are employed. In parallel, a two-dimensional convolutional neural network (2D-CNN) is used to learn hierarchical deep features, which are grouped into LLF, MLF, and high-level features (HLF). The handcrafted LLF and MLF are then concatenated with the corresponding DL-based LLF, MLF, and HLF to construct a discriminative multilevel feature representation. The resulting feature vector is classified using Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and XGBoost classifiers under a 10-fold and leave-one-subject-out (LOSO) cross-validation protocol. Performance is evaluated using accuracy, precision, recall, F1-score, and balanced accuracy. Experimental results demonstrate that the proposed framework consistently outperforms existing state-of-the-art methods. Among the evaluated classifiers, SVM achieves the highest performance, yielding accuracies of 93.07% and 98.70% for 4-class and 2-class emotion classification, respectively, on the DEAP dataset, and 93.85% and 98.78% for 4-class and 2-class emotion classification, on the DREAMER dataset. These results demonstrate that concatenation of handcrafted with DL-based features at different abstraction level significantly enhances feature discriminability and classification robustness. The proposed framework provides an effective solution for EEG-based emotion recognition, with promising applications in affective computing, human–computer interaction, and intelligent mental health monitoring.

Keywords Convolution Neural Networks (CNN); Electroencephalogram (EEG); Emotion Recognition; Multilevel Feature.

1. Introduction

Emotion recognition plays a critical role in mental health monitoring and treatment. Compared to traditional methods and other physiological signal-based approaches for mental health assessment, electroencephalogram (EEG) based emotion recognition method has gained much attention due to its several technological advantages; it is rich in spatiotemporal information, provides direct insights into neural processes, and enables dynamic variations in brain activity related to various emotional states [1], [2]. machine learning (ML)-based emotion recognition studies typically involve the extraction of handcrafted features across various domains: time, frequency, time-frequency, nonlinear, and their combinations from

EEG signals [3], [4]. These features carry relevant information that can distinguish different emotional states [5], [6]. Various handcrafted feature-based emotion classification studies have explored different signal decomposition techniques to decompose the EEG signals into their different bands [7], [8], [9]. Although ML algorithms and handcrafted features using statistical learning can yield good domain-invariant features for EEG-based emotion classification, extraction of handcrafted features requires manual selection and evaluation of relevant channels and features. This process can also introduce human bias and may not always capture the most relevant features for emotion classification. Moreover, handcrafted features might not fully exploit the rich

information (high level-features) embedded in EEG signals, potentially leading to suboptimal performance in recognizing complex emotional states. Deep learning (DL) techniques, on the other hand, leverage artificial neural networks (ANNs) to learn high-level features from raw EEG data automatically. DL-based architectures can reveal complex relationships and patterns that may be challenging to identify with handcrafted features. Over the past few years, DL has gained more popularity in EEG-based emotion recognition because it can automatically extract hierarchical patterns from raw EEG signals. Recurrent neural networks (RNNs), especially the long short-term memory (LSTM) networks, and convolutional neural networks (CNNs) are the two most commonly used architectures in such studies [10], [11], [12], [13]. However, they also have some drawbacks, such as the necessity of a huge amount of labeled data for model training, making it computationally intensive and thereby limiting their application in real-world wearable devices. Recent research states that lightweight DL-based [14] models are well suited for emotion recognition and mental health studies, especially when real-world applications are required. Also, a few previous studies [15], [16] have reported that combining handcrafted and DL-based features can significantly improve performance in mental health classification tasks. A summary of different EEG-based literatures for emotion classification and mental health analysis are presented in Table 1.

Despite these advances, existing hybrid methods [15], [16] commonly perform direct concatenation/fusion of handcrafted and DL features. Other studies [11], [13], have explored hierarchical, layer-wise, or multi-level DL-based feature fusion captured from different DL layers. However, these strategies do not explicitly align handcrafted EEG features with DL-based representations according to their abstraction levels. In the proposed framework, low-, mid-, and high-level representations are defined based on the progressive abstraction of the features and are established across both handcrafted and DL-based domains. Accordingly, handcrafted features are concatenated with CNN-derived representations at corresponding levels before concatenation. This level-wise cross-domain correspondence distinguishes the proposed framework from direct hybrid concatenation and from fusion strategies that combine multiple learned representations without explicit correspondence between handcrafted and learned features. Accordingly, this study proposes a multilevel feature concatenation framework for EEG-based emotion classification. A 2D-CNN is used to extract hierarchical representations from different network stages, which are categorized into low-, mid-, and high-level features according to their degree of abstraction.

Handcrafted EEG features derived from statistical-, time-, frequency-, and nonlinear domains are organized into corresponding feature levels. The matched handcrafted and CNN-derived representations are subsequently concatenated at each level to obtain a unified feature representation for emotion classification. The main contributions of this paper are listed below:

1. In this study, multi-domain handcrafted features: statistical, time-domain, frequency-domain, and nonlinear features are extracted from the preprocessed EEG signals and categorized into low- and mid-level representations.
2. A lightweight 2-D CNN model is tailored to extract comprehensive features from EEG that may be missed by traditional feature extraction techniques. These features are categorized into low-, mid-, and high-level abstractions.
3. Finally, handcrafted features and DL features are selectively concatenated at the relevant abstraction level, based on statistical significance analysis, to enhance the emotion classification performance.

II. Method

The details of the multilevel feature-based emotion classification technique using EEG signals are illustrated in Fig. 1, and the corresponding computational steps are presented in Algorithm 1. The model primarily consists of two parts: handcrafted feature extraction and DL-based feature extraction. In the final stage, these features are concatenated and classified using the support vector machine (SVM), K-Nearest Neighbors (KNN) and eXtreme Gradient Boosting (XG-Boost) classifiers.

A. Dataset Description

Two publicly available benchmark datasets for emotion recognition, a dataset for emotion analysis using EEG, physiological and video signals (DEAP) [17] and a database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-self-Devices (DREAMER) [18] are used in this study. The details of the datasets are given in Table 2. **Preprocessing and Segmentation:** For the preprocessed DEAP dataset, the initial 3-second baseline is removed, and the remaining 60-second EEG signal is segmented into twenty non-overlapping 3-second epochs. Each epoch is represented as Eq. (1) [19]:

$$E_i = \{x_1, x_2, \dots, x_L\}, i = 1, 2, \dots, 20, \quad (1)$$

Table 1. Related Work on EEG-based Emotion Recognition and mental health studies.

Study	Database Used	Feature Representation	Fusion/concatenation Strategy	Emotion Classes	Classifier/Model	Accuracy (%)
Bajaj et al. [7]	Custom database	Handcrafted Based on Multivariate decomposition	No Fusion	4	MC-LS-SVM	91.04
Anuragi et al. [8]	DEAP, SEED	Handcrafted (entropy and energy feature)	No Fusion	2 3	ANN, k-NN, Ensemble Bagged Tree with 10-fold CV and LOSOCV	77.9 95.7
Sharma et al. [3]	Custom database	Handcrafted Based on DWT	No Fusion	4	Bi-LSTM	82.00
Gupta et al. [9]	SEED, DEAP	Handcrafted FAWT-based	No Fusion	4 3	Random Forest, SVM	71.43 90.48
Khateb et al. [5]	DEAP	Handcrafted Time-, wavelet-, frequency-domain	No Fusion	9	SVM with 10-fold cross-validation and LOOCV	65.82
Wagh et al. [6]	SEED	Handcrafted Time- and frequency-domain	No Fusion	3	SVM	71.52
Sheykhivand et al. [10]	Custom database	DL-based	Multilevel DL-based feature fusion	2 3	CNN + LSTM	97.42 96.78
Gao et al. [11]	SEED	DL-based Based on (TFR) using CWT	Multilevel DL-based feature concatenation	3	MCNN	91.45
Tao et al. [12]	DEAP, DREAMER	DL-based	No fusion	2 3	ACRNN	93.52 98.04
Piedrabuena et al. [13]	InfoGain	Hybrid DL-based	Model fusion	3	Hybrid Model	98.59 (overall)
Nsugbe et al. [15]	Custom database	Handcrafted+ DL-based	Direct fusion	2	Decision Tree	98.00
Malekzadeh et al. [16]	Freiburg Bonn	Handcrafted +CNN	Direct fusion	2	CNN-RNN	99.45
Proposed	DEAP	Handcrafted+DL-based	Level-wise (multilevel)low /mid/high feature concatenation	(2,4)	SVM	(98.70, 93.07)
	DREAMER			(2,4)		(98.78, 93.85)

where, E_i denotes the i th EEG epoch, $L = 3 \times fs = 384$ samples, and $fs = 128$ Hz is the sampling frequency. Each epoch was normalized using Z-score normalization as Eq. (2) [19]:

$$x_i' = \frac{x_i - \mu}{\sigma}, \tag{2}$$

where, x_i is the EEG sample, x_i' is the normalized

sample, μ is the mean, and σ is the standard deviation (SD) of the corresponding epoch. This normalization ensures each epoch has a mean of zero and a variance of 1. Similarly, for the preprocessed DREAMER dataset, the last 180 seconds of each trial were selected and segmented into sixty non-overlapping 3-second epochs using the same segmentation process.

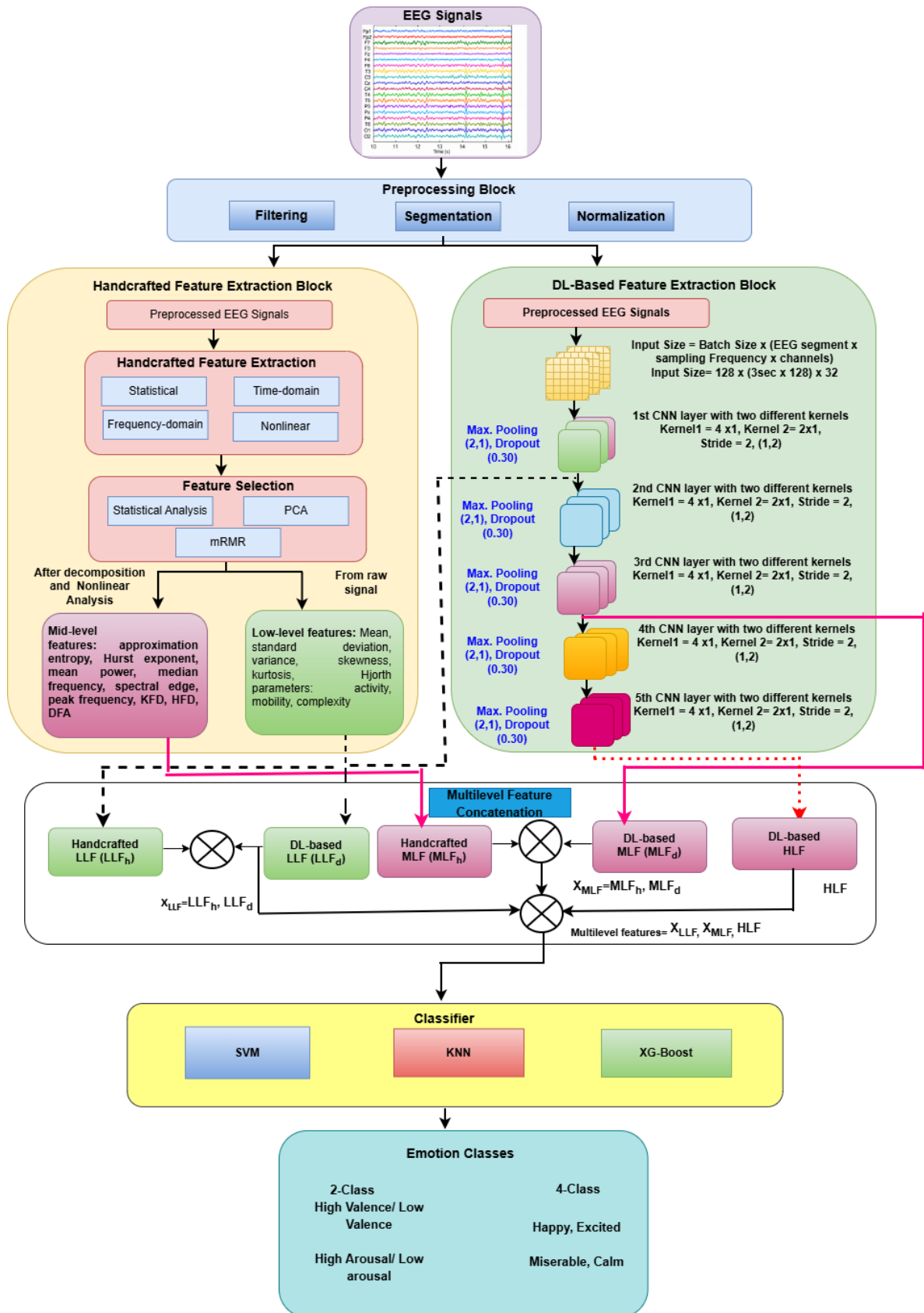


Fig. 1. Block Diagram of the Proposed Multilevel Feature Concatenation Framework for EEG-based Emotion Classification

Table 2. Description of DEAP and DREAMER datasets used for emotion classification

Parameters	DEAP dataset	DREAMER dataset
Number of subjects	32 (16 males + 16 females, average age range: 19–37 years)	23 (14 males + 9 females, average age range: 22–23 years)
Number of trials per subject	40	18
Collected signals	EEG, EOG, GSR, BVP, RSP, EMG, skin temperature	EEG and ECG
Number of EEG channels	32	14
Number of physiological peripheral signals	8	4
Signal duration	63 seconds	65–393 seconds (mean = 199 seconds)
Sampling frequency	128 Hz	128 Hz
Emotion rating	valence, arousal, dominance, like/dislike, and familiarity	Valence (V), arousal (A), dominance
Scale of emotion rating	1–9	1–5

Each epoch was normalized using the Z-score normalization described in Eq. (2). After segmentation, the class distributions of the DEAP and DREAMER datasets are presented in Table 3.

Emotion Label Generation: In this study, we focused only on arousal and valence ratings for both databases. Valence/arousal ratings for the DEAP dataset are 1-9, while those for the DREAMER dataset are 1-5 on the Self-Assessment Manikin (SAM) [20] scale. For the DEAP database, a rating <5 is considered low, and a rating ≥ 5 is considered high for both valence and arousal. Similarly, for DREAMAER dataset, ratings in the range of <3 is considered low, while ratings ≥ 3 are considered high for both valence and arousal planes.

Table 3. Class distribution after segmentation of DEAP and DREAMER datasets (H: High, L: Low, V: Valence, A: Arousal)

Dataset	HV– HA	LV– HA	HV– LA	LV– LA	Total
DEAP	9160	5920	5320	5200	25600
DREAMER	10560	7440	4620	2220	24840

B. Handcrafted Feature Extraction

After preprocessing, features from various domains: statistical, time, frequency, and nonlinear are extracted to capture more relevant information from the multichannel EEG data. In this study, mean (μ), standard deviation (σ), variance (σ^2), skewness, and kurtosis are extracted as statistical features. The mathematical formulations of the extracted statistical features are given in the equations (Eq. (3)-Eq. (7)) [21]:

$$Mean(\mu) = \frac{1}{N} \sum_{i=1}^N x_i \quad (3)$$

$$Variance(\sigma^2) = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2 \quad (4)$$

It measures the degree of dispersion of EEG signal amplitudes around their mean by computing the average of the squared deviations from the mean.

$$sd = \sqrt{\sigma^2} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2} \quad (5)$$

$$Skewness = \frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - \mu}{\sigma} \right)^3 \quad (6)$$

Skewness measures the asymmetry of the EEG signal amplitude distribution relative to its mean.

$$Kurtosis = \frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - \mu}{\sigma} \right)^4 \quad (7)$$

It represents the degree of peakedness and the presence of heavy tails in the EEG signal amplitude distribution. In the above equations, x_i represents the input EEG signal and N denotes the total number of EEG samples. In the time-domain analysis, the Hjorth parameters: activity (a_h), mobility (M_h), and complexity (C_h) are extracted from the EEG signals, and their mathematical formulations are presented in the equations (Eq. (8), Eq. (9), Eq. (10)) [22]:

$$a_h = \sigma^2(x_i) \quad (8)$$

It measures the variance of the EEG signal, representing its overall signal power.

$$M_h = \sqrt{\frac{\sigma^2(x'_i)}{\sigma^2(x_i)}} \quad (9)$$

It represents the rate of change of the EEG signals.

Here, x'_i denotes the 1st derivative of x_i .

$$C_h = \frac{M'_h}{M_h} \quad (10)$$

It measures the variation in frequency, reflecting the waveform irregularity of the EEG signal. Here, M'_h denotes the 1st derivative of M_h . To extract frequency domain features: alpha (α) (8–13 Hz), beta (β) (13–30 Hz), theta (θ) (4–8 Hz), delta (δ) (0.5–4 Hz), and gamma (γ) (30–44 Hz), are separated using Welch's

periodogram method. For each frequency band, power spectral density (PSD) and other frequency domain features: mean frequency, mean power, peak frequency, median frequency, and spectral edge are determined. The PSD of the input EEG signal $\mathbf{x}_i$ of frequency j is computed as Eq. (11) [23]:

Algorithm 1: Algorithm of the proposed multilevel feature concatenation framework

Input: Preprocessed labeled EEG dataset $X \in R^{(N \times T \times C \times L)}$, where N = number of subjects, T = number of trials, C = number of EEG channels, and L = number of samples per window (W); corresponding class labels Y

Output: Predicted emotion class labels $\hat{Y}$.

1. Initialize handcrafted feature set $\mathbf{h} = \{\emptyset\}$, deep feature set $\mathbf{d} = \{\emptyset\}$, and final multilevel feature set $\mathbf{F}_{Multi} = \{\emptyset\}$
2. For each subject $s = 1$ to N do
3. For each trial $t = 1$ to T do
4. Segment the EEG signal into non-overlapping 3-second windows (W).
5. For each segment (w) do
6. For each channel $c = 1$ to C do
7. Extract handcrafted features from the statistical, time-, frequency-, and nonlinear domains.
8. End For
9. $\mathbf{h}$ = handcrafted feature vector
10.
11. End For
12. End For
13. End For
14. Apply PCA, statistical analysis, and mRMR independently to the training portion of $\mathbf{h}$ to obtain the optimal handcrafted feature vector $\mathbf{h}'$.
15. Divide the selected ($\mathbf{h}'$) into low- and mid- level handcrafted features LLF_h and MLF_h
16. For each subject $s = 1, \dots, N$ do
17. For each trial $t = 1$ to T do
18. Segment the EEG signal into non-overlapping 3-second windows (W).
19. For each window $w = 1, \dots, W$ do
20. Construct the EEG segment: $X(s, t, w) \in R^{(32 \times 384)}$
21. Construct the CNN input tensor: $Xc(s, t, w) \in R^{(32 \times 384 \times 1)}$
where 32 represents EEG channels, 384 represents temporal samples, and 1 represents the singleton CNN input-channel dimension.
22. Feed $Xc(s, t, w)$ into the proposed 2D-CNN model.
23. End For
24. End For
25. End For
26. **Block 1:** $F1(s, r, w) \in R^{(16 \times 384 \times 32)}$, **Block 2:** $F2(s, r, w) \in R^{(8 \times 384 \times 64)}$, **Block 3:** $F3(s, r, w) \in R^{(4 \times 384 \times 128)}$, **Block 4:** $F4(s, r, w) \in R^{(2 \times 384 \times 256)}$, **Block 5:** $F5(s, r, w) \in R^{(1 \times 384 \times 512)}$
27. Extract:
Low-level deep feature (LLF_d) = Flatten($F1$), mid-level deep feature $[(MLF)]_d$ = Flatten($F3$), and high-level deep feature (HLF) = Flatten($F5$).
28. Compute the feature concatenation:
 $\mathbf{F}_{LLF} = \{LLF_h, LLF_d\}$,
 $\mathbf{F}_{MLF} = \{MLF_h, MLF_d\}$,
Multilevel Feature ($\mathbf{F}_{Multi}$) = $\{\mathbf{F}_{LLF}, \mathbf{F}_{MLF}, \text{HLF}\}$
29. Apply SVM, KNN, and XG-Boost on $\mathbf{F}_{Multi}$ with 10-fold and LOSO cross-validation
30. Compute accuracy, precision, recall, F1-score, balance accuracy for performance evaluation
31. Return the predicted class labels $\hat{Y}$

$$E_j = \lim_{i \rightarrow \infty} \frac{1}{i} |\hat{X}_i(j)|^2, \quad (11)$$

where, $\hat{X}_i(j)$ denotes the Fourier Transform of $\mathbf{x}_i$. Mean power represents the average spectral power within a specified frequency band. Median frequency is the frequency that divides the total spectral power into two equal halves. Peak frequency corresponds to the frequency at which the maximum spectral power occurs. Spectral edge frequency is the frequency below which 90% of the total spectral power is contained. Amongst the most common nonlinear features, the selected features are crucial in emotion recognition as they capture complex patterns and relationships present in EEG, leading to better and more robust identification of human emotions [24]. In this study, the nonlinear features: Hurst exponent (H), approximate entropy ($AppEn$), Higuchi fractal dimension (HFD), Katz fractal dimension (KFD), and detrended fluctuation analysis (DFA) are extracted. The details of these features are given as follow:

H , measures the long-term memory and persistence of the EEG signal, indicating the presence of long-range temporal correlations. It is represented as Eq. (12) [24]:

$$H = \frac{\log(R/\sigma)}{\log(N)}, \quad (12)$$

here, R is the range of cumulative deviations, σ is the standard deviation of the EEG signal, and N is the total number of sample points.

$$AppEn(S_N, m, r) = \ln\left(\frac{C_m(r)}{C_{m+1}(r)}\right), \quad (13)$$

$AppEn$, given in Eq. (13) [24] quantifies the regularity and complexity of the EEG signal. Here, S_N represents the EEG signal, m represents the pattern length, r is the similarity threshold, and $C_m(r)$ and $C_{m+1}(r)$ denote the average probabilities of finding similar patterns of length m and $m+1$, respectively. HFD is estimated from the relationship $L(k) \propto k$, where $L(k)$ is the average curve length of the EEG signal at interval k , and D represents the Higuchi fractal dimension (HFD) obtained from the slope of the $\log L(k)$ versus $\log(1/k)$ plot.

$$KFD = \frac{\log(n)}{\log(n) + \log \frac{d}{L}} \quad (14)$$

KFD represented in Eq. (14) [24] quantifies the geometric complexity of the EEG signal, where n is the number of steps, L is the total curve length, and d is the maximum distance between the first point and any other point in the signal.

$$Y(m) = \sum_{i=1}^m (x_i - \mu) \quad (15)$$

$$DFA = \sqrt{\frac{1}{N} \sum_{m=1}^N [Y(m) - Y_n(m)]^2} \quad (16)$$

DFA given in Eq. (16) [24] quantifies the long-range temporal correlations of the EEG signal. Here, x_i is the original EEG signal, μ is the mean of the EEG signal, $Y(m)$ given in Eq. (15) is the integrated signal profile,

$Y_n(m)$ is the local trend of the profile in a window of length n , and N is the total number of samples.

C. Selection of Handcrafted Features

Three different feature selection techniques: principal component analysis (PCA), statistical analysis, and minimum Redundancy Maximum Relevance (mRMR) are independently applied to determine the most relevant features from the handcrafted feature vector. The statistical feature selection method includes the Shapiro-Wilk (SW) or Kolmogorov-Smirnov (KS) test, which is first used in statistical analysis to determine the feature distribution. For feature significance a paired t-test is conducted for features with a normal distribution, and the Wilcoxon signed-rank (WSR) test is performed for features with a skewed distribution. To reduce redundant features, the Pearson correlation coefficient (PCC) with a threshold value of 0.80 is applied. Further, PCA is used to reduce the dimensionality of the data. The mRMR [25] method is used to select features that are highly relevant to the target variable while minimizing redundancy amongst the extracted features. Let $F = \{F_1, F_2, \dots, F_M\}$ denote the initial set of M candidate features extracted from the EEG signals. Here, F_j represents the j^{th} candidate feature, and Y denotes the corresponding emotion class label. During feature selection, S represents the set of features that have already been selected. Thus, an unselected candidate feature satisfies $F_j \in F/S$.

1. Maximum relevance:

The relevance of a candidate feature F_j with respect to the class label Y is measured using mutual information as Eq. (17) [25]:

$$D(F_j, Y) = I(F_j; Y), \quad (17)$$

where, $D(F_j, Y)$ is the relevance of feature F_j to the class label, and $I(F_j; Y)$ denotes the mutual information between F_j and Y . A higher value of $D(F_j, Y)$ indicates that the feature contains more information about the target emotion class.

2. Minimum redundancy:

The redundancy of a candidate feature F_j with respect to the already selected feature set S is calculated as Eq. (18) [25]:

$$R(F_j, S) = \frac{1}{|S|} \sum_{F_s \in S} I(F_j, F_s), \quad (18)$$

where, $R(F_j, S)$ denotes the average redundancy of a candidate feature F_j with the selected features, F_s denotes an individual feature in the selected set S , $I(F_j, F_s)$ represents the mutual information between F_j and F_s , and $|S|$ denotes the number of features currently present in S .

3. mRMR criterion

The mRMR score of a candidate feature is defined as Eq. (19) [25]:

$$mRMR(F_j) = D(F_j, Y) - R(F_j, S), \quad (19)$$

where $mRMR(F_j)$ is the mRMR score of candidate feature F_j . At each iteration, the candidate feature having the highest mRMR score is selected and added to S . Therefore, mRMR favors features that are highly relevant to the emotion class while being minimally redundant with the features already selected. Out of the above three feature selection methods, the optimal method is determined based on emotion classification performance using the SVM classifier on handcrafted features alone. For the DEAP dataset, out of 2016 extracted handcrafted features, mRMR, statistical analysis, and PCA retained 575, 592, and 457 features, respectively. Similarly, for the DREAMER dataset, out of 882 handcrafted features, mRMR, statistical analysis, and PCA retained 252, 332, and 229 features, respectively. The emotion classification results presented in Table 6 and Table 7 for handcrafted features using SVM on both datasets are based on mRMR-selected features, which achieved the highest performance compared to the other feature selection methods. Therefore, mRMR is selected as the final feature selection method for the proposed framework.

D. Extraction of DL-based (CNN) Feature

In this study, we implement a lightweight CNN model (including only five stages) for emotion classification. This compact architecture supports the effective integration with the proposed multilevel feature concatenation approach and is well suited for future embedded hardware-based implementations. In the proposed technique, the EEG signals are first transformed into a 2-D tensor and fed into the convolutional layers for feature extraction. These layers apply filters to the input EEG data, capturing essential patterns. Table 4 provides the detailed architecture of the proposed 2D CNN model. Also, it uses a batch size of 128, a 3-second window size, 250 epochs, and a learning rate of 10^{-4} . The output of the convolution layer is given by Eq. (20) [26],

$$Y = (W_k * X + b_k), \quad (20)$$

where, $X \in R^{(3 \times 128) \times 32}$, W_k and b_k are the weights of the kernel and bias, respectively. The output of the convolution operation is passed through a Leaky ReLU (δ) activation function, which is represented as Eq. (21) [26]:

$$Y_A = \delta(W_k * X + b_k), \quad (21)$$

where, δ is defined as:

$$\delta = \begin{cases} x_\alpha & \text{if } x_\alpha \geq 0 \\ \frac{x_\alpha}{a_\alpha} & \text{if } x_\alpha < 0 \end{cases},$$

here, x_α denotes the input value received from the convolution layer, and $\frac{x_\alpha}{a_\alpha}$ denotes negative slope coefficient of the LeakyReLU function. Further, to avoid

overfitting of the model, a dropout layer is applied. Finally, a dropout and a batch normalization (BN) layer are added to obtain the final feature map, which is represented by Eq. (22) [26]:

$$Y_{Fi} = BN(dropout(\delta(W_k * X + b_k))), \quad (22)$$

where, Y_{Fi} represents the feature maps of 5 different stages of the proposed 2D-CNN model, and BN is the batch normalization, which is defined as Eq. (23) [26]:

$$BN = \frac{x_j - \mu(x_i)}{\sigma^2(x_i) + \epsilon} * \text{gamma} + \text{beta}, \quad (23)$$

here, x_j represents an individual element within the input vector x_i , ϵ is a small constant added to prevent division by zero. Additionally, beta and gamma are trainable parameters that adjust the normalized values. Each selected feature map (from the 1st, the 3rd, and the last layers) is transformed into a 1-dimensional feature vector using a flattening operation. The flattened low-level, mid-level, and high-level feature vectors are subsequently concatenated with the optimized handcrafted EEG feature vector to construct the proposed multilevel feature representation for emotion classification. The trainable parameter counts across successive CNN stages are 2004, 16640, 66000, 210067, and 1000600, respectively, resulting in a total of 1295311.

E. Multi-Domain Feature Concatenation

Based on the complexity and nature of information they capture, handcrafted features are categorized as low- and mid-level features [26]. Basic features directly extracted from raw EEG signals: μ , σ , σ^2 , kurtosis, skewness, and Hjorth parameters: a_h , M_h , and C_h , are considered low-level features. On the other hand, some other features, spectral edge, mean power, median frequency, peak frequency, H, AppEn, HFD, KFD, and DFA, which involve decomposition, nonlinear analysis, and fractal- based transformation, are considered mid-level features. In the CNN architecture, the first layer extracts simple features and deeper layers capture complex features [27], [28]. Therefore, to uncover the domain specific knowledge and class discriminative features, features at multiple levels of the CNN have been extracted. The DL features are extracted after each pooling layer to have compressed features without losing meaningful information. Based on the statistical analysis in Table 5 and Fig. 3, the DL features are categorized as low-, medium-, and high-level features. The low-level features (F_{LLF}) from both the handcrafted and DL features are concatenated, followed by the concatenation of mid-level features (F_{MLF}) extracted from both. Finally, the combined LLF and MLF are further concatenated with the high-level features (HLF) to create a comprehensive feature representation as shown in Fig. 1. Mathematically, this process can be

Table 4. Architecture of the proposed DL model for emotion classification

Stage	Input Size	Layer Configuration	No. of Filters	Output Feature Map/Vector Size
Input	-	EEG epoch	-	(32×384×1)
Block 1	(32×384×1)	Conv2D (4×1), LeakyReLU, Dropout (0.30), Conv2D (1×2), Batch Normalization, MaxPooling (2×1)	32	(16×384×32)
Block 2	(16×384×32)	Conv2D (4×1), LeakyReLU, Dropout (0.30), Conv2D (1×2), Batch Normalization, MaxPooling (2×1)	64	(8×384×64)
Block 3	(8×384×64)	Conv2D (4×1), LeakyReLU, Dropout (0.30), Conv2D (1×2), Batch Normalization, MaxPooling (2×1)	128	(4×384×128)
Block 4	(4×384×128)	Conv2D (4×1), LeakyReLU, Dropout (0.30), Conv2D (1×2), Batch Normalization, MaxPooling (2×1)	256	(2×384×256)
Block 5	(2×384×256)	Conv2D (4×1), LeakyReLU, Dropout (0.30), Conv2D (1×2), Batch Normalization, MaxPooling (2×1)	512	(1×384×512)
Feature Extraction	(1×384×512)	Flatten	-	196608
Total Trainable Parameters	-	Complete CNN model	-	1295311

expressed as Eq. (24), Eq. (25), and Eq. (26):

$$\mathbf{F}_{LLF} = \{LLF_h, LLF_d\}, \quad (24)$$

$$\mathbf{F}_{MLF} = \{MLF_h, MLF_d\} \quad (25)$$

$$\text{Multilevel Feature } (\mathbf{F}_{Multi}) = \{\mathbf{F}_{LLF}, \mathbf{F}_{MLF}, HLF\} \quad (26)$$

where ‘,’ denotes concatenation of feature levels and the subscripts *h* and *d* indicate the handcrafted and DL features, respectively.

F. Emotion Classification and Hyperparameter Tuning

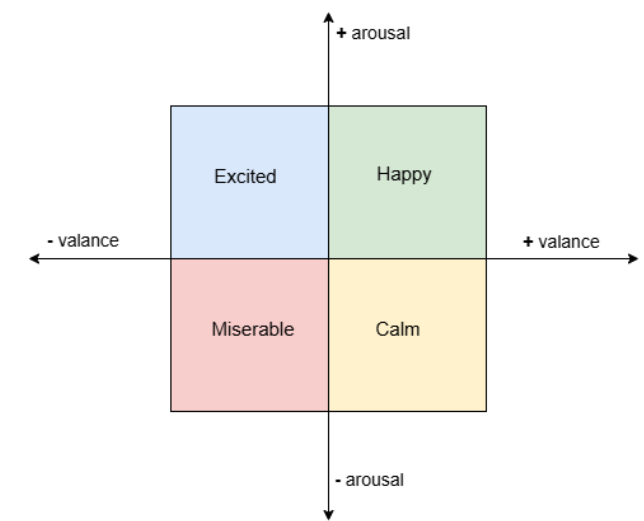


Fig. 2. Representation of emotion rating in 2-D valence-arousal (V-A) plane [31]

To classify 4-emotion classes (Happy, Excited, Miserable, and Calm) from the valence-arousal plane as shown in Fig. 2 and 2-emotion classes (valence (high/low) and arousal (high/low)) [29], [30], three different classifiers are used. For all three classifiers: SVM, KNN, and XGBoost hyperparameters are optimized using a grid search method. For SVM, the search space included: kernel = {linear, polynomial, RBF}, C = {0.1, 1, 10, 100}, and tolerance parameter (tol) = {0.1, 0.01, 0.001, 0.0001}. For KNN, n-neighbors = {3, 4, 5, 6} are considered. For XGBoost, n-estimators = {50, 100, 150, 200}, max-depth = {5, 7, 9, 11}, and learning rate = {0.001, 0.01, 0.1, 1} are used. The optimal hyperparameters selected are a polynomial kernel of 3rd degree, C = 1, and tol = 0.001 for SVM; n-neighbors = 4 for KNN; and n_estimators = 200, max. depth = 11, and learning rate = 0.1 for XGBoost. Hyperparameter tuning is performed using only the training data and independently in every training fold.

SVM Optimization and Classification Loss: The final multilevel feature vector is fed into the SVM classifier for emotion classification. The optimization problem is formulated as Eq. (27) and Eq. (28) [32] below:

$$\min_{w,b,\xi} \frac{1}{2} ||w||^2 + c \sum_{i=1}^N \xi_i, \quad (27)$$

Subjected to

$$Y_i(w^T \phi(F_{Multi,i}) + b) \geq 1 - \xi_i, \quad (28)$$

here, $\mathbf{F}_{Multi}$ is the final multilevel feature vector of the *i*th EEG sample, Y_i is its class label, *w* and *b* are the SVM weight vector and bias, ξ_i is the slack variable, *c* is the

regularization parameter, and $\phi(\cdot)$ denotes the kernel mapping. The corresponding hinge classification loss is given by Eq. (29) [32],

$$SVM_{Loss} = \frac{1}{N} \sum_{i=1}^N \max(0, 1 - Y_i f(F_{Multi,i})), \quad (29)$$

where, $f(F_{Multi,i}) = w^T \phi(F_{Multi,i}) + b$ is the SVM decision function. The polynomial kernel used in this study is determined as in Eq. (30) [32],

$$K(F_{Multi,i}, F_{Multi,j}) = (\gamma x_i^T x_j + r)^d, \quad (30)$$

where, γ , r , and d denote the kernel coefficient, independent term, and polynomial degree, respectively.

G. Cross-Validation Strategy

The model is evaluated using a subject-independent 10-fold and LOSO cross-validation. For 10-fold cross-validation, the data are divided into 10 subject-independent folds, where 9 folds are used for training and 1 fold for testing in each iteration, and the final performance is obtained by averaging across all folds. In addition, LOSO cross-validation is performed by excluding all epochs of each subject from the training set and using them exclusively for testing, with the procedure repeated until every subject has served as the test subject.

III. Results

In this study, all experiments are performed in the Jupyter Notebook IDE of Spyder 3.7.12 on a laptop equipped with a 1.80 GHz CPU, 8 GB MS RAM on Windows 10, and 4GB GPU RAM. To accelerate training and evaluation processes, the laptop's GPU is utilized. Specifically, TensorFlow GPU version 2.4.0 is installed, along with CUDA Toolkit version 11.0 and cuDNN version 8, which are compatible with the system's GPU. In addition, the PyTorch library is used to implement the proposed 2D-CNN model. The proposed study consists of three main experiments. In the first experiment, emotion classification is performed based on a selected handcrafted EEG feature set only. In the second experiment, DL features are extracted, and classification is performed. Lastly, in the third experiment, handcrafted and DL-based features are concatenated at multiple abstraction levels and classification is carried out. To assess the effectiveness of the proposed emotion recognition model, accuracy, Precision, Recall, and F1-Score [33], and balanced accuracy metrics are used in this study.

A. Study Significance of Feature Concatenation

To study the significance of different feature representations: handcrafted, DL-based, and concatenated (handcrafted and DL-based), SHapley Additive exPlanations (SHAP) [34], [35] analysis is conducted on the best-performing SVM classifier. SHAP importance scores are computed separately for

each dataset and then averaged across all datasets. The average SHAP values across all the datasets are presented in Table 5. The results show that, when multilevel features from both handcrafted and DL-

Table 5. Feature significance analysis for different feature sets and abstraction levels using SHAP

Feature Type	Feature Level	Average SHAP Value
Handcrafted (h)	Low	0.6792
	Mid	0.6895
	Combined	0.7125
DL-based (d)	Low	0.6675
	Mid	0.7112
	Combined	0.7442
Concatenated	Low (h and d)	0.7265
	Mid (h and d)	0.7415
	Low(h and d), Mid (h and d), High(h)	0.8246

based methods are concatenated, the average SHAP value reaches 0.8246, which is higher than using handcrafted and DL features alone.

B. Study Significance of Multilevel Feature Concatenation

To study the compatibility of feature concatenation at different levels, handcrafted low-level features are first concatenated with each level of DL-based features. Similarly, handcrafted mid-level features are concatenated with each level of DL-based features. Subsequently, mutual information (MI) analyses between handcrafted (h) and DL-based features at different abstraction levels are computed as Eq. (31) [25]:

$$MI(h, DL) = \sum_x \sum_y p(x, y) \log \frac{p(x, y)}{p(x)p(y)}, \quad (31)$$

where, $p(x, y)$ is the joint probability distribution of handcrafted and DL-based features, and $p(x)$ and $p(y)$ are their corresponding marginal probability distributions. MI measures the amount of shared information between the two variables.

Fig. 3 shows the MI between handcrafted and DL-based features at different abstraction levels. The results demonstrate that the MI between low-level handcrafted features and low-level DL-based features falls within the range of 0.60 to 0.95, the highest among all other combinations, i.e., low-level handcrafted and mid-level DL and low-level handcrafted and high-level DL. A similar result is also observed when mid-level handcrafted features are concatenated with mid-level DL features, showing an MI in the range of 0.64 to 0.91. Furthermore, when the handcrafted and DL-based features are concatenated at multiple levels with proper

level selection, the SHAP value is also the highest, i.e. 0.8246.

C. Ablation Study

1. 10-Fold Cross-Validation:

To investigate the effectiveness of each feature representation of the proposed framework, an ablation study is conducted. Table 6 and Table 7 present the emotion classification performance for 2-class and 4-class using handcrafted, DL-based, and multilevel features on the DEAP, DREAMER, and schizophrenia datasets with a 10-fold cross-validation protocol.

Results show that, for the DEAP dataset, the proposed multilevel feature set achieved the highest

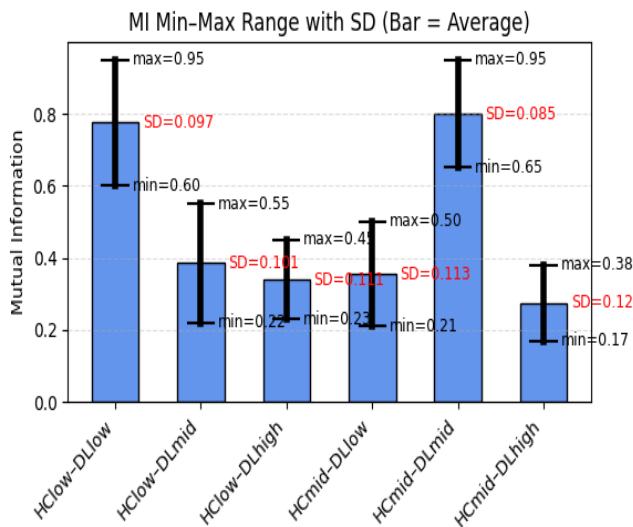


Fig. 3. Range of MI at different feature concatenation level

classification accuracies of 93.07% and 98.70% for the 4-class and 2-class emotion classification tasks, respectively, with corresponding precision, recall, and F1-scores of 0.92, 0.93, 0.91 and 0.97, 0.96, 0.96 using an SVM classifier. Compared with traditional handcrafted multidomain features, the proposed framework improved classification accuracy and precision from 72.72% to 93.07% and from 0.73 to 0.93 for the 4-class task, and from 79% to 98.70% and from 0.78 to 0.97 for the 2-class task, respectively. It also outperformed the MIMF feature by 19.84% and 12.40%, and the DL-based features by 4.23% and 6.67%, respectively, for 4-class and 2-class emotion classification using the SVM Classifier. Furthermore, results show that, among the evaluated classifiers, SVM achieved higher accuracies than KNN and XG-Boost for all feature representations. A similar trend is also observed for the DREAMER dataset. The proposed multilevel feature set achieved the highest classification accuracies of 93.85% and 98.78% for the 4-class and 2-class emotion classification tasks,

respectively, with corresponding precision, recall, and F1-scores of 0.91, 0.91, 0.92 and 0.97, 0.96, 0.96 using SVM classifier. Compared with traditional handcrafted multidomain features, the accuracies and precision increase by 20.68%, 19%, 17.78%, and 20% for the 4-class and 2-class emotion classification tasks, respectively. It also outperformed the MIMF feature by 4.15% and 5.15 %, and the DL-based features by 1.00% and 0.60%, respectively using an SVM Classifier for 4-class and 2-class emotion classes. Similar to the DEAP dataset, the SVM classifier consistently outperformed KNN and XG-Boost across all feature representations. To further investigate the generalizability of the proposed multilevel feature concatenation framework, previous experiments are conducted on an additional EEG-based schizophrenia dataset [37]. Although this dataset is intended for schizophrenia classification rather than emotion recognition, schizophrenia is characterized by altered neural activity and impaired emotion regulation. Therefore, it provides a meaningful benchmark for assessing whether the proposed multilevel feature can generalize across different EEG-based classification tasks while capturing neurophysiological patterns relevant to both neurological disorders and affective processing. Using the SVM classifier, the proposed framework achieved the highest classification accuracy of 98.02%, with a precision of 0.97, recall of 0.97, and F1-score of 0.98. Results represent an improvement of 8.02 % over handcrafted multidomain features, 0.95 % over DL-based features, and a slightly similar result to the MIMF [38] approach.

2. LOSO Cross-Validation:

Furthermore, LOSO cross-validation is performed to assess the subject-independent generalization of the proposed framework. Table 8 shows that for DEAP dataset, the proposed multilevel feature representation with SVM classifier achieved accuracies of 91.75% and 95.98% for 4-class and 2-class classification, respectively. Compared with KNN, the SVM achieved accuracy improvements of 2.82% and 0.22%, while improvements of 2.45% and 0.96% were obtained over XGBoost for the 4-class and 2-class classification settings. Similarly, for the DREAMER dataset, the proposed multilevel feature representation with SVM achieved accuracies of 91.75% and 96.88%, providing improvements of 2.49% and 3.65% over KNN and 2.16% and 1.63% over XG-Boost for 4-class and 2-class classification, respectively. For the schizophrenia dataset, the proposed multilevel feature with SVM achieved an accuracy of 96.12%, with precision, recall, and F1-scores of 0.95, 0.95, and 0.96, respectively. This result is 2.56% and 0.92% higher than the KNN and XG-Boost classifier, respectively. Table 9 summarizes the mean accuracy, SD and confidence

interval (CI) at 90% obtained using the 10-fold and LOSO cross-validation (CV) protocols. The proposed multilevel feature concatenation framework with SVM achieved the highest accuracy among the evaluated classifiers for both datasets. Under 10-fold CV, accuracies of 93.07% and 98.70% were obtained for 4-class and 2-class classification, respectively, for the DEAP dataset, while the DREAMER dataset achieved 93.85% and 98.78%. Under LOSO CV, the

corresponding accuracies were 91.45% and 95.98% for DEAP and 91.04% and 96.88% for DREAMER. The relatively low SD values (0.70-2.17%) with CI at 90% indicate stable and consistent performance across folds and subjects.

3. Class Imbalance Effect Study

Class imbalance may bias conventional accuracy toward the majority class and potentially overestimate

Table 6. Ablation study under 10-fold cross validation for classification of 2 and 4 emotion classes using DEAP and DREAMER datasets

Database	Feature Type	Classifier	Emotion Classes							
			4-Classes				2-Classes			
			Accuracy (%)	Precision	Recall	F1-Score	Accuracy (%)	Precision	Recall	F1-Score
DEAP	Multidomain	SVM	72.72	0.72	0.73	0.70	79.00	0.78	0.79	0.77
		KNN	69.70	0.68	0.69	0.68	76.18	0.73	0.76	0.75
		Boosting	71.01	0.70	0.69	0.71	77.90	0.78	0.76	0.76
	Handcrafted Features	SVM	78.23	0.79	0.78	0.76	86.30	0.87	0.86	0.85
		KNN	75.10	0.77	0.75	0.76	83.78	0.83	0.81	0.81
		Boosting	77.45	0.78	0.77	0.77	85.45	0.86	0.84	0.85
	Features using MIMF	SVM	88.97	0.87	0.88	0.87	92.03	0.91	0.93	0.91
		KNN	86.01	0.85	0.86	0.86	90.17	0.89	0.91	0.91
		Boosting	87.67	0.87	0.88	0.87	91.06	0.91	0.92	0.92
	DL-based Features	SVM	93.07	0.92	0.93	0.91	98.70	0.97	0.96	0.96
		KNN	90.20	0.89	0.90	0.89	96.16	0.95	0.95	0.95
		Boosting	91.70	0.91	0.91	0.90	97.12	0.96	0.96	0.95
DREAMER	Multilevel (Handcrafted + DL-based) Features	SVM	73.17	0.73	0.73	0.72	81.00	0.80	0.79	0.80
		KNN	70.60	0.69	0.70	0.70	78.97	0.77	0.76	0.78
		Boosting	72.10	0.71	0.72	0.71	79.87	0.78	0.79	0.79
	Features using MIMF	SVM	89.70	0.87	0.88	0.87	93.63	0.92	0.91	0.92
		KNN	86.95	0.86	0.87	0.86	91.23	0.90	0.89	0.89
		Boosting	88.16	0.88	0.88	0.88	92.16	0.91	0.90	0.90
	DL-based Features	SVM	93.75	0.91	0.93	0.91	98.18	0.97	0.96	0.97
		KNN	90.87	0.91	0.90	0.90	95.87	0.95	0.95	0.96
		Boosting	92.17	0.92	0.91	0.90	97.06	0.96	0.96	0.95
	Multilevel (Handcrafted + DL-based) Features	SVM	93.85	0.91	0.91	0.92	98.78	0.97	0.96	0.97
		KNN	91.16	0.90	0.89	0.90	96.13	0.95	0.94	0.94
		Boosting	92.19	0.91	0.91	0.90	97.15	0.96	0.95	0.95

classification performance. Therefore, this study also calculates balanced accuracy (BA) as Eq. (32) [39] to evaluate class-wise performance more fairly.

$$BA = \frac{1}{CL} \sum_{c=1}^C \frac{TP_{cl}}{TP_{cl} + FN_{cl}}, \tag{32}$$

here, CL is the total number of classes, TP_{cl} denotes the true-positive samples for class cl , and FN_{cl} denotes the

for the 4-class and 2-class settings on the DEAP dataset and 91.13% and 96.11% on the DREAMER dataset respectively. For the schizophrenia dataset, a balanced accuracy of 95.86 % is obtained for the 2-class setting. To study the statistical significance of our results, a two-way analysis of variance (ANOVA) test is conducted for performance comparison of different

Table 7. Ablation study under 10-fold cross validation for binary classification of schizophrenia dataset

Database	Feature Type	Classifier	Emotion Classes			
			Binary classes			
			Accuracy (%)	Precision	Recall	F1-Score
Schizophrenia	Multidomain	SVM	90.00	0.89	0.89	0.88
	Handcrafted	KNN	87.45	0.86	0.87	0.86
	Features	Boosting	88.97	0.87	0.88	0.88
	Features using MIMF	SVM	99.00	0.98	0.97	0.98
		KNN	96.13	0.95	0.95	0.93
		Boosting	97.56	0.96	0.97	0.97
	DL-based Features	SVM	97.07	0.95	0.96	0.95
		KNN	94.32	0.92	0.92	0.91
		Boosting	95.68	0.93	0.93	0.92
	Multilevel (Handcrafted + DL-based) Features	SVM	98.02	0.97	0.97	0.98
		KNN	95.46	0.95	0.93	0.93
		Boosting	97.10	0.96	0.95	0.96

false-negative samples for class cl . Balanced accuracy therefore gives equal weight to the recall of every class. The SVM classifier with 10-fold cross-validation achieved balanced accuracies of 91.75% and 96.00%

feature types across multiple datasets. A two-way ANOVA was performed by considering feature type and dataset as the two factors. The total variation is expressed as Eq. (33) [40]:

Table 8. Ablation study under LOSO cross-validation for 2-class and 4-class emotion classification using different datasets

Database	Feature Type	Classifier	Emotion Classes							
			4-Classes				2-Classes			
			Accuracy (%)	Precision	Recall	F1-Score	Accuracy (%)	Precision	Recall	F1-Score
DEAP	Multilevel (Handcrafted + DL-based) Features	SVM	91.45	0.91	0.90	0.89	95.98	0.95	0.93	0.94
		KNN	88.63	0.88	0.86	0.87	95.76	0.93	0.92	0.93
		Boosting	89.0	0.90	0.88	0.88	95.02	0.95	0.94	0.93
DREAMER	Multilevel (Handcrafted + DL-based) Features	SVM	91.75	0.89	0.90	0.90	96.88	0.95	0.94	0.95
		KNN	89.26	0.88	0.87	0.88	93.23	0.93	0.93	0.92
		Boosting	89.29	0.89	0.87	0.88	95.25	0.94	0.93	0.93
Sizophrenia	Multilevel (Handcrafted + DL-Bbased) Features	SVM	–	–	–	–	96.12	0.95	0.95	0.96
		KNN	–	–	–	–	93.56	0.93	0.91	0.91
		Boosting	–	–	–	–	95.20	0.94	0.93	0.94

$$SS_T = SS_A + SS_B + SS_{AB} + SS_E, \tag{33}$$

(CIs) for the 2-class and 4-class emotion classification

Table 9. Calculation of mean accuracy with standard deviation (SD) and CI at 90% under 10-fold, and LOSO cross-validation protocols for emotion classification using different datasets

Database	Cross-validation protocol	Feature Type	Classifier	Emotion Classes			
				4-Classes		2-Classes	
				Mean Accuracy (%)	±SD	Mean Accuracy	±SD
DEAP	10-fold	Multilevel Features	SVM	93.07	1.47	98.70	0.78
			KNN	90.20	2.42	96.16	1.17
			Boosting	91.70	2.19	97.12	0.98
DREAMER		Multilevel Features	SVM	93.85	1.51	98.78	0.70
			KNN	91.16	2.07	96.13	0.98
			Boosting	92.19	2.32	97.15	1.23
DEAP	LOSO	Multilevel Features	SVM	91.45	2.17	95.98	1.67
			KNN	88.63	3.04	95.76	1.89
			Boosting	89.00	3.18	95.02	2.01
DREAMER		Multilevel Features	SVM	91.04	2.02	96.88	0.98
			KNN	89.26	2.98	93.23	1.98
			Boosting	89.29	3.22	95.25	1.03

where, SS_A , SS_B , SS_{AB} , and SS_E denote the sum of squares due to feature type, dataset, their interaction, and error, respectively. The corresponding F-statistics were calculated as in Eq. (34) [40]:

$$F_A = \frac{MS_A}{MS_E}, F_B = \frac{MS_B}{MS_E}, F_{AB} = \frac{MS_{AB}}{MS_E} \tag{34}$$

where, MS denotes the corresponding mean square. A $p < 0.05$ is considered statistically significant. The test is applied separately to accuracy, precision, recall, and F1-score. The results show a significant difference in the means of accuracy, precision, recall, and F1-score across feature types ($p < 0.05$). Furthermore, the confusion matrices (CM) with 90% confidence intervals

tasks with 10-fold cross-validation on the DEAP and DREAMER datasets are presented in Fig. 4 and Fig. 5 to demonstrate the statistical robustness of the proposed framework.

D. Comparison Study

The comparison of the proposed multilevel feature concatenation framework with state-of-the-art methods [11], [15], [6], [34], and [13] evaluated on the DEAP and DREAMER datasets using a 10-fold cross validation protocol is given in Table 10. The proposed framework achieves emotion classification accuracies of 98.70% and 93.07% on the DEAP dataset, and 98.78% and

Table 10. Comparison study of the proposed method with existing methods for DEAP and DREAMER datasets

Method	Year	Dataset	Classifier	Cross-validation Protocol	Emotion Classes	
					2-Class	4-Class
MCNN [11]	2021	DEAP	SVM	10-fold	93.7	89.10
		DREAMER			92.37	88.79
Intelligent Combiner [15]	2021	DEAP	SVM	10-fold	91.14	88.07
		DREAMER			90.86	87.54
Time-frequency analysis [6]	2022	DEAP	SVM	10-fold	76.30	70.65
		DREAMER			75.70	71.25
Multimodal features [34]	2023	DEAP	SVM	10-fold	79.00	72.72
		DREAMER			81.00	73.17
hybrid BiGRU + BiLSTM [13]	2025	DEAP	SVM	10-fold	97.12	92.33
		DREAMER			97.45	93.32
Proposed Method	2026	DEAP	SVM	10-fold	98.70	93.07
		DREAMER			98.78	93.85

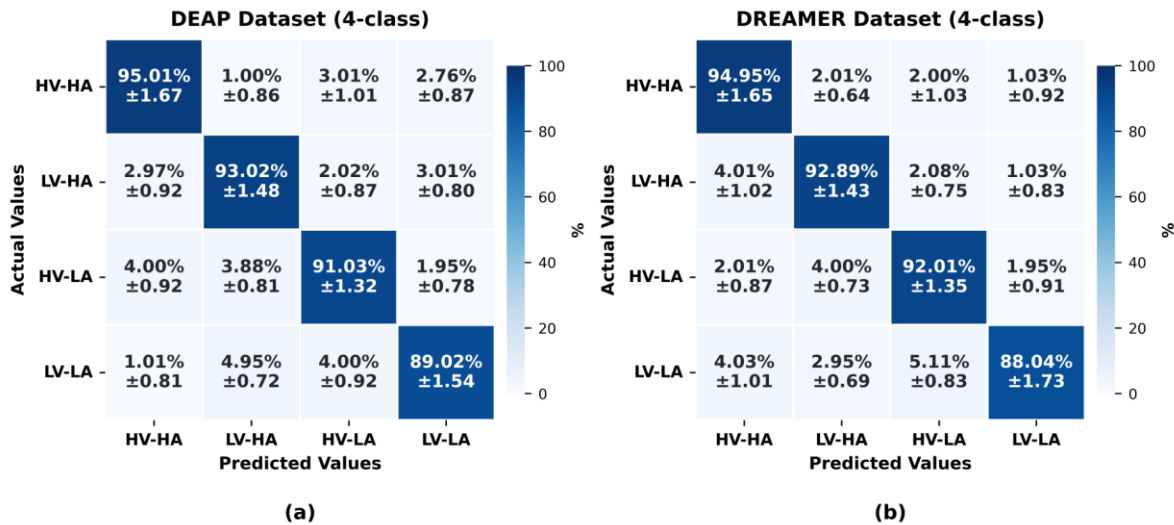


Fig. 4. Confusion Matrix (CM) for 4-class emotion classification for DEAP and DREAMER datasets

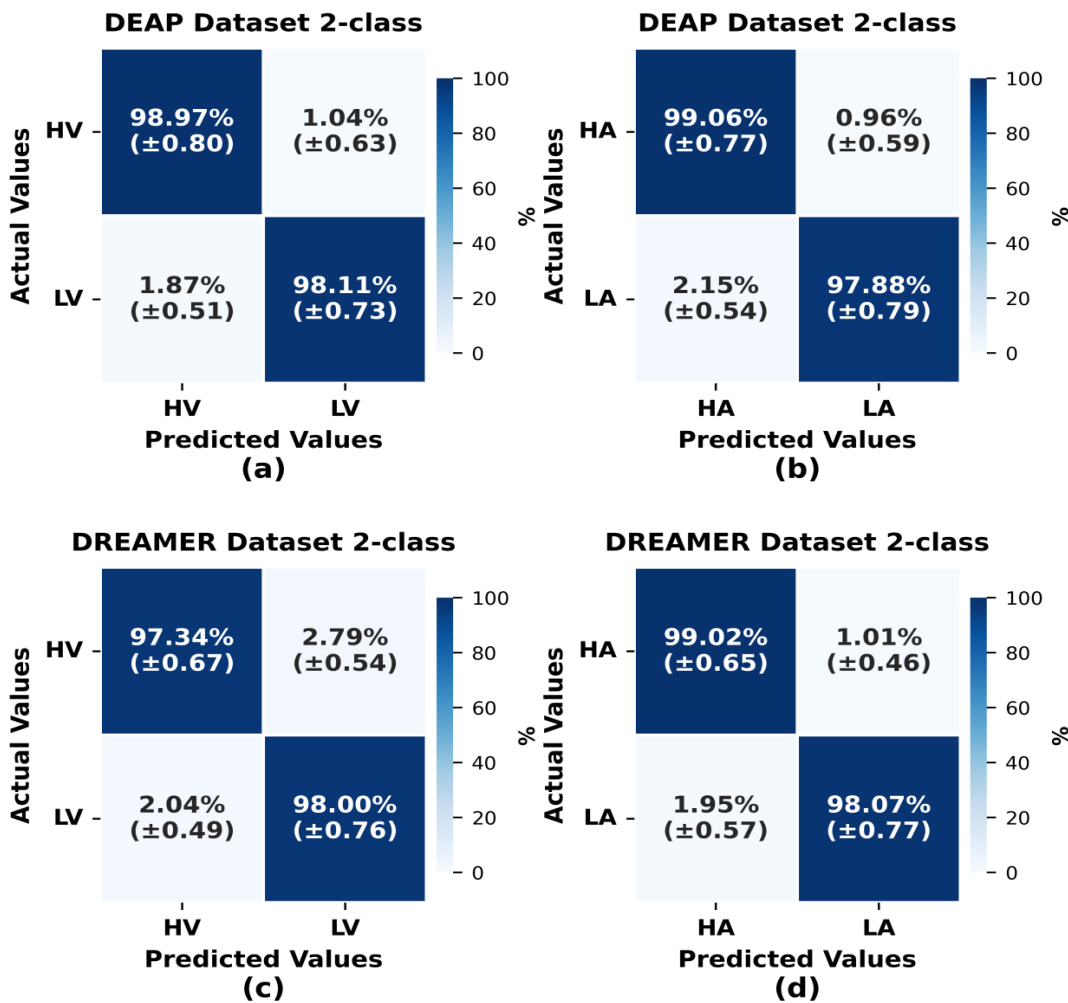


Fig. 5. Confusion Matrix (CM) for 2-class emotion classification for DEAP and DREAMER datasets

93.85% on the DREAMER dataset for 2-class and 4-class classification, respectively. Compared with the
closely related hybrid BiGRU + BiLSTM [13], the proposed method achieves improvements in

Table 11. Complexity Comparison of the proposed Model with existing models

Model	Architecture	FLOPs (M)	Parameters (K)	Inference Time (sec)	Accuracy (%)	Sensitivity	F1-Score
Conformer [41]	CNN-Transformer	32229	1297	40	93.02	0.91	0.90
EEGNex [42]	CNN	108389	95	66	87.67	0.86	0.87
Proposed Model	CNN	3100	1295	19	98.70	0.97	0.96

classification accuracy of 1.58 % and 0.74 % for 2-class classification and 1.33 % and 0.53 % for 4-class classification on the DEAP and DREAMER datasets, respectively. Compared with MCNN [11], the proposed framework achieves improvements of 5.00 % and 6.41 % for 2-class classification and 3.97 % and 5.06 % for 4-class classification on similar dataset and experimental setting. The proposed method also consistently outperforms the handcrafted feature-based approaches [6], [34] across both datasets and a similar validation protocol for 2-class, and 4-class classification setting. Similarly, compared with the Intelligent Combiner [15], which directly fuses handcrafted and DL-based features, the proposed framework achieves higher accuracies by 7.56 % and handcrafted and DL-based features, the proposed framework achieves higher accuracies by 7.56 % and 7.92 % for 2-class classification and by 5.00% and 6.31% for 4-class classification on the DEAP and DREAMER dataset, respectively. Furthermore, the receiver operating characteristic (ROC) analysis in Fig. 6 shows that the proposed multilevel features achieve the highest area under the curve (AUC) of 0.91. The complexity comparison of the proposed framework with existing models is presented in Table 11. Results show that the proposed CNN architecture demonstrates substantially reduced computational complexity (3100M FLOPs) compared to Conformer [41] and EEGNex [42]. Despite higher parameterization, it achieves a competitive inference time (19 sec), indicating an efficient balance between computational cost and practical deployability for EEG-based emotion recognition. In terms of classification performance, the proposed model achieves the highest classification accuracy of 98.70% with sensitivity of 0.97, outperforming Conformer (93.02%, 0.91) and EEGNex (87.67%, 0.86). It also attains a higher F1-score of 0.96, compared with 0.90 for Conformer and 0.87 for EEGNex.

IV. Discussions

A. Feature Concatenation Significance Analysis

Table 5 demonstrates that the proposed multilevel feature concatenation framework effectively combines handcrafted and DL-based EEG representations. The SHAP analysis further supports this observation, as the concatenated multilevel representation obtained an average SHAP value of 0.8246, higher than the individual handcrafted and DL-based feature representations. This indicates that when both handcrafted and DL-based features are concatenated, their overall contribution to the model's prediction increases.

B. Multilevel Feature Concatenation Significance Analysis

The MI analysis, as shown in Fig. 3 further explains the compatibility between handcrafted and DL-based features at different abstraction levels. The MI between low-level handcrafted and low-level DL-based features ranged from 0.60 to 0.95, which was higher than the other combinations. Similarly, the MI between mid-level handcrafted and mid-level DL features ranged from 0.64 to 0.91. Furthermore, integrating handcrafted and DL-

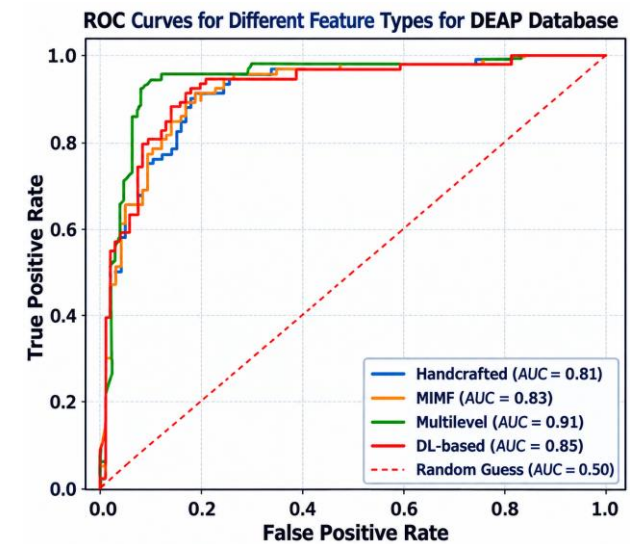


Fig. 6. ROC curves of 2-class emotion classification for different feature types using DEAP dataset

based features across multiple levels, with appropriate level selection, yields the highest SHAP value of 0.8246. This reestablishes the importance of multilevel feature concatenation with proper level selection as already hypothesized. A previous study [36] reported that different CNN layers progressively capture abstract representations, and combining features from these layers improves overall performance. However, they used DL-based features alone and did not incorporate handcrafted features at multiple abstraction levels. In contrast, our approach leverages both handcrafted and DL-based features across multiple levels of abstraction, resulting in a more informative and discriminative feature set that enhances model performance. Moreover, since the developed DL-based feature extraction model is lightweight, the overall approach is well-suited for real-world applications of emotion recognition and mental health monitoring.

C. Classification Performance and Generalizability

The ablation results obtained using both 10-fold and LOSO cross-validation, as shown in Table 6, Table 7, and Table 8, further demonstrate the effectiveness of the proposed feature representation. Across DEAP and DREAMER, the multilevel feature representation consistently produced high classification performance, while the SVM classifier achieved higher accuracies than KNN and XG-Boost across the evaluated feature representations. The schizophrenia dataset further demonstrates the generalizability of the framework across a different EEG-based classification task. Although this dataset is intended for schizophrenia classification rather than emotion recognition, the obtained results indicate that the multilevel feature representation can capture neurophysiological patterns relevant to EEG-based classification.

D. Classifier Performance and Stability

The consistent advantage of SVM can be attributed to its maximum-margin optimization, which effectively constructs optimal decision boundaries in the high-dimensional and discriminative feature space generated by the proposed multilevel feature representation. In contrast, KNN is more sensitive to overlapping feature distributions and the curse of dimensionality, while the tree-based partitioning strategy of XGBoost may not fully capture the continuous relationships among the fused EEG features. The relatively low SD values (0.70–2.17%) with CI at 90%, as shown in Table 9, indicate stable and consistent performance across folds and subjects. Compared with KNN and XG-Boost, the SVM provided accuracy improvements of up to 3.65% and 2.16%, respectively, demonstrating the effectiveness of the proposed multilevel feature representation SVM classifier.

E. Class Imbalance Effect Analysis

The balanced accuracy results further indicate that class imbalance has a limited influence on the overall performance of the proposed framework. The relatively small differences between accuracy and balanced accuracy suggest that the reported classification performance is not substantially dominated by the majority class. The consistently high balanced accuracy across datasets further indicates that the proposed multilevel feature representation provides robust class-wise discrimination under imbalanced conditions.

F. Comparison and Discussion

Although the compared methods in Table 10 use the same DEAP and DREAMER datasets, SVM classifier, and 10-fold cross-validation protocol, they differ in their feature representation and integration strategies. MCNN [11] primarily relies on deep convolutional features, whereas [6] and [34] emphasize handcrafted time-frequency and multimodal features, respectively. The Intelligent Combiner [15] combines handcrafted and DL-based features, while the hybrid BiGRU + BiLSTM method [13] focuses on sequential deep feature learning. In contrast, the proposed framework integrates handcrafted and CNN-derived representations at corresponding feature levels, enabling complementary low-, mid-, and high-level information to be jointly exploited. The consistent improvement over these approaches across both datasets and classification settings suggests that level-wise feature integration provides a more effective representation of heterogeneous EEG features for emotion classification. Additionally, the ROC analysis in Fig. 6 shows that the proposed multilevel features achieve an AUC of 0.91, further confirming their effectiveness for EEG-based emotion classification. Furthermore, Table 11 shows that the proposed framework provides a favorable balance between computational complexity and model size. Unlike Conformer, which incorporates Transformer-based processing, and EEGNex, which requires substantially higher computational operations, the proposed model maintains shorter inference time with a comparable number of parameters. This indicates an effective balance between computational efficiency and classification capability for EEG-based emotion recognition.

G. Limitations and future directions

The 10-fold cross-validation uses epoch-level partitioning; therefore, LOSO cross-validation is additionally performed to assess subject-independent generalization. The slightly reduced LOSO performance highlights the importance of subject-wise evaluation, which will be further investigated using larger and more diverse EEG datasets. Balanced accuracy is also reported to account for class imbalance and provide an equitable assessment of class-wise performance. The multilevel feature

representation increases feature dimensionality and computational requirements. Future work will investigate more compact feature aggregation and dimensionality-reduction strategies, along with automated learning of complementary handcrafted representations. Since the present evaluation uses controlled benchmark datasets [17], [18], validation on diverse populations and real-world EEG recordings is needed to further assess generalizability. SHAP and MI analyses provide model-level interpretability, while independent neurophysiological validation could further strengthen their physiological interpretation. Although the proposed framework is compared with relevant lightweight EEG models [41], [42], broader evaluation on emerging lightweight and hardware-oriented architectures will further assess practical applicability. Future optimization may support real-time and wearable EEG emotion recognition, while larger and more diverse populations will be important for evaluating its potential in objective affective-state assessment.

V. Conclusion

This study proposes a multilevel feature concatenation framework that integrates handcrafted EEG features with DL features extracted from different abstraction levels for emotion classification. The proposed framework generates a more discriminative feature representation and consistently outperforms individual handcrafted and DL feature sets as well as existing state-of-the-art methods. Among the evaluated classifiers: SVM, KNN, and XG-Boost, the SVM achieved the best performance, yielding classification accuracies of 98.70% and 93.07% on the DEAP dataset and 98.78% and 93.85% on the DREAMER dataset for 2-class and 4-class emotion classification, respectively. These findings demonstrate the effectiveness of the proposed multilevel feature concatenation framework for enhancing EEG-based emotion classification. Future work will focus on evaluating the proposed framework on larger subject-independent datasets, assessing its robustness under noisy wearable EEG conditions, optimizing its computational efficiency, and validating its real-time performance on embedded hardware platforms.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

The data used in this study are publicly available datasets. The details and access information for the datasets are provided in references [17] and [18].

Author Contribution

Kabita Devi conceptualized and designed the study, performed data preprocessing, feature extraction, model development, experimental analysis, and manuscript writing. Bhabesh Deka supervised the research work, conceptualized the methodology, contributed to methodology development, and critically reviewed the manuscript. Debarun Chakraborty contributed to the methodological mathematical formulation and reviews the write-up. All authors reviewed and approved the final version of the manuscript and agreed to be accountable for all aspects of the work, ensuring its integrity and accuracy.

Competing Interests

The authors declare no competing interests.

Ethical approval and consent to participate

Not applicable, as this study used only publicly available datasets and did not involve the collection of our own dataset.

References

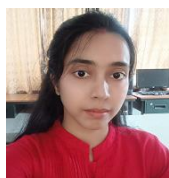
- [1] M. Jafari, A. Shoeibi, M. Khodatars, S. Bagherzadeh, A. Shalbaf, D. L. Garcia, J. M. Gorriz, and U. R. Acharya, "Emotion recognition in EEG signals using deep learning methods: A review," *Computers in Biology and Medicine*, vol. 165, pp. 107450, 2023, Elsevier, <https://doi.org/10.1016/j.combiomed.2023.107450>.
- [2] E. H. Houssein, A. Hammad, and A. A. Ali, "Human emotion recognition from EEG-based brain-computer interface using machine learning: A comprehensive review," *Neural Computing and Applications*, vol. 34, no. 15, pp. 12527–12557, 2022, Springer, <https://doi.org/10.1007/s00521-022-07292-4>.
- [3] R. Sharma, R. B. Pachori, and P. Sircar, "Automated emotion recognition based on higher order statistics and deep learning algorithm," *Biomedical Signal Processing and Control*, vol. 58, pp. 101867, 2020, Elsevier, <https://doi.org/10.1016/j.bspc.2020.101867>.
- [4] A. Nalwaya, K. Das, and R. B. Pachori, "An automated framework for human emotion detection from multichannel EEG signals," *IEEE Sensors Journal*, vol. 24, no. 13, pp. 20920–

- 20927, IEEE, 2024, <https://doi.org/10.1109/JSEN.2024.3398050>.
- [5] M. Khateeb, S. M. Anwar, and M. Alnowami, "Multidomain feature fusion for emotion classification using DEAP dataset," *IEEE Access*, vol. 9, pp. 12134-12142, IEEE, 2021, <https://doi.org/10.1109/ACCESS.2021.3051281>.
- [6] K. P. Wagh and K. Vasanth, "Performance evaluation of multi-channel electroencephalogram signal (EEG)-based time frequency analysis for human emotion recognition," *Biomedical Signal Processing and Control*, vol. 78, pp. 103966, Elsevier, 2022, <https://doi.org/10.1016/j.bspc.2022.103966>.
- [7] V. Bajaj and R. B. Pachori, "Human emotion classification from EEG signals using multiwavelet transform," in *Proceedings of the 2014 International Conference on Medical Biometrics*, IEEE, pp. 125–130, 2014, <https://doi.org/10.1109/ICMB.2014.29>.
- [8] A. Anuragi, D. S. Sisodia, and R. B. Pachori, "EEG-based cross-subject emotion recognition using Fourier–Bessel series expansion based empirical wavelet transform and NCA feature selection method," *Information Sciences*, vol. 610, pp. 508–524, Elsevier, 2022, <https://doi.org/10.1016/j.ins.2022.07.121>.
- [9] V. Gupta, M. D. Chopda, and R. B. Pachori, "Cross-subject emotion recognition using flexible analytic wavelet transform from EEG signals," *IEEE Sensors Journal*, vol. 19, no. 6, pp. 2266–2274, 2019, IEEE, [10.1109/JSEN.2018.2883497](https://doi.org/10.1109/JSEN.2018.2883497).
- [10] S. Sheykhivand, Z. Mousavi, T. Y. Rezaii, and A. Farzamnia, "Recognizing emotions evoked by music using CNN-LSTM networks on EEG signals," *IEEE Access*, vol. 8, pp. 139332–139345, IEEE, 2020, <https://doi.org/10.1109/ACCESS.2020.301188>.
- [11] Z. Gao, R. Li, C. Ma, L. Rui, and X. Sun, "Core-brain network-based multilayer convolutional neural network for emotion recognition," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–9, IEEE, 2021, <https://doi.org/10.1109/TIM.2021.3090164>.
- [12] W. Tao, C. Li, R. Song, J. Cheng, Y. Liu, F. Wan, and X. Chen, "EEG-based emotion recognition via channel-wise attention and self attention," *IEEE Transactions on Affective Computing*, vol. 14, no. 1, pp. 382–393, IEEE, 2020, <https://doi.org/10.1109/TAFFC.2020.3025777>.
- [13] N. Piedrabuena, W. M. Elmannai and Y. Wang, "Emotion Recognition From EEG Signals Using a Hybrid Deep Learning Approach," *IEEE Access*, vol. 14, pp. 11035-11047, IEEE, 2025, <https://doi.org/10.1016/j.bspc.2024.106620>.
- [14] S. Wang, X. Zhang, and R. Zhao, "Lightweight CNN-CBAM BiLSTM EEG emotion recognition based on multiband DE features," *Biomedical Signal Processing and Control*, vol. 103, pp. 107435, Elsevier, 2026, <https://doi.org/10.1109/ACCESS.2025.3642249>.
- [15] E. Nsugbe, O. W. Samuel, M. G. Asogbon, and G. Li, "Intelligence combiner: A combination of deep learning and handcrafted features for an adolescent psychosis prediction using EEG signals," in *Proceedings of the 2022 IEEE International Workshop on Metrology for Industry 4.0 and IoT*, pp. 92–97, IEEE, 2022, <https://doi.org/10.1109/MetroInd4.0IoT54413.2022.9831741>.
- [16] A. Malekzadeh, A. Zare, M. Yaghoobi, H. R. Kobravi, and R. Alizadehsani, "Epileptic seizures detection in EEG signals using fusion handcrafted and deep learning features," *Sensors*, vol. 21, no. 22, pp. 7710, MDPI, 2021, <https://doi.org/10.3390/s21227710>.
- [17] S. Koelstra, C. Mühl, M. Soleymani, J. S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2011, <https://doi.org/10.1109/T-AFFC.2011.15>.
- [18] S. Katsigiannis and N. Ramzan, "DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices," *IEEE Journal of Biomedical Health and Informatics*, vol. 22, no.1, pp. 98-107, IEEE, 2017, <https://doi.org/10.1109/JBHI.2017.2688239>.
- [19] K. Cabello-Solorzano, I. O. de Araujo, M. Peña, L. Correia, and A. J. Tallón-Ballesteros, "The impact of data normalization on the accuracy of machine learning algorithms: A comparative analysis," in *18th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2023)*, Lecture Notes in Networks and Systems, vol. 750, 2023, pp. 344-353, Springer, https://doi.org/10.1007/978-3-031-42536-3_33.
- [20] M. M. Bradley and P. J. Lang, "Measuring emotion: The self-assessment manikin and the semantic differential," *Journal of Behavior Therapy and Experimental Psychiatry*, vol. 25, no. 1, pp. 49–59, Elsevier, 1994, [https://doi.org/10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9).

- [21] R. Nawaz, K. H. Cheah, H. Nisar, and V. V. Yap, "Comparison of different feature extraction methods for EEG-based emotion recognition," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 3, pp. 910–926, Elsevier, 2020, <https://doi.org/10.1016/j.bbe.2020.04.005>.
- [22] 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, IEEE, 2014, <https://doi.org/10.1109/TAFFC.2014.2339834>.
- [23] I. Hussain, M. A. Hossain, R. Jany, M. A. Bari, M. Uddin, A. R. M. Kamal, Y. Ku, and J. S. Kim, "Quantitative of evaluation of EEG biomarkers for prediction of sleep stages," *Sensors*, vol. 22, no. 8, pp. 3079, MDPI, 2022, <https://doi.org/10.3390/s22083079>.
- [24] M. M. Rahman, A. K. Sarkar, M. A. Hossain, and M. A. Moni, "EEG-based emotion analysis using non-linear features and ensemble learning approaches," *Expert Systems with Applications*, vol. 207, pp. 118025, Elsevier, 2022, <https://doi.org/10.1016/j.eswa.2022.118025>.
- [25] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226–1238, IEEE, 2005, <https://doi.org/10.1109/TPAMI.2005.159>.
- [26] N. Gupta, C. K. Verma, et al, "Convolutional Neural Networks: Architecture, Feature Learning, and Evolution," SSRN, 2025.
- [27] L. Alzubaidi, J. Zhang, A.J. Humaidi et al., "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions," *Journal of Big Data*, vol. 8, no. 1, pp. 53, Springer, 2021, <https://doi.org/10.1186/s40537-021-00444-8>.
- [28] T. Talaei Khoei, H. O. Slimane, and N. Kaabouch, "Deep learning: Systematic review, models, challenges, and research directions," *Neural Computing and Applications*, vol. 35, no. 31, pp. 23103–23124, 2023, <https://doi.org/10.1007/s00521-023-08957-4>.
- [29] M. Kołodziej, A. Majkowski, P. Tarnowski, and R. J. Rak, "Recognition of visually induced emotions based on electroencephalography," in *2015 IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, vol. 2, pp. 701–704, IEEE, 2015, <https://doi.org/10.1109/IDAACS.2015.7341394>.
- [30] E. J. Choi and D. K. Kim, "Arousal and valence classification model based on long short-term memory and DEAP data for mental healthcare management," *Healthcare Informatics Research*, vol. 24, no. 4, pp. 309–316, Korean Society of medical Info., 2018, <https://doi.org/10.4258/hir.2018.24.4.309>.
- [31] J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161, American Psychological Association, 1980, <https://psycnet.apa.org/doi/10.1037/h0077714>.
- [32] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995, <https://doi.org/10.1007/BF00994018>.
- [33] D. M. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," *arXiv preprint arXiv: 2010.16061*, 2020.
- [34] K. Devi and B. Deka, "EEG-based emotion classification with per class feature importance using multi-domain features," in *Proceedings of the 2023 International Conference on Computational Intelligence, Networks and Security (ICCINS)*, IEEE, 2023, pp.1-6, <https://doi.org/10.1109/ICCINS58907.2023.10450043>.
- [35] S. Lundberg, "SHAP (SHapley Additive exPlanations)," GitHub repository. Available: <https://github.com/slundberg/shap>, 2019.
- [36] S. U. Amin, M. Alsulaiman, G. Muhammad, M. A. Bencherif, and M. S. Hossain, "Multilevel weighted feature fusion using convolutional neural networks for EEG motor imagery classification," *IEEE Access*, vol. 7, pp. 18940–18950, IEEE, 2019, <https://doi.org/10.1109/ACCESS.2019.2895688>.
- [37] E. Olejarczyk and W. Jernajczyk, "Graph-based analysis of brain connectivity in schizophrenia," *PLOS ONE*, vol. 12, no.11, pp. e0188629, Public Library of Science San Francisco, CA USA 2017, <https://doi.org/10.1371/journal.pone.0188629>.
- [38] K. Das and R. B. Pachori, "Schizophrenia detection technique using multivariate iterative filtering and multichannel EEG signals," *Biomedical Signal Processing and Control*, vol. 67, pp. 102525, Elsevier, 2021, <https://doi.org/10.1016/j.bspc.2021.102525>.
- [39] K. H. Brodersen, C. S. Ong, K. E. Stephan, and J. M. Buhmann, "The balanced accuracy and its posterior distribution," in *Proc. 20th International*

- Conference on Pattern Recognition (ICPR), 2010, pp. 3121–3124, IEEE, <https://doi.org/10.1109/ICPR.2010.764>.
- [40] J. Rydén and S. E. Alm, "The effect of interaction and rounding error in two-way ANOVA: Example of impact on testing for normality," *Journal of Applied Statistics*, vol. 37, no. 10, pp. 1695–1701, Taylor & Francis, 2010, <https://doi.org/10.1080/02664760903143925>.
- [41] Y. Song, Q. Zheng, B. Liu, X. Gao, "EEG conformer: Convolutional transformer for EEG decoding and visualization," *IEEE Trans. on Neural Syst. and Rehab. Eng.*, vol. 31, pp. 710–719, 2022, IEEE, <https://doi.org/10.1109/TNSRE.2022.3230250>.
- [42] X. Chen, X. Teng, H. Chen, Y. Pan, P. Geyer, "Towards reliable signal decoding for Electroencephalogram: A benchmark study to EEGNeX," *Biomedical Signal Processing and Control*, vol. 87, pp. 105475, 2024, Elsevier, <https://doi.org/10.1016/j.bspc.2023.105475>.

Author Biography



Kabita Devi is currently pursuing her Ph.D. at the Intelligent Imaging and Vision Research Laboratory (IIVRL), Department of Electronics and Communication Engineering, Tezpur University, Assam, India. Her research focuses on biomedical signal analysis using machine learning and artificial intelligence techniques. Her primary areas of interest include electroencephalogram (EEG)-based signal processing, EEG analysis during meditation practices, multimodal physiological signal analysis, emotion recognition, and mental health monitoring. She is also actively involved in physiological signal data collection, preprocessing, and experimental analysis for mental health assessment and emotion classification studies. She completed her Master of Technology (M.Tech.) degree from Tezpur University in 2020 with specialization in Bioelectronics from Dept. of ECE.



Dr. Bhabesh Deka is a Professor and Head of the Department (HoD) in the Department of Electronics and Communication Engineering at Tezpur University, Assam, India. He received his doctoral degree from IIT Guwahati in 2011 in the field of Image Processing. His research interests include Digital and Biomedical Signal/Image Processing, Machine and Computer Vision, Machine Learning and Deep Learning for

biomedical and precision agriculture applications, microwave and millimeter-wave biosensors for smart healthcare, and embedded machine learning on FPGA-based SoC applications. He has been a Senior Member of IEEE, USA since 2018 and a Fellow, IETE, India. He has published more than 60 research articles in reputed journals and conferences and authored/edited books in reputed international publications.



Dr. Debarun Chakraborty received his Ph.D. degree from Intelligent Imaging and Vision Research Laboratory (IIVRL), Department of Electronics and Communication Engineering, Tezpur University. His research focuses on computer vision and Deep Learning for precision agriculture, leveraging UAV-based sensing imagery and multimodal data fusion. He is particularly interested in designing optimized fusion architectures on CUDA-enabled GPGPU systems to advance intelligent models for both agriculture and biomedical applications. He has developed innovative solutions for crop management and EEG-based mental health disorder detection, significantly improving data analysis and interpretation. Currently, he serves as a Project Research Scientist-C at Tezpur University under an ICMR-funded project, focusing on multimodal deep learning using Ear-EEG and Photoplethysmography (PPG) to develop an AI-based mental health alert system integrated with smartphone applications.

Graphical Abstract

