Biomedical Signal Processing and Control 92 (2024) 106002

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Novel multimodal emotion detection method using Electroencephalogram and Electrocardiogram signals

Purnata Saha [a], Ali K. Ansaruddin Kunju [b], Molla E. Majid [c,*], Saad Bin Abul Kashem [d], Mohammad Nashbat [b,*], Azad Ashraf [b], Mazhar Hasan [b], Amith Khandakar [e,*], Md Shafayet Hossain [e], Abdulrahman Alqahtani [f,g], Muhammad E.H. Chowdhury [e]

[a] Department of Electrical and Electronics Engineering, University of Dhaka, Dhaka, Bangladesh
[b] Chemical Engineering Department, University of Doha for Science and Technology, Doha, Qatar
[c] Computer Applications Department, Academic Bridge Program, Qatar Foundation, Doha, Qatar
[d] Department of Computing Science, AFG College with the University of Aberdeen, Doha, Qatar
[e] Department of Electrical Engineering, Qatar University, Doha, Qatar
[f] Department of Medical Equipment Technology, College of Applied, Medical Science, Majmaah University, Majmaah City 11952, Saudi Arabia
[g] Department of Biomedical Technology, College of Applied Medical Sciences in Al-Kharj, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia

ARTICLE INFO

Keywords:
Emotion recognition system (ERS)
Electroencephalogram (EEG)
Electrocardiogram (ECG)
Photoplethysmography (PPG)
Time Domain (TD)
Frequency-domain (FD)
Time-frequency domain (TFD)
1D segmentation model
Machine Learning

ABSTRACT

Emotion Recognition Systems (ERS) play a pivotal role in facilitating naturalistic Human-Machine Interactions (HMI). The research has utilized a dataset with diverse physiological signals, including Electroencephalogram (EEG), Photoplethysmography (PPG), and Electrocardiogram (ECG), to detect emotions evoked by video stimuli. The study has addressed challenges with EEG data, particularly prefrontal channels contaminated by eye blink artifacts. To tackle this, a novel 1D deep learning model, MultiResUNet3p, effectively generated clean EEG signals. Extensive features have been extracted from each modality (TD, FD, TFD), and the study identified that combining 112 features from EEG and ECG achieved the highest accuracy. The emotion classification task encompassed six emotions, and the model demonstrates outstanding performance with 96.12% accuracy in binary classification (Positive vs. Negative) and 94.25% accuracy in a complex multiclass classification of six emotions (Happy, Anger, Disgust, Fear, Neutral, and Sad). This research underscores the potential of integrating multiple physiological signals and advanced techniques to significantly improve emotion recognition accuracy, particularly in real-world scenarios involving naturalistic Human-Machine Interactions (HMI).

1. Introduction

Emotions constitute a complex interplay of psychological and physiological components, encompassing subjective experiences, physiological responses, and behavioral expressions. Detecting emotions holds great promise for advancing human–computer interaction, personalized healthcare, and informed decision-making across diverse domains. This capability allows for the discernment and understanding of human emotions from various cues, enabling technology to respond empathetically and enhance user experiences. However, it also raises ethical concerns regarding privacy and responsible use[10].

Emotional shifts often coincide with physiological changes, including variations in heart rate, blood pressure, facial expressions, and other bodily indicators[16]. Contemporary researchers are actively involved in developing Emotion Recognition Systems (ERS) using cues such as facial expressions[7], speech[6], and physiological signals, including electroencephalogram (EEG)[12], temperature (TEMP)[21], electrocardiogram (ECG)[28], electromyogram (EMG)[2], galvanic skin response (GSR)[22], and photoplethysmography (PPG)[29].

Incorporating physiological data into Emotion Recognition Systems (ERS) offers distinct advantages over relying solely on observable facial, vocal, or behavioral cues. While individuals can manipulate or feign

* Corresponding authors.
E-mail addresses: purnatasaha.1999@gmail.com (P. Saha), Aliyarukunju.kunju@udst.edu.qa (A.K. Ansaruddin Kunju), mollamajid@gmail.com (M.E. Majid), saad.kashem@afg-aberdeen.edu.qa (S. Bin Abul Kashem), Mohammad.nashbat@UDST.edu.qa (M. Nashbat), Azad.ashraf@udst.edu.qa (A. Ashraf), mazhar.hasanzia@udst.edu.qa (M. Hasan), amitk@qu.edu.qa (A. Khandakar), shafayet.iut08@gmail.com (M. Shafayet Hossain), ama.alqahtani@psau.edu.sa (A. Alqahtani), mchowdhury@qu.edu.qa (M.E.H. Chowdhury).

https://doi.org/10.1016/j.bspc.2024.106002
Received 7 September 2023; Received in revised form 15 December 2023; Accepted 29 January 2024
Available online 13 February 2024

these visible signals, tampering with physiological data is significantly more challenging. Consequently, the inclusion of physiological indicators enhances the reliability and precision of emotion identification within ERS frameworks.

EEG has emerged as a promising avenue for emotion recognition, underpinned by its impressive performance[13]. The foundation of EEG-based emotion detection hinges on feature extraction techniques encompassing linear or nonlinear analyses within the temporal, spectral, or time–frequency domains. Sarma et al.[30]introduce a novel approach addressing the variability of emotional responses within EEG signals. By leveraging instantaneous phase synchronization measurements and a majority voting algorithm, they pinpoint potent EEG segments from pertinent channels. This technique attains a classification accuracy of up to 95 % for higher frequency sub-bands (β and γ) using 15 channels, marked by distinct segment locations for positive/neutral and negative emotions. Their approach exhibits enhanced efficiency through channel and data reduction in comparison to prevailing methods. Chanel et al. [8]distinguish three emotions (negative, positive, and neutral) using EEG features. The accuracy achieved with time–frequency EEG features stands at 63 %. However, combining EEG features with pairwise mutual information (MI) features and peripheral attributes elevates accuracy to 70 %. Post rejection of non-confident samples, accuracy further heightens to 80 %. Zheng et al.[42] evaluate six popular EEG-derived features for emotion recognition. The graph regularized extreme learning machine based on differential entropy (DE) emerges as the top-performing feature, averaging 69.67 % on the Database for Emotion Analysis using Physiological Signals (DEAP) dataset and 91.07 % on the SJTU Emotion EEG Dataset (SEED) dataset. Power spectral density (PSD) features extracted from all bands (delta, theta, alpha, beta, gamma) of specific channels (FT7, FT8, T7, T8, TP7, TP8) yielded machine learning model accuracies of around 76.19 % and 57.3 % on the DEAP and SEED datasets, respectively (Altamirano Asher [4]. Ding et al. [9]introduce a fresh EEG emotion recognition method using dispersion entropy (DispEn) for feature extraction across four frequency bands. Addressing imbalanced emotional labels, a random oversampling algorithm is employed. Results on the DEAP dataset reveal high recognition accuracy for both valence and arousal dimensions compared to other statistical features.

The inclusion of additional physiological signals alongside EEG has indeed yielded notable enhancements in specific instances; however, EEG remains unparalleled, consistently showcasing superior performance. This supremacy of EEG is exemplified in a study where EEG and speech signals were jointly employed to detect emotions within the Multi-modal Emotion Database[39]. Four baseline algorithms, namely Identification-vector and Probabilistic Linear Discriminant Analysis (I-vector + PLDA), Temporal Convolutional Network (TCN), Extreme Learning Machine (ELM), and Multi-Layer Perception Network (MLP), were formulated to evaluate the database and assess the efficacy of Automatic Emotion Recognition(AER) methods. Two fusion strategies, one at the feature-level and the other at the decision-level, were devised to harness both external and internal cues of human emotional status. The outcomes underscore the EEG's superiority, achieving higher accuracy in emotion recognition than speech signals (88.92 % in an anechoic room and 89.70 % in a natural noisy room, as opposed to 64.67 % and 58.92 %, respectively). Tong et al.[38] harness the DEAP dataset, encompassing EEG and peripheral physiological signals, to establish continuous wearable emotion recognition. Employing machine learning techniques and a reduced number of EEG channels coupled with one PPG signal, the study attains arousal and valence classification accuracies of 68 % and 66 %, respectively. This underscores the potential of employing wearable headbands for real-time emotion monitoring. In the realm of emotion recognition systems (ERS), both ECG and PPG signals hold significance, proving their efficacy in emotion detection across various investigations[27]. The growing popularity of wearable devices has bolstered the use of PPG signals in emotion detection[31]. Yang et al.[40] harnessed convolutional neural networks (CNN) to efficiently discern subjects' emotions, achieving a 75.4 % accuracy for 3 classes among 47 subjects. Shahid et al. [32]delved into four emotional states—disgust, fear, happiness, and sadness—employing frequency transformations, statistical measures, and multiple preprocessing techniques to extract 35 features from bioelectrical signals. Employing the Ensemble Bagged Trees classifier, their system achieved an 85.7 % accuracy, substantiating its adeptness in detecting the subject's emotional state.

In the realm of physiological signals, artifacts often distort signal quality. EEG recordings are particularly susceptible to various interferences like ocular, muscular, and cardiac artifacts. Mashhadi et al. [24] proposed deep learning-based denoising methods to eliminate electro-oculogram (EOG) from EEG. They deployed a 2D U-Net model, typically used for image segmentation tasks, to process EEG signals by rendering them as images. Independent component analysis (ICA), a common approach, is also applied for artifact removal in EEG[34]. Issa et al. introduced an automatic wavelet-based approach for selectively correcting EOG artifacts within EEG independent components, mitigating the limitations of discarding entire components. This method outperforms others in terms of accuracy, both temporally and spectrally, marking a significant stride toward precise EOG artifact eradication in EEG analysis [20]. Given that our dataset comprised prefrontal EEG data alone, abrupt EOG artifacts were evident across all channels. ICA's effectiveness was limited due to the confined prefrontal electrode coverage, hindering accurate ocular artifact capture. To address this, we employed a 1D segmentation model, MultiResUNet3p, integrated with deep supervision, for EOG artifact removal.

In light of the aforementioned EEG artifacts, our efforts focus on denoising the EEG signals by eliminating EOG artifacts. Subsequently, we have extracted features from the time domain, frequency domain, and time–frequency domain of EEG, PPG, and ECG signals. We have conducted an initial analysis using solely EEG features and then sequentially incorporated additional modalities to assess their impact on emotion detection outcomes. The following key contributions of this study are emphasized:

- A novel Deep learning-based approach to reconstruct the EEG signal while preserving its morphology by removing the EOG artifacts.
- Provide comprehensive multimodal performance analysis of EEG, ECG and PPG in detecting emotions.
- Investigate a large number of time, frequency, and time–frequency features for emotion recognition.
- Novel multimodal network for emotion recognition outperforming the state-of-the-art reported performance.

This article has been structured into four sections. In Section 2, we provide a comprehensive description of the datasets and the conceptual framework of our proposed methodology. In Section 3, the findings of this study are summarized. Finally, the concluding discussion and future prospects are given in Section 4.

2. Methodology

This section initiates with an outline of our methodology. Initially, we will provide a comprehensive explanation of the dataset utilized in our experimental study and the preparatory steps taken for data handling. Subsequently, we will delve into the intricate details of our devised technique for mitigating eye blink artifacts in EEG signals, followed by a thorough description of the features employed in emotion detection.

2.1. Overview

Our proposed approach's comprehensive structure is illustrated in Fig. 1. In the dataset, EEG, PPG, and ECG signals were concurrently recorded during exposure to diverse emotional video stimuli. These

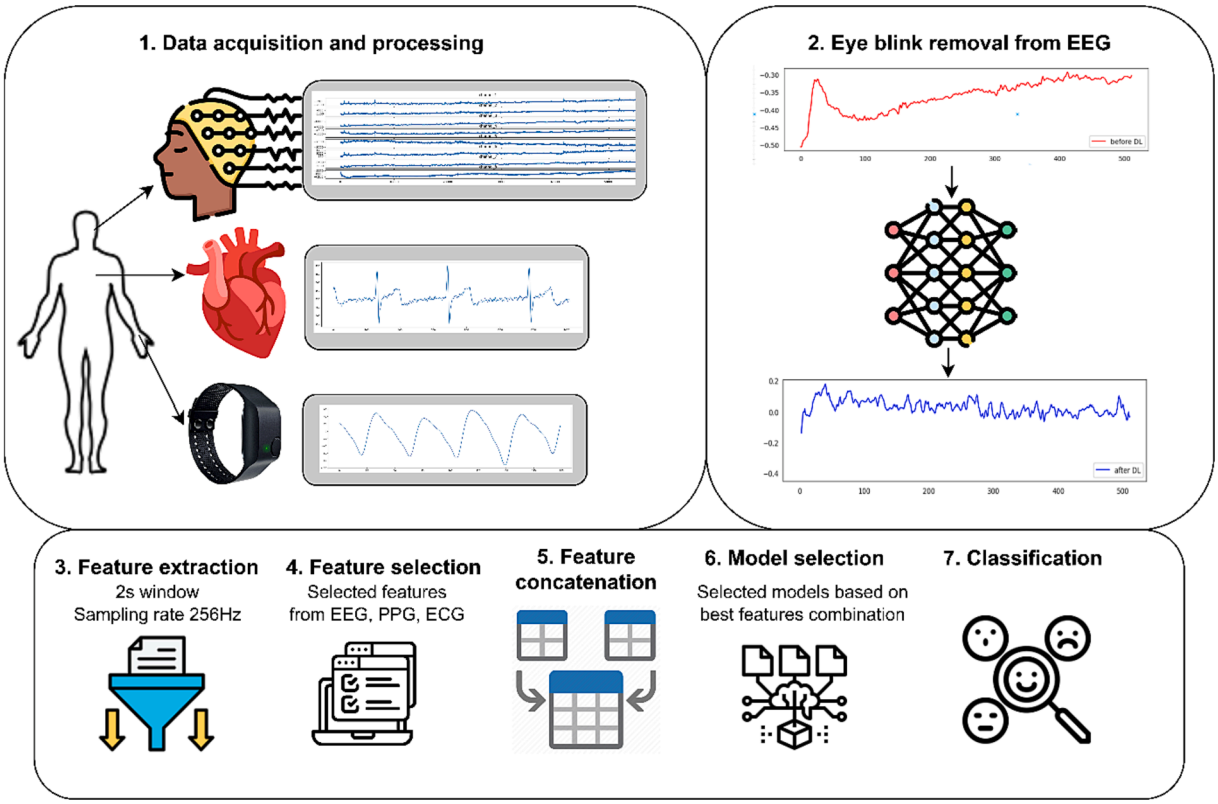

Fig. 1. General framework of our proposed methodology.

signals underwent several initial preprocessing stages, including resampling, filtering, segmentation, and normalization. Subsequently, a multitude of features from the time domain, frequency domain, and time–frequency domain were extracted from these signals. The optimal number of features was determined and then input into the most effective machine learning model for the purpose of emotion detection.

2.2. Dataset description

Utilizing the publicly available and extensive dataset titled 'ECSMP: A dataset on emotion, cognition, sleep, and multi-model physiological signals,' we addressed the emotion classification task[15]. The study involved a total of 87 healthy college students from Southeast University, encompassing 31 males and 56 females, with an average age of 23.68 ± 2.12 years. However, simultaneous recordings of EEG, PPG, and ECG for different emotions were not available for all 87 participants. Thus, we focused on the subset of 52 participants for whom we obtained simultaneous physiological signals corresponding to emotions. Prior to the experiment, participants were thoroughly briefed about the procedures and precautions, providing informed consent through signed documents. The Institutional Ethics Committee for Clinical Research of Zhongda Hospital, affiliated with Southeast University, granted approval for the study (Approval No. 2019ZDSYLL073-P01).

The experimental setup occurred in a tranquil, softly lit room, with participants seated at a 0.5-meter distance from the screen. They engaged in a task involving video-watching following a resting period. During the resting phase, participants were directed to fixate their gaze on a white cross displayed against a black screen for approximately 3 min, alternating between periods of having their eyes open and closed. The video -watching task aimed to elicit emotions and featured six distinct video clips, each thoughtfully selected to evoke specific emotional states. These clips included 'World Heritage In China' (neutral emotions), 'The Conjuring' (fear), 'Nuan Chun' (sadness), 'Top Funny Comedian' (happiness), 'Never Talk to Strangers' (anger), and 'The Fly' (disgust). These video choices were made with precision, drawing from prior research to consistently provoke the intended emotional responses.

Each participant was exposed to a set of randomized videos , each lasting around five minutes. The experimental procedure, as depicted in Fig. 2, followed this sequence. The emotional cycle model aided in identifying discrete positions of six emotions (neutral, fear, sad, happy, anger, and disgust) elicited by respective videos. While engaging in the video -watching task, participants' physiological signals were recorded, including EEG (Electroencephalogram), ECG (Electrocardiogram), PPG (Photoplethysmography), ACC (Accelerometer), EDA (Electrodermal Activity), and TEMP (Temperature). Given that each signal was captured by distinct devices, the sampling frequency differed across these signals. The devices' names and corresponding sampling frequencies are detailed in Table 1.

2.3. Preprocessing

Various forms of interference and distortions can affect biomedical signals, including baseline drift, motion artifacts, and power line noise [23]. In the datasets used in this study, these interferences impact EEG, PPG, and ECG signals. Specifically, EEG signals were collected from specific channels located in the prefrontal region of the head, namely Fp1, Fpz, Fp2, AF3, AF4, F7, F8, and A2. However, these channels were notably susceptible to Electrooculogram (EOG) artifacts, primarily originating from eye blinks, which led to significant disruptions in the EEG data quality. To counteract these challenges, a selection of effective preprocessing techniques was thoughtfully chosen and applied, alongside a sophisticated strategy aimed at preserving the signal's inherent morphology to the greatest extent possible. Given the multimodal nature of this physiological-based issue, distinct preprocessing steps were adopted for each modality.

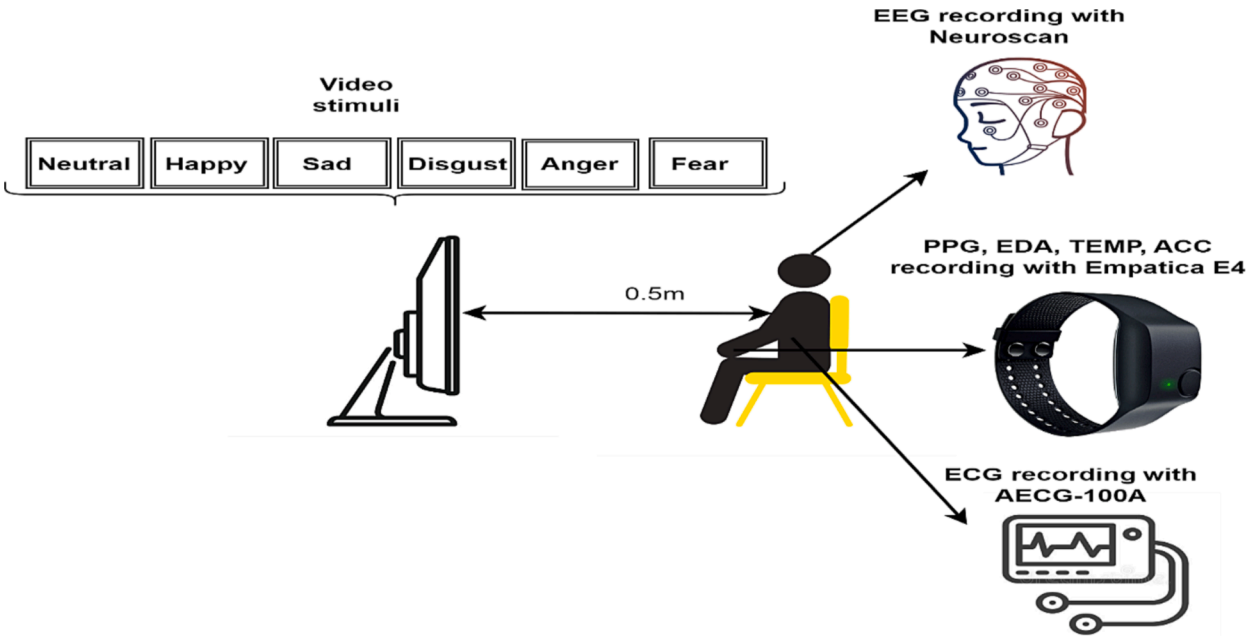

Fig. 2. Experimental setup [15].

Table 1
Devices used in data collection.

Signal Modality	Device used
EEG	Neuroscan Synamps RT
PPG, EDA, ACC, TEMP	Wristband Empatica E4
ECG	AECG-100A

2.3.1. EEG data preprocessing

Raw EEG signals are susceptible to noise, prompting the application of a 50 Hz notch filter to eliminate power line interference and reduce artifacts. During EEG analysis, the presence of baseline drift, manifested as undesirable amplitude shifts in the signals, posed a significant challenge. Unaddressed, these drifts could compromise the accuracy of results. To mitigate this issue, an approach involving polynomial fitting to the baseline drift was adopted, followed by its subtraction from the signal. This technique employed a polynomial of order 6, chosen based on trial and error. The signal was then resampled at 256 Hz to align with other modalities. Subsequently, the data was segmented using a window size of 512 (equivalent to 2 s), and z-score normalization was applied within each segment.

$$signal_i(norm) = \frac{signal_i - \mu_i}{\sigma_i} \tag{1}$$

here, μ_i and σ_i are the mean and standard deviation of the i^{th} segment, respectively.

Despite undergoing multiple processing stages, the persistence of the eye blink artifact within the data is evident in Fig. 3. Ordinarily, the Independent Component Analysis (ICA) technique is efficacious in eliminating Electrooculography (EOG) artifacts[33]. Nonetheless, within this specific dataset, the application of ICA did not yield satisfactory outcomes. The EEG channels employed in this study were situated exclusively in the prefrontal region of the scalp, which is particularly susceptible to EOG artifacts. Consequently, all the channels (Fp1, Fpz, Fp2, AF3, AF4, F7, F8, A2) were significantly influenced by EOG artifacts, particularly those arising from eye blinks. This scenario likely accounts for the inefficacy of ICA in mitigating the EOG artifact in this particular context. To rectify the issue of eye blink artifact elimination from the segmented EEG data, we employed a 1D segmentation model known as deep supervised MultiResUNet3p. This model harnesses deep learning techniques to generate EEG signals with removed eye blink artifacts, effectively eradicating the eye blink-related artifacts from the segmented data. A visual representation of EEG segments

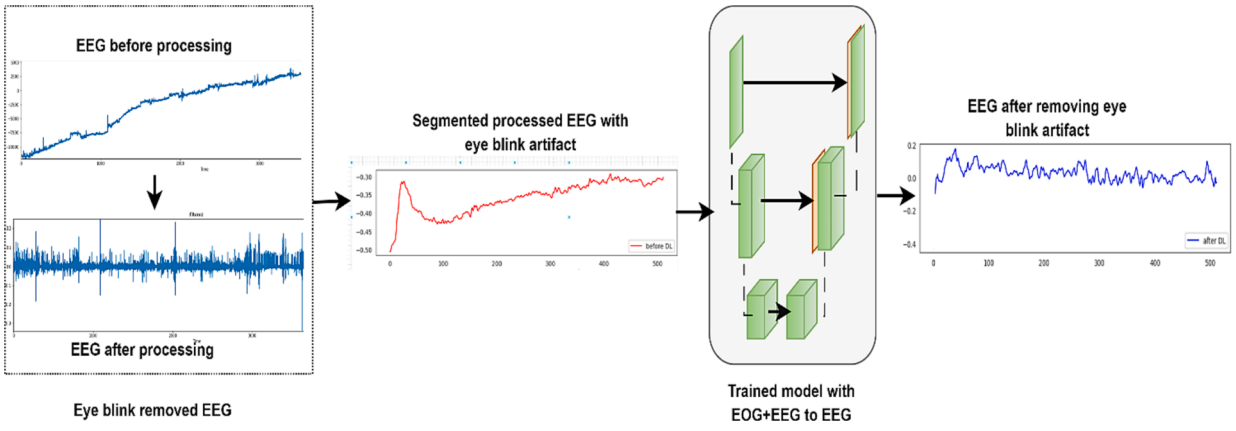

Fig. 3. EEG segment before and after Eye blink artifact removal.

before and after the removal of eye blink artifacts is depicted in Fig. 3.

2.3.2. Eye blink removal from EEG using 1D segmentation model

Built upon our earlier work, this approach draws from the successful implementation of the Multi-Residual UNet Model to eliminate electrooculogram (EOG) and electromyogram (EMG) artifacts from corrupted electroencephalogram (EEG) signals[17]. In this current study, we harnessed the EEGdenoiseNet dataset[41], a publicly accessible repository of meticulously curated physiological segments from diverse origins. Comprising 4514 pristine EEG segments, 3400 pure EOG segments, and 5598 clean EMG segments, this dataset served as our foundation.

By blending untainted EEG and EOG signals using techniques delineated by EEGdenoiseNet authors, we generated semi-synthetic noisy EEG segments. Respecting the original dataset's sampling rate of 256 Hz, we aligned our EEG data accordingly. From the pool of 4514 clean EEG segments, we crafted 34,000 semi-synthetic corrupted EEG segments spanning ten Signal-to-Noise Ratio (SNR) tiers. Notably, a subset of clean EOG segments was judiciously reintegrated during this amalgamation process. Of these segments, 27,800 were earmarked for training, while the remaining 6200 were dedicated for validation, meticulously preventing any potential data leakage. Upon validation, our approach demonstrated a remarkable temporal correlation of 90.21 % after undergoing 128 epochs.

With our trained model, we successfully generated EEG signals devoid of eye blink artifacts. The MultiResUNet3P model for segmentation efficiently amalgamates principles from both the MultiResUNet [19]and UNet3P[18] networks in a unified framework. Distinguishing itself from UNet3P, MultiResUNet3P incorporates residual paths in place of skip connections. The fundamental architecture of MultiResUNet3P is presented in Fig. 4.

UNet3P employs comprehensive skip connections, amalgamating interconnections between encoders and decoders alongside intra-connections among decoder sub-networks. However, while UNet3 + effectively integrates both small and large-scale feature maps within each decoder layer, it continues to utilize direct skip connections for both inter- and intra-connections. This introduces a semantic gap challenge in the foundational UNet framework, a concern that advanced architectures like UNet++ and MultiResUNet have sought to address. In response, MultiResUNet introduces diverse configurations of convolutional blocks to replace skip connections or similar structures. By doing so, the approach effectively reduces semantic gaps, enhancing the model's capability to acquire and generate EEG features, especially in scenarios where EEG signals are contaminated with EOG noise.

This Residual block is used in the architecture for inter- and intra-connections in MultiResUNet3P and is expected to reduce the semantic gaps by replacing direct skip connections. This particular ResPath represents the one produced from X^1_{EN} and X^2_{DE} among the inter- and intra- ResPaths, respectively, in Fig. 5.

2.3.3. PPG and ECG data processing

The original PPG data was collected at a sampling frequency of 64 Hz, while the ECG data was collected at 512 Hz. The ECG signal contained power line noise at 50 Hz, which was addressed by applying a notch filter with a center frequency of approximately 50 Hz to eliminate the noise. Subsequently, a higher-order Butterworth bandpass filter was used with cutoff frequencies of 0.05 Hz to 60 Hz for ECG signals. The PPG data did not exhibit power line noise, allowing for the application of a bandpass filter with cutoff frequencies of 0.5 Hz to 10 Hz. Both PPG and ECG signals were then resampled to 256 Hz. The data was segmented into 2-second intervals without any overlap. Additionally, the segmented data was normalized using the z-score normalization formula to ensure consistency and comparability across the dataset.

2.4. Feature Extraction

Feature extraction from EEG signals during each subject's action interval involved utilizing a range of methodologies. Time-domain (TD) features, as introduced by Tkach et al. [37], encompassed pure TD attributes, statistical parameters (such as mean, deviation, absolute value), and features derived from the signal envelope. Frequency domain (FD) features, based on the work of Altın and Er [5], described the signal's frequency domain characteristics, including dominant frequency, spectral roll-off point, spectral deformation, skewness, and kurtosis of the power spectrum. Time-frequency domain (TFD) features were obtained through the discrete wavelet transform (DWT) and included attributes like band power, mean absolute value, waveform length, root mean square value, standard deviation, and fractal length of the extracted wavelets, in accordance with Stergiou et al. [36].

In total, each EEG channel segment yielded 50 TD features, 24 FD features, and 117 TFD features, culminating in 191 features per channel. The dataset encompassed data from 8 EEG channels, amounting to a cumulative total of 1528 features (191 × 8). Similarly, for ECG, we also extracted 191 features using the same methodology. Concerning PPG,

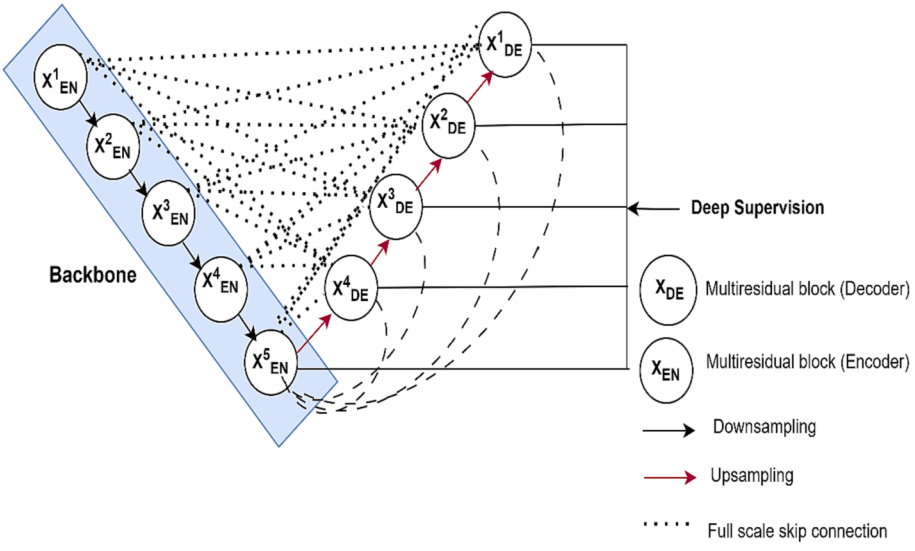

Fig. 4. Architecture of MultiResUNet3P . adapted from [17]

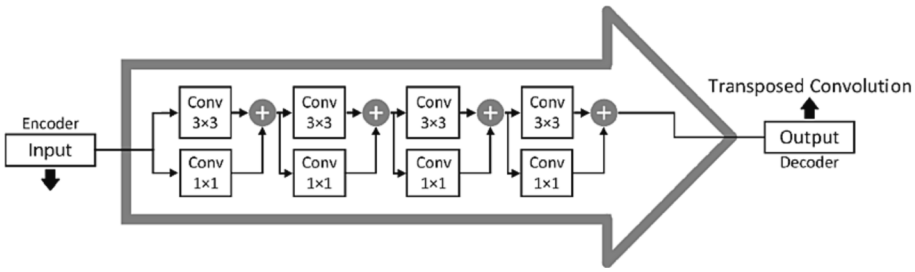

Fig. 5. Residual Path architecture .
adapted from [19]

93 time-domain features and 14 statistical features were extracted. The statistical features were computed from the preprocessed signal, while the time-domain features were extracted from the PPG signal along with its first and second derivatives, following the approach outlined in. The comprehensive details of the features from different domains are furnished in Table 2. However, the substantial number of features could potentially introduce elevated variance in predictions when using machine learning models, underscoring the necessity for employing feature selection techniques.

Table 2
TD, FD and TFD Features used for emotion detection (Nazmul Islam [26]).

Time domain		Frequency domain	Time-Frequency domain
1. Enhanced mean absolute value. 2. Enhanced waveform length. 3. Mean absolute value. 4. Waveform length. 5. Zero crossing. 6. Slope sign. 7. Root mean square. 8. Avg. amplitude change. 9. Difference in absolute standard deviation. 10. Zero-order moment. 11. Log detector. 12. Modified mean absolute value. 13. Total energy. 14. Sparseness. 15. Myoplasm percentage rate. 16. Irregularity factor. 17. Simple square. integral.18. Waveform length ratio (WLR) .	19. Variance. 20. Signal Kurtosis. 21. Auto regressive coefficients. 22. Sigma signal entropy. 23.Estimate signal entropy. 24. Standard deviation. 25. Peak to peak. 26. Willison amplitude. 27. Max to mean ratio. 28. Number of biases. 29. Descriptor of signal entropy. 30. Shape factor. 31. Integrated absolute value. 32. V order is the square root of variance. 33. Maximum fractal length. 34. Average power. 35. Shape factor.	1. Mean frequency. 2. Median frequency. 3. Spectral entropy. 4. Max frequency.5. Power spectral (PS) skewness. 6. Max value. 7. PS kurtosis. 8. PS mean frequency. 9. PS median frequency. 10. Signal to artifact ratio. 11. SNR. 12. Max to min drop of PS. 13. Spectral centroid. 14. Spectral flatness. 15. Spectral slope. 16. Spectral spread. 17. Spectral crest. 18. Spectral deformation. 19. Spectral roll-off. 20. Spectral flux. 21. Spectral skewness. 22. Spectral decrease. 23. Max ratio.	1. Band Power of wavelets of the signal. 2. Standard Deviation of wavelets of the signal. 3. Mean Absolute Value of wavelets of the signal. 4. Fractal Length of wavelets of the signal. 5. Waveform length of wavelets of the signal. 6. Entropy for levels. 7. Root Mean Square of wavelets of the signal.

2.5. Feature selection

After feature extraction, the process of feature selection has gained significance, as the extensive number of features could potentially lead to a substantial decrease in model performance due to overfitting. Keeping this in consideration, we delve into the exploration of the following feature selection techniques:

1. Decision Tree-Based Selection[3]
2. Principal Component Analysis (PCA)[1]
3. ReliefF[35]
4. Least Absolute Shrinkage and selection operator (LASSO)[25]

The feature selection procedure was executed independently for EEG, PPG, and ECG features, given the utilization of signals from these three modalities. Subsequently, the chosen features from each modality were amalgamated, followed by an additional round of feature selection aimed at identifying the optimal number of features to yield the most favorable outcomes.

2.6. Classification algorithm

The machine learning models underwent 5-fold cross-validation. We have used 4 folds for training and 1-fold for testing each time. The distribution of signal segments for training and testing is outlined in Table 3**.** For training purposes, 80 % of the recordings were employed, and another 20 % of the 33,732 samples were assigned for testing. Feature extraction was carried out after splitting the data. The selected features from training samples were used to train and an unseen 20 % test set per fold was used to predict the emotional state. The average Accuracy and F1 score of 5 runs are presented as the final performance results.

As previously mentioned, our study involved the analysis of signals from three distinct modalities, with the best performance achieved through the combination of EEG and ECG features. After optimizing the extracted features, we proceeded to evaluate various classical machine learning algorithms, including Support Vector Machine (SVM), XGBoost Classifier (XGB), Extra Trees Classifier (ETC), Random Forest Classifier (RFC), MLP Classifier, and K-Nearest Neighbors (kNN) Classifier. These algorithms were systematically assessed, and their overall performance outcomes were documented. This research showcases the implementation of two distinct classification schemes using the dataset: i) Binary emotion classification (Positive vs. Negative emotion) and ii) Six emotion classification.

Table 3
Distribution of Training and Test set.

Train	26,986
Test	6746

2.7. Data augmentation

To counteract the challenge of an imbalanced dataset, which can introduce bias to the model, we employed data augmentation techniques during the training phase. The Synthetic Minority Oversampling Technique (SMOTE) was utilized to rectify the dataset imbalance by generating synthetic samples for the minority class[11]. SMOTE is an oversampling technique aimed at alleviating overfitting problems linked to random oversampling. It functions within the feature space, generating new instances by resampling from neighboring positive instances. The process involves selecting a random minority class instance and identifying its k nearest neighbors (where k = 5 in this case). One neighbor is then randomly chosen, and a synthetic sample is created along the line connecting the selected instance and its neighbor. The implementation of SMOTE involves enriching the dataset with these synthetic samples, effectively augmenting the minority class and achieving a more balanced distribution among the various classes. This augmentation strategy is beneficial for building a robust model by addressing class imbalance. Notably, SMOTE was exclusively applied to the training dataset to ensure unbiased training. Before applying SMOTE, the training set comprised 26,986 samples, and post-augmentation, the count rose to 32,652 samples, with each class containing 5,442 segments.

2.8. Evaluation criteria

The evaluation of classifier performance employs two metrics: accuracy and F1 score. These metrics are derived from the values of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). Notably, due to a balanced dataset across all subjects, accuracy and F1 score are closely aligned, and either metric can effectively represent the predictive performance of a classification model in this context. The metrics employed in this study are detailed as follows:

1) Accuracy: Accuracy quantifies the proportion of correct predictions made by a classifier. It is calculated as the ratio of accurate predictions to the total number of predictions. This can be expressed through Equation (2), where accuracy is computed by dividing the sum of True Positives and True Negatives (representing correct predictions) by the sum of True Positives, True Negatives, False Positives, and False Negatives (representing all predictions).

$$Accuracy(\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (2)$$

2) F1 Score: The F1 score evaluates a model's classification accuracy by accounting for both its precision (correct positive predictions) and its ability to correctly identify positive instances (recall). Calculated as the harmonic mean of precision and recall, the F1 score offers a comprehensive assessment of model performance. In this case, the F1 score is computed using Equation (3), which employs True Positives, False Positives, and False Negatives.

$$F1Score(\%) = \frac{2 \times TP}{2 \times TP + FP + FN} \times 100\% \quad (3)$$

3. Results and discussion

This study aims to assess the effectiveness of feature selection techniques and the importance of different modalities in detecting emotions. Through comprehensive evaluations of multiple machine learning algorithms in various classification scenarios, it was found that the Random Forest algorithm exhibited superior performance in initial assessments. The subsequent analysis compares the outcomes of different feature selection methods for both six-class and two-class classification scenarios. The evaluation commenced by focusing on EEG performance and subsequently incorporated other modalities progressively to augment the overall detection outcomes.

3.1. Combination of different modalities

Initially, we embarked on identifying optimal EEG features post feature extraction. Employing decision tree-based feature selection methods, which exhibited superior performance, we adopted them for generating optimal outcomes. Following feature extraction, approximately 1528 EEG features were derived from 8 channels. Through trial and error, we narrowed this down to 120 features using the RF-based feature selection method. Solely utilizing these selected EEG features, our model achieved notable accuracies of 90.23 % for six-class classification and 91.41 % for two-class classification. Subsequently, we turned our attention to ECG feature selection, incorporating the top 50 features alongside EEG's best features. A combined total of 170 features led to the selection of around 112 features, resulting in enhanced performance – reaching 94.25 % accuracy for six-class classification and 96.12 % for two-class classification. Incorporating PPG data, we selected 65 features from a pool of 107 PPG features and combined them with EEG features. Furthermore, we integrated EEG and PPG data, selecting 100 features from the merged dataset. In this case, while the improvement over baseline EEG performance was marginal, the combined EEG and PPG features achieved an accuracy of 91.23 % for six-class classification and 94.05 % for binary emotion classification. Combining features from 3 modalities (EEG, PPG, ECG) we achieved 92.56 % accuracy for six-class classification and 93.71 % for binary emotion classification. A comprehensive overview of the multimodal data's performance can be found in Fig. 6.

3.2. Binary and multi-class classification

Table 4 illustrates the performance outcomes of diverse models in both the two-class and six-class classification scenarios. The classification repertoire encompassed SVM, ETC, RFC, XGBoost, AdaBoost, MLP, and KNN, all utilized for distinguishing between various emotion categories. In the two-class classification, the distinction was made between happy versus all negative emotions, while the six-class classification encompassed emotions such as disgust, happy, fear, sad, angry, and neutral. For each context, the models were trained separately and subsequently evaluated on the test set. In the six-class classification, a single situation entailed neutral, happy, fear, sad, anger, and disgust emotions. The models were appropriately trained and assessed for each scenario. Notably, the highest accuracy achieved in two-class emotion detection was 96.12 %, accompanied by an F1 score of 96.73 %. In the context of the six-class emotion classification, the obtained accuracy stood at 94.25 %, with an accompanying F1 score of 93.45 %.

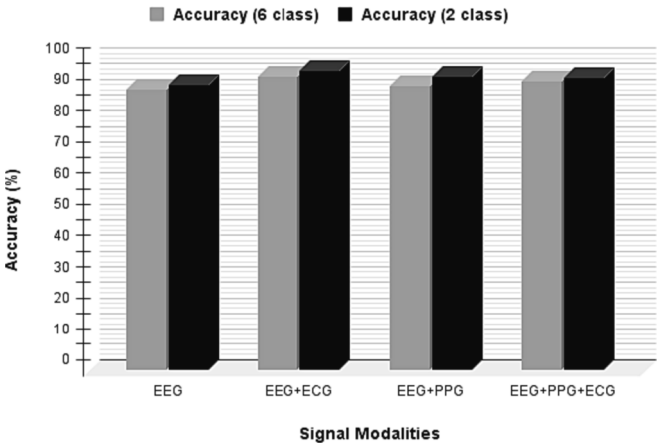

Fig. 6. Performance of different modalities in detecting emotion.

Table 4
Accuracy (in percentage) of binary and six-emotions classification.

Model	Six-emotions classification	Binary emotions classification				
		Fear	Anger	Disgust	Sad	Neutral
RFC	**94.25**	95.67	95.23	96.09	**96.12**	94.78
ETC	94.10	93.43	95.36	94.25	95.89	92.38
MLP	71.18	71.65	76.40	73.70	79.42	70.43
SVM	67.36	66.45	63.87	65.23	68.25	65.39
XGB	93.18	96.02	91.69	95.73	95.31	92.32

3.3. Selection of feature optimization technique

This study assesses various feature selection techniques, encompassing Tree Based Selection through Extra Trees Classifier (ETC) and Random Forest Classifier (RFC), Principal Component Analysis, ReliefF, and LASSO, in conjunction with machine learning models for classification purposes. The summarized outcomes of feature selection and classification are presented in Table 5. For this evaluation, the top 100 features are selected from the most effective feature combination (EEG + ECG) to gauge the performance of each feature selection method. The results from the model with the most favorable overall performance are showcased. Notably, the analysis reveals that Tree Based Feature selection demonstrates superiority, achieving an overall accuracy of 90.67 % and 94.23 % in six-class and two-class emotion detection, respectively, using an RFC model. ReliefF and LASSO techniques also exhibit promise, attaining overall accuracies of 85.45 % and 82.34 % in six-class emotion detection, respectively.

The EEG-measured features of six emotions were calculated for the EEG segments. The correlation coefficients between the top 4 features and emotions are shown in Table 6**.** Top features with lower p values and higher correlation justify the importance of features and strong relationship selected features and emotion.

3.4. Performance improvement by MultiResUNet3p

As previously mentioned, the performance of EEG significantly improved with the application of the 1D segmentation model MultiResUNet3p to eliminate eye blink artifacts. In this section, we present a comparative analysis of EEG features before and after the removal of eye blink artifacts. We also explored the use of Variational Mode Decomposition (VMD) initially to address EOG artifacts. Although VMD effectively mitigated horizontal eye movement artifacts, complete removal of eye blink artifacts was not achieved using this method. Fig. 7 illustrates EEG segments before and after applying both VMD and deep learning (DL) techniques. Initially, with conventional processing steps, EEG accuracy reached approximately 79.25 %. After implementing VMD, accuracy improved to 89.56 % for two-class emotion detection and 83 % for six-class emotion classification. Upon utilizing MultiResUNet3p to eliminate eye blink, accuracy further increased to 90.23 % for two-class and 91.41 % for six-class classification. An attempt to combine VMD and DL resulted in performance degradation due to the loss of crucial components.

Table 5
Performance of different feature selection method.

Feature Selection Method	Model Name	Accuracy (%)		F1 score (%)	
		Binary	Six-class	Binary	Six-class
Decision Tree-Based Selection using ETC	ETC	93.89	90.23	95.02	90.35
Decision Tree-Based Selection using RFC	RFC	94.23	90.67	95.43	91.24
PCA	ETC	76.13	68.20	75.47	69.35
LASSO	RFC	86.94	82.34	87.49	82.06
ReliefF	ETC	87.88	85.45	87.29	84.13

Fig. 7 clearly illustrates a noticeable improvement in the signals' quality after the application of the 1D segmentation model. The EEG signals post-Deep Learning (DL) exhibit a significantly enhanced quality compared to their pre-DL counterparts. To assess the statistical significance of this performance enhancement, a *t*-test was conducted on the predicted emotions before and after applying DL, yielding a p-value of 0.0126. Given that this p-value is less than the conventional significance level of 0.05, it can be concluded that the MultiResUnet3p has indeed led to a statistically significant improvement in the results.

Table 7 provides a summary of various studies that have employed physiological signals from diverse modalities for emotion recognition. Upon reviewing the table, it becomes evident that a significant portion of previous research has utilized EEG data for emotion detection. In some instances, the combination of other physiological signals like ECG and PPG, or PPG and GSR, has demonstrated promising outcomes. In the current study, we conducted an analysis using the ECSMP open-source database, featuring multimodal data encompassing six distinct emotions. Previously, Gao et al. [14]showcased the efficacy of MODFA-a1, FE, and PSD features extracted from the gamma band, achieving a 96.81 % accuracy in binary classification and a 42.17 % accuracy in six-emotion classification. In our research, we extend this by incorporating physiological signals from additional modalities to enhance performance, resulting in a substantial improvement in the accurate detection of multiclass emotions.

4. Conclusion

The practical importance of emotion detection systems lies in their ability to enhance human–computer interaction, tailor content and services, and improve mental health support. ERS can aid in the early detection of mental health issues by identifying signs of stress, depression, or anxiety. Emotion detection can be used in security applications, such as lie detection during interviews or border control, to assess emotional responses.

Emotional biometrics can enhance security by using unique emotional patterns for user identification. This study centered on the detection of emotions using a publicly accessible dataset that encompassed diverse physiological signals. The dataset, comprising data from 52 subjects, underwent preprocessing and feature extraction across various domains. Different feature selection methods were assessed to identify the most pertinent features for training machine learning models. The investigation evaluated the performance of different modalities for emotion detection, with EEG serving as the baseline and other modalities subsequently integrated to enhance effectiveness. Overcoming the challenge of eliminating eye blink artifacts from EEG signals was complex due to the prefrontal nature of all channels. To address this, we employed a novel technique, the deep supervised 1D MultiResUNet3P, resulting in significant performance enhancement. The proposed approach achieved competitive results akin to state-of-the-art techniques on this dataset, achieving 94.25 % accuracy for six-class and 96.12 % for two-class emotion detection. Notably, while prior research predominantly excelled in binary classification, our approach adeptly handled the intricate task of multiclass classification. It should be noted, however, that PPG data for most subjects were severely distorted, consequently not yielding performance improvements when incorporated as a modality. The potential for enhancing multimodal performance hinges on reconstructing PPG signals from the distorted data. Successfully achieving this reconstruction could potentially elevate the overall efficacy of the multimodal approach.

Funding

This work was made possible by student grant # QUST-1-CENG-2023–941 from the Qatar University and is also supported via funding from Prince Sattam Bin Abdulaziz University project number (PSAU/

Table 6
Correlation analysis between EEG-measured features and ground truth emotion.

Features	Neutral		Anger		Happy		Sad		Fear		Disgust	
	r-value	p-value	r-value	p-value	r-value	p-value	r-value	p-value	r-value	p-value	r-value	p-value
Enhanced***WavLen Ch1_EEG	0.09	0.004	0.021	<0.001	−0.01	0.002	0.04	<0.001	0.03	0.004	0.02	<0.001
PsSkewness***Ch1 EEG	−0.12	0.012	0.31	0.005	0.30	<0.001	0.12	0.007	0.19	<0.001	0.02	<0.001
MaxRatio_ECG	−0.01	0.02	0.02	<0.001	0.02	<0.001	−0.4	<0.001	−0.06	0.03	0.02	0.002
Kurtosis_ECG	0.09	<0.001	0.04	0.006	0.09	0.002	−0.1	0.01	0.05	0.007	0.02	0.004

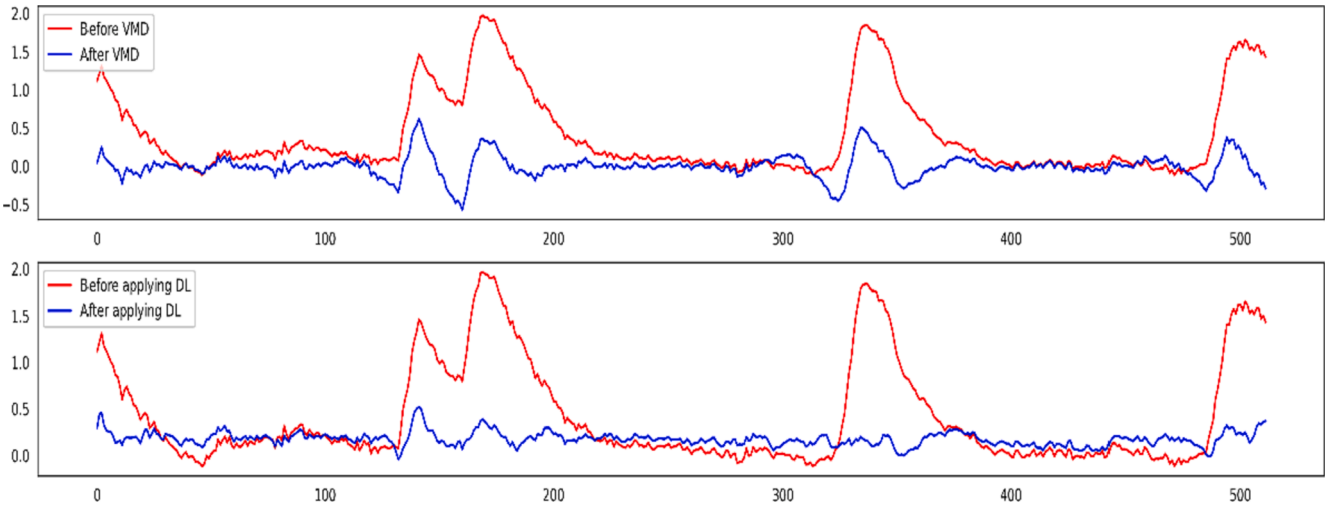

Fig. 7. Comparison of Eye blink removal with VMD and DL.

Table 7
Summary of previous studies based on emotion detection.

Author	Database	Modality	Classes	Method	Best Model	Overall Accuracy
Sarma et al.[30]	SEED	EEG	3(Positive, Negative, Neutral)	Instantaneous phase synchronization measurements and a majority voting algorithm	KNN	95.00
Weiss et al. (Altamirano Asher [4]	SEED, DEAP	EEG	2 (Arousal, Valance)	Measuring PSD features from all bands of 6 channels (FT7, FT8, T7, T8, TP7, TP8)	MLP, RFC, SVM	76.19
Ding et al.[9]	DEAP	EEG	2 (Arousal, Valance)	Dispersion Entropy computed for the four typical frequency bands	SVM	74.81
Yang et al.[40]	Self-collected	ECG, PPG	3(Positive, Negative, Neutral)	Time domain features were transformed into feature vector for feeding in CNN	CNN	75.4
Wang et al.[39]	Multi-modal Emotion Database	EEG, Speech	4(Neutral, Sad, Angry, Happy)	I-vector + PLDA and TCN was used for Speech signals, ELM and MLP was used in EEG after fusion in feature level or decision level	Fusion of AVER, WAVER, RF, ET	89.7
Tong et al.[38]	DEAP	EEG, PPG	2 (A,V)	Statistical features of PPG and total power, Ratio of delta to total, Ratio of alpha to total was extracted and fed into ML models	Adaboost	65.34
Ismail et al.[31]	DREAMER	EEG, ECG	2 (Arousal, Valance)	Peaks, HRV, Amplitude, Energy ratio, multiscale entropy, meanIBI and some more statistical features were extracted	SVM	67.84
Gao et al.[14]	ECSMP	EEG	2 (Positive, Negative) 6 (Fear, Sad, Angry, Disgust, Hap, Neutral)	MODFA-a1, FE, and PSD features were extracted only from gamma band	SVM	Binary: 96.81 Six-class: 42.17
Our method	**ECSMP**	**EEG, ECG, PPG**	**2 (Positive, Negative) 6(Fear, Sad, Angry, Disgust, Hap, Neutral)**	**TD, FD, TFD features from different modalities (EEG, ECG, PPG)**	**RFC**	**Binary: 96.12 Six-class: 94.25**

2023/R/1444). The statements made herein are solely the responsibility of the authors. The open access publication cost is covered by Qatar National Library.

CRediT authorship contribution statement

Purnata Saha: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing. **Ali K. Ansaruddin Kunju:** . **Molla E. Majid:** Funding acquisition, Methodology, Supervision,

Writing – original draft, Writing – review & editing. **Saad Bin Abul Kashem: . Mohammad Nashbat: . Azad Ashraf:** Conceptualization, Methodology, Software, Validation, Writing – original draft. **Mazhar Hasan:** Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing. **Amith Khandakar: . Md Shafayet Hossain: . Abdulrahman Alqahtani: . Muhammad E.H. Chowdhury:** Supervision, Validation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] H. Abdi, L.J. Williams, Principal component analysis: Principal component analysis, Wiley Interdiscip. Rev. Comput. Stat. 2 (2010) 433–459, https://doi.org/10.1002/wics.101.

[2] Abtahi, F., Ro, T., Li, W., Zhu, Z., 2018. Emotion Analysis Using Audio/Video, EMG and EEG: A Dataset and Comparison Study, in: 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). Presented at the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), IEEE, Lake Tahoe, NV, pp. 10–19. 10.1109/WACV.2018.00008.

[3] G. Alfian, M. Syafrudin, I. Fahrurrozi, N.L. Fitriyani, F.T.D. Atmaji, T. Widodo, N. Bahiyah, F. Benes, J. Rhee, Predicting Breast Cancer from Risk Factors Using SVM and Extra-Trees-Based Feature Selection Method, Computers 11 (2022) 136, https://doi.org/10.3390/computers11090136.

[4] K. Altamirano Asher Weiss, F. Concatto, R. Celeste Ghizoni Teive, A.R. Garcia Ramirez, On-line recognition of emotions via electroencephalography, IEEE Lat. Am. Trans. 20 (2022) 806–812, https://doi.org/10.1109/TLA.2022.9693565.

[5] C. Altın, O. Er, Comparison of Different Time and Frequency Domain Feature Extraction Methods on Elbow Gesture's EMG, Eur. J. Interdiscip. Stud. 2 (2016) 35, https://doi.org/10.26417/ejis.v2i3.p35-44.

[6] S. Bhattacharya, S. Borah, B.K. Mishra, A. Mondal, Emotion detection from multilingual audio using deep analysis, Multimed. Tools Appl. 81 (2022) 41309–41338, https://doi.org/10.1007/s11042-022-12411-3.

[7] F.Z. Canal, T.R. Müller, J.C. Matias, G.G. Scotton, A.R. de Sa Junior, E. Pozzebon, A.C. Sobieranski, A survey on facial emotion recognition techniques: A state-of-the-art literature review, Inf. Sci. 582 (2022) 593–617, https://doi.org/10.1016/j.ins.2021.10.005.

[8] G. Chanel, J.J.M. Kierkels, M. Soleymani, T. Pun, Short-term emotion assessment in a recall paradigm, Int. J. Hum.-Comput. Stud. 67 (2009) 607–627, https://doi.org/10.1016/j.ijhcs.2009.03.005.

[9] X.-W. Ding, Z.-T. Liu, D.-Y. Li, Y. He, M. Wu, Electroencephalogram Emotion Recognition Based on Dispersion Entropy Feature Extraction Using Random Oversampling Imbalanced Data Processing, IEEE Trans. Cogn. Dev. Syst. 14 (2022) 882–891, https://doi.org/10.1109/TCDS.2021.3074811.

[10] Y. Du, R.G. Crespo, O.S. Martínez, Human emotion recognition for enhanced performance evaluation in e-learning, Prog. Artif. Intell. 12 (2023) 199–211, https://doi.org/10.1007/s13748-022-00278-2.

[11] A. Fernandez, S. Garcia, F. Herrera, N.V. Chawla, SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary, J. Artif. Intell. Res. 61 (2018) 863–905, https://doi.org/10.1613/jair.1.11192.

[12] S. Gannouni, A. Aledaily, K. Belwafi, H. Aboalsamh, Electroencephalography based emotion detection using ensemble classification and asymmetric brain activity, J. Affect. Disord. 319 (2022) 416–427, https://doi.org/10.1016/j.jad.2022.09.054.

[13] S. Gannouni, A. Aledaily, K. Belwafi, H. Aboalsamh, Emotion detection using electroencephalography signals and a zero-time windowing-based epoch estimation and relevant electrode identification, Sci. Rep. 11 (2021) 7071, https://doi.org/10.1038/s41598-021-86345-5.

[14] Z. Gao, X. Cui, W. Wan, W. Zheng, Z. Gu, Long-range correlation analysis of high frequency prefrontal electroencephalogram oscillations for dynamic emotion recognition, Biomed. Signal Process. Control 72 (2022) 103291, https://doi.org/10.1016/j.bspc.2021.103291.

[15] Z. Gao, X. Cui, W. Wan, W. Zheng, Z. Gu, ECSMP: A dataset on emotion, cognition, sleep, and multi-model physiological signals, Data Brief 39 (2021) 107660, https://doi.org/10.1016/j.dib.2021.107660.

[16] J.J. Gross, R.W. Levenson, Emotional suppression: Physiology, self-report, and expressive behavior, J. Pers. Soc. Psychol. 64 (1993) 970–986, https://doi.org/10.1037/0022-3514.64.6.970.

[17] M.S. Hossain, S. Mahmud, A. Khandakar, N. Al-Emadi, F.A. Chowdhury, Z. B. Mahbub, M.B.I. Reaz, M.E.H. Chowdhury, MultiResUNet3+: A Full-Scale Connected Multi-Residual UNet Model to Denoise Electrooculogram and Electromyogram Artifacts from Corrupted Electroencephalogram Signals, Bioengineering 10 (2023) 579, https://doi.org/10.3390/bioengineering10050579.

[18] Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., Han, X., Chen, Y.-W., Wu, J., 2020. UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation.

[19] N. Ibtehaz, M.S. Rahman, MultiResUNet : Rethinking the U-Net architecture for multimodal biomedical image segmentation, Neural Netw. 121 (2020) 74–87, https://doi.org/10.1016/j.neunet.2019.08.025.

[20] M.F. Issa, Z. Juhasz, Improved EOG Artifact Removal Using Wavelet Enhanced Independent Component Analysis, Brain Sci. 9 (2019) 355, https://doi.org/10.3390/brainsci9120355.

[21] Jamal S, K.M., Kamioka, E., 2019. Emotions detection scheme using facial skin temperature and heart rate variability. MATEC Web Conf. 277, 02037. 10.1051/matecconf/201927702037.

[22] Kaiser, M.S., Bandyopadhyay, A., Ray, K., Singh, R., Nagar, V. (Eds.), 2022. Proceedings of trends in electronics and health informatics: TEHI 2021, Lecture notes in networks and systems. Presented at the International Conference on Trends in Electronics and Health Informatics, Springer, Singapore.

[23] S. Mahmud, M.S. Hossain, M.E.H. Chowdhury, M.B.I. Reaz, MLMRS-Net: Electroencephalography (EEG) motion artifacts removal using a multi-layer multi-resolution spatially pooled 1D signal reconstruction network, Neural Comput. Appl. 35 (2023) 8371–8388, https://doi.org/10.1007/s00521-022-08111-6.

[24] Mashhadi, N., Khuzani, A.Z., Heidari, M., Khaledyan, D., 2020. Deep learning denoising for EOG artifacts removal from EEG signals.

[25] R. Muthukrishnan, R. Rohini, LASSO: A feature selection technique in predictive modeling for machine learning, in: 2016 IEEE International Conference on Advances in Computer Applications (ICACA). Presented at the 2016 IEEE International Conference on Advances in Computer Applications (ICACA), IEEE, Coimbatore, India, 2016, pp. 18–20, https://doi.org/10.1109/ICACA.2016.7887916.

[26] M. Nazmul Islam Shuzan, M.E.H. Chowdhury, M. Bin Ibne Reaz, A. Khandakar, F. Fuad Abir, Md. Ahasan Atick Faisal, S. Hamid Md Ali, A.A.A. Bakar, M. Hossain Chowdhury, Z.B. Mahbub, M. Monir Uddin, M. Alhatou, Machine learning-based classification of healthy and impaired gaits using 3D-GRF signals, Biomed. Signal Process. Control 81 (2023) 104448, https://doi.org/10.1016/j.bspc.2022.104448.

[27] S. Pan, W. Li, Fuzzy-HMM modeling for emotion detection using electrocardiogram signals, Asian J. Control 22 (2020) 2206–2216, https://doi.org/10.1002/asjc.2375.

[28] F. Panahi, S. Rashidi, A. Sheikhani, Application of fractional Fourier transform in feature extraction from ELECTROCARDIOGRAM and GALVANIC SKIN RESPONSE for emotion recognition, Biomed. Signal Process. Control 69 (2021) 102863, https://doi.org/10.1016/j.bspc.2021.102863.

[29] M.Á. Sánchez-Cifo, F. Montero, M.T. Lopez, A methodology for emotional intelligence testing in elderly people with low-cost EEG and PPG devices, J. Ambient Intell. Humaniz. Comput. 14 (2023) 2351–2367, https://doi.org/10.1007/s12652-022-04490-9.

[30] P. Sarma, S. Barma, Emotion Recognition by Discriminating EEG Segments With High Affective Content From Automatically Selected Relevant Channels, IEEE Trans. Instrum. Meas. 71 (2022) 1–12, https://doi.org/10.1109/TIM.2022.3147876.

[31] S.N.M. Sayed Ismail, N.A. Ab. Aziz, S.Z. Ibrahim, A comparison of emotion recognition system using electrocardiogram (ECG) and photoplethysmogram (PPG), J. King Saud Univ. - Comput Inf. Sci. 34 (2022) 3539–3558, https://doi.org/10.1016/j.jksuci.2022.04.012.

[32] H. Shahid, A. Butt, S. Aziz, M.U. Khan, S.Z. Hassan Naqvi, Emotion Recognition System featuring a fusion of Electrocardiogram and Photoplethysmogram Features, in: 2020 14th International Conference on Open Source Systems and Technologies (ICOSST). Presented at the 2020 14th International Conference on Open Source Systems and Technologies (ICOSST), IEEE, Lahore, Pakistan, 2020, pp. 1–6, https://doi.org/10.1109/ICOSST51357.2020.9333021.

[33] N. Siew Cheok, P. Raveendran, Removal of EOG Artifacts Using ICA Regression Method, in: N.A. Abu Osman, F. Ibrahim, W.A.B. Wan Abas, H.S. Abdul Rahman, H.-.-N. Ting (Eds.), 4th Kuala Lumpur International Conference on Biomedical Engineering 2008, IFMBE Proceedings, Springer, Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 226–229, https://doi.org/10.1007/978-3-540-69139-6_59.

[34] Z. Song, T. Fang, S. Li, L. Niu, Y. Zhang, S. Le, G. Zhan, X. Zhang, H. Li, M. Zhao, H. Jiang, L. Zhang, X. Kang, Removing EOG Artifacts from the EEG signal of Methamphetamine Addicts, in: 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). Presented at the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE, Mexico, 2021, pp. 500–503, https://doi.org/10.1109/EMBC46164.2021.9629660.

[35] N. Spolaor, E.A. Cherman, M.C. Monard, H.D. Lee, ReliefF for Multi-label Feature Selection, in: 2013 Brazilian Conference on Intelligent Systems. Presented at the 2013 Brazilian Conference on Intelligent Systems (BRACIS), IEEE, Fortaleza, Brazil, 2013, pp. 6–11, https://doi.org/10.1109/BRACIS.2013.10.

[36] N. Stergiou, G. Giakas, J.E. Byrne, V. Pomeroy, Frequency domain characteristics of ground reaction forces during walking of young and elderly females, Clin. Biomech. 17 (2002) 615–617, https://doi.org/10.1016/S0268-0033(02)00072-4.

[37] D. Tkach, H. Huang, T.A. Kuiken, Study of stability of time-domain features for electromyographic pattern recognition, J. Neuroengineering Rehabil. 7 (2010) 21, https://doi.org/10.1186/1743-0003-7-21.

[38] Z. Tong, X. Chen, Z. He, K. Tong, Z. Fang, X. Wang, Emotion Recognition Based on Photoplethysmogram and Electroencephalogram, in: 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC). Presented at the 2018 IEEE 42nd Annual Computer Software and Applications Conference

(COMPSAC), IEEE, Tokyo, Japan, 2018, pp. 402–407, https://doi.org/10.1109/COMPSAC.2018.10266.

[39] Q. Wang, M. Wang, Y. Yang, X. Zhang, Multi-modal emotion recognition using EEG and speech signals, Comput. Biol. Med. 149 (2022) 105907, https://doi.org/10.1016/j.compbiomed.2022.105907.

[40] C.-J. Yang, N. Fahier, W.-C. Li, W.-C. Fang, A Convolution Neural Network Based Emotion Recognition System using Multimodal Physiological Signals, in: In: 2020 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-Taiwan). Presented at the 2020 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-Taiwan), IEEE, Taoyuan, Taiwan, 2020, pp. 1–2, https://doi.org/10.1109/ICCE-Taiwan49838.2020.9258341.

[41] H. Zhang, M. Zhao, C. Wei, D. Mantini, Z. Li, Q. Liu, EEGdenoiseNet: a benchmark dataset for deep learning solutions of EEG denoising, J. Neural Eng. 18 (2021) 056057, https://doi.org/10.1088/1741-2552/ac2bf8.

[42] Zheng, W.-L., Zhu, J.-Y., Lu, B.-L., 2016. Identifying Stable Patterns over Time for Emotion Recognition from EEG.