IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 30, No. 10, October, 2025, pp. 2739-2746
Journal homepage: http://iieta.org/journals/isi

EEG-Based Machine Learning for Emotional Stress Recognition in the Valence–Arousal Space

Jefferson Sanchez-Vivanco, Myriam Hernandez-Alvarez*

Departamento de Informática y Ciencias de la Computación, Facultad de Ingeniería de Sistemas, Escuela Politécnica Nacional, Quito 170524, Ecuador

Corresponding Author Email: myriam.hernandez@epn.edu.ec

https://doi.org/10.18280/isi.301019

Received: 15 September 2025
Revised: 15 October 2025
Accepted: 23 October 2025
Available online: 31 October 2025

Keywords:
emotional stress recognition, EEG signals, BCI, stress induction

ABSTRACT

Emotional stress impacts mental health and cognitive function, influencing human performance by affecting memory and attention. This study generated its own dataset of electroencephalographic (EEG) signals recorded using a Brain-Computer Interface (BCI) device. Participants were induced to stress using math tasks with strict time constraints. This dataset was used to identify significant features for the detection of emotional stress. The EEG signals were labeled according to their respective positions on the valence-arousal plane. Significant quadrant-specific thresholds relevant to stress were determined for classification and subsequent analysis. The class imbalance was mitigated using resampling methods. Feature extraction was performed using techniques in time, frequency, and time-frequency domains for obtaining a comprehensive signal representation. Principal Component Analysis (PCA) was applied to the extracted features to reduce dimensionality and improve model generalization. The features served as inputs to various CNN architectures to identify the optimum models for recognizing stress. The best recognition accuracy of 90.2% was obtained in the recognition of stress-related emotional states. The findings demonstrate the effectiveness of the combination of EEG signal processing and machine learning algorithms in the detection of stress levels in the valence-arousal emotional space.

1. INTRODUCTION

Emotional stress significantly influences mental and physical health, contributing to conditions like anxiety, depression, and other stress-related disorders. In academic performance, stress is a key factor that can affect cognitive function, memory, and attention [1, 2].

Accurate identification of emotional stress is critical in clinical and occupational settings, and educators and researchers could benefit from gaining real-time insights into the emotional state of students to design personalized interventions and develop adaptive learning strategies [3].

Electroencephalography (EEG) has emerged as an effective, non-invasive method to detect emotional states, leveraging signals captured by Brain-Computer Interface (BCI) devices. EEG signals can be systematically analyzed in the valence-arousal emotional plane, a widely accepted model in affective computing research [4].

The valence-arousal model categorizes emotions based on two dimensions: valence (positive or negative emotions) and arousal (intensity of emotional response). Identifying stress accurately within this space involves defining precise quadrant boundaries. Machine learning techniques, particularly convolutional neural networks (CNNs), have shown effectiveness in analyzing EEG signals, leveraging their capability to manage multi-dimensional and complex data.

This research is an approach to the problem of detecting stress as a result of facing math problems, leveraging the use of a BCI device to capture EEG signals. The research starts generating an original EEG dataset with the participation of a heterogeneous population that includes elderly people. The dataset goes through preprocessing to balance its classes to avoid bias toward the majority class. After that, a feature extraction process is defined in different domains, followed by a dimension reduction using PCA. The resulting dataset is tested with CNN configurations to determine the most effective models [5].

In this paper, Chapter 2 mentions some related work and emphasizes the gap that our research aims to fill, which is the acquisition of a relevant dataset with balanced classes and the identification of the combination of characteristics that works best to identify stress in the valence-arousal plane.

Chapter 3 briefly describes the stress induction protocol applied in the experiments, details labeling according to the location in the valence-arousal plane, the techniques to balance the dataset, the feature extraction algorithms in the different domains, the application of PCA to reduce feature dimensionality, and the process to get the best-performing CNN system.

Chapter 4 shows the best configuration used for the CNN algorithm and accuracy results obtained using features in the time, frequency, time-frequency, and a combination of

characteristics in all domains.

Chapter 5 presents conclusions and the direction of possible future research.

2. RELATED WORK

The convolutional neural networks (CNN) algorithms have shown useful results in stress detection using EEG signals as inputs. Martínez-Rodrigo et al. [6] used CNN algorithms to extract spectral features using an AlexNet architecture, obtaining 84% accuracy in differentiating between two classes: stress and non-stress conditions. They do not extract features before applying the CNN algorithm, but the input is raw data, so the processing times, although not given, must be longer than if feature data is the input [7-12].

Other works [13-15] presented CNN and Deep learning applications, achieving a 96% accuracy in the detection of stress in a process that, according to the authors, is not suitable for real-time use due to the necessity of pre-processing a large volume of raw data.

Jaloli et al. [16] explored stress detection using a CNN algorithm in a non-EEG dataset obtained using wearable sensors. The accuracy was 99.82%, suggesting better results than those achieved using EEG. Yet, this paper also uses raw data to process with CNN and extract features automatically, which again requires longer processing times and, therefore, cannot be used in real-time applications.

While these studies deliver valuable results, they focus only on features extracted via the CNN method using raw data and lengthy processing. This gap justifies our work, which aims to integrate feature extraction in the time, frequency, and time-frequency domains with CNN architectures. The use of pre-obtained features enables faster processing that could facilitate real-time applications useful for EEG-based stress recognition applied to multiple fields.

Furthermore, those studies did not consider the valence-arousal space for labeling emotional states, which limits the granularity of the classification. In contrast, our approach uses this model to facilitate the individual's self-report of their feelings.

3. MATERIALS AND METHODS

This work generates a dataset called EEGstress1-epn with EEG data collected from 11 participants using a BCI device that detects these signals in real-time in different brain areas and reports them as CSV files. The device is called Emotiv Epoc+, shown in Figure 1, and has 14 electrodes corresponding to the channels, according to the 10-20 system presented in Figure 2: AF3, F7, F3, FC, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4, plus two reference signals CMS and DRL located in the P3 and P4 channels.

Figure 1. Brain computer interface emotiv Epoc+

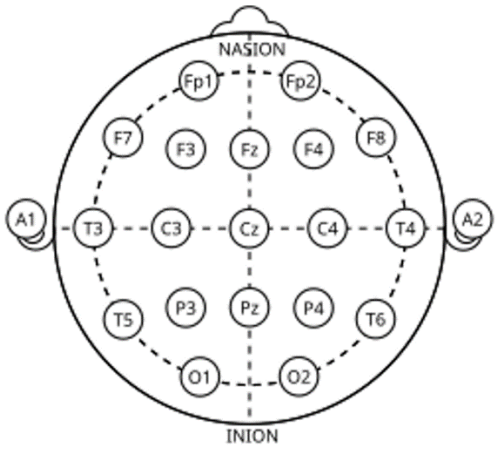

Figure 2. 10-20 system for electrode placement

3.1 Stress induction protocol

The present work obtains its own data set of EEG signals using a 14-channel BCI device. The dataset was called EEGstress1-epn. The participants were induced to stress by being challenged with arithmetic tasks with a strict time limit to respond. Immediate visual feedback was provided to indicate if the answers were correct or incorrect, aiming to increase anxiety and cognitive load.

The participants were 11, 7 male and 4 female, with ages between 26 and 79 years old. They were informed beforehand that their performance would be evaluated to increase psychological pressure and reinforce stress induction. The participants had to qualify their level of stress in the valence-arousal plane using the Self-Assessment Manikin (SAM) presented in Figure 3. Valence evaluated if the feeling was positive, neutral, or negative, and arousal estimated the intensity of the perceived emotion.

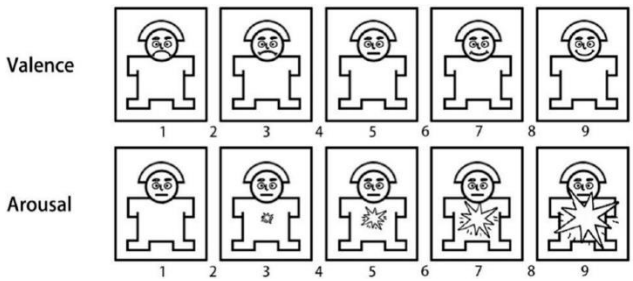

Figure 3. Self-assessment manikin

Also, a baseline was previously recorded where the participants were asked to relax. EEG signals recorded during relaxation and the execution of the math tasks generated the EEGstress1-epn dataset.

The classes in the valence-arousal plane were stress, calm, and other. The stress state includes high stress with a valence lower than three and arousal over 5, mild stress with a valence between 3 and 5, and arousal between 4 and 6. These are negative emotions with high activation linked to anxiety or cognitive overload. Calm is classified when the valence is between 4 and 6, and arousal is lower than 4, which is a positive feeling with low activation. Other values in the valence-arousal plane were considered in the "other" category.

This classification approach allows for a nuanced differentiation of emotional states while preventing misclassification of ambiguous conditions. The plane

valence–arousal is shown in Figure 4.

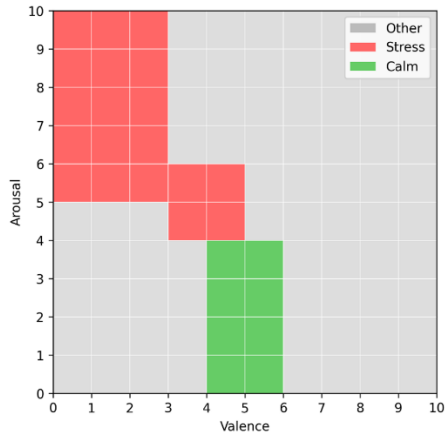

Figure 4. Valence-arousal for stress recognition

3.2 Strategies to overcome class imbalance

It is necessary to generate a balanced dataset with each class receiving enough samples to prevent the learning algorithm from always predicting the majority label. In applications like stress detection, it is crucial to avoid misclassifying, for instance, high-stress instances as calm or vice versa. One method to ensure a balanced dataset is to use resampling techniques to guarantee that every stress category contributes equally to the model, alleviating bias, improving recall on minority classes, and yielding more reliable performance.

In this work, three popular resampling techniques were used: RandomOverSampler, Borderline-SMOTE, and RandomUnderSampler. After applying them, we chose the method that gives the best performance in combination with a feature selection algorithm.

3.2.1 RandomOverSampler

This method duplicates randomly selected instances from minority classes until all classes have equal representation. By oversampling to the number of cases in the majority class, the classifier sees each class equally often during training.

3.2.2 Borderline-SMOTE

This method is an extension of the Synthetic Minority Over-sampling Technique (SMOTE). Borderline-SMOTE focuses synthetic sample generation on minority examples near the decision boundary. It identifies minority samples whose nearest neighbors include majority instances ("danger" samples) and creates new points along the line segments joining each "danger" sample to its minority neighbors. This targeted oversampling sharpens the classifier's ability to distinguish between classes in the critical boundary region, reducing the risk of creating noisy, out-of-distribution examples.

3.2.3 RandomUnderSampler

It is an algorithm that removes samples from the majority classes until class counts are balanced. While it avoids overfitting by reducing redundant data, it may discard potentially informative examples, and is thus more suitable when the majority class vastly outnumbers the minority and the lost information is unlikely to be critical.

Table 1 reports the number of EEG epochs of 60 seconds each, per class ("Stress", "Calm", "Others"). The table presents the epochs in the original dataset and after applying the three resampling techniques. Notice how oversampling methods balance all classes to the size of the majority class (1,010 samples), while under-sampling reduces each class to the size of the smallest class (129 samples).

Table 1. Class sample counts before and after resampling

Technique	Stress	Calm	Other	Total
Original	129	141	1,010	1,280
RandomOverSampler	1,010	1,010	1,010	3,030
Borderline-SMOTE	1,010	1,010	1,010	3,030
RandomUnderSampler	129	129	129	387

3.3 Feature extraction

Feature extraction was performed for each channel using a 1x7680 array, where various features in the time, frequency, and time-frequency domains were extracted [17-19]. The value of 7680 represents the number of data samples recorded from each channel during a 60-second interval. With a sampling rate of 128 Hz (i.e., 128 samples per second), it was obtained $128 \times 60 = 7680$ samples were obtained.

Feature extraction leveraged diverse techniques in time, frequency, and time-frequency domains. The characteristics extracted in the time domain are computed directly from the raw EEG signal over time and give its statistical and structural behaviour. They were:

Median: The central value of the signal.

Mean: The average signal amplitude in a window.

Variance: Evaluates the spread of the signal values around the mean.

Maximum: The highest amplitude in the window.

Minimum: The lowest amplitude.

Root mean square or RMS: Amplitude and duration in one measure.

Peak to peak value: The difference between the maximum and minimum values, which reflects the signal range.

Energy (Eng): For each 60-s window, the EEG signal x(n) is measured in microvolts (μV). Eng is computed as the sum of squared amplitudes over the N samples of the window, and therefore has units of μV^2.

Average power (Avg) is obtained by normalizing the energy by the number of samples in the window, which yields μV^2 per sample (equivalently proportional to $\mu V^2/s$ given the fixed sampling frequency of 128 Hz).

Line length: The Sum of absolute differences between consecutive samples to evaluate signal complexity.

Nonlinear energy: Evaluates changes in signal structure by combining products of successive samples.

First differences: Computes the average of absolute changes between consecutive points to evaluate signal variability.

Area Under the Curve (AUC): It represents the integral of the signal and is useful for comparing overall amplitude or energy between segments of the signal.

Kurtosis: Measures the extremity of values in the signal distribution. High kurtosis indicates more outliers.

Skewness: Evaluates the asymmetry of the amplitude distribution.

Hjorth Parameters: Represents the mean frequency or signal smoothness. These parameters include Hjorth complexity, which measures the change in frequency. High complexity suggests fast-changing patterns.

Petrosian Fractal dimension (PFD): Estimates the signal complexity by analyzing changes in signal direction.

Hurst exponent: It is a statistical measure that evaluates how the signal changes over time, showing persistence and randomness.

In the frequency domain, the features describe the distribution of energy obtained using the Fourier transform. In this domain, the characteristics obtained were:

Five bands: Delta in the 1 to 4 Hz range, Theta in the 4 to 8 Hz range, Alpha in the 8 to 13 Hz range, Beta in the 13 to 30 Hz range, and Gamma for frequencies greater than 30 Hz. Each band is related to a type of brain activity. Band powers are computed from the power spectral density (PSD) of each 60-s window. For each band B $\in$ {delta, theta, alpha, beta, gamma}, the band energy is defined as the integral of the PSD over the corresponding frequency range, which results in units of μV^2. To reduce inter-subject variability, we use relative band-energy ratios: each band feature reported in Table 2 corresponds to the band energy divided by the total energy in the 1–45 Hz range, yielding a dimensionless ratio.

Spectral Entropy: It computes the irregularity or disorder of the power spectral density (PSD) distribution. High entropy indicates a flat and noisy spectrum, while low entropy suggests a dominant frequency.

The frequency-time domain is obtained using wavelet transforms. Among these methods, we used:

Wavelet variance: Measures the variability of the wavelet coefficients to assess signal fluctuations over time.

Wavelet STD: The standard deviation of wavelet coefficients to evaluate dispersion in time and frequency.

Wavelet Absolute mean: Computes the average magnitude of wavelet coefficients, to indicate the level of signal activity.

Wavelet energy: Is the sum of the squares of wavelet coefficients divided by their length to measure localized energy.

Wavelet entropy: It is the normalized magnitude of the wavelet coefficient that interprets a probability distribution. It evaluates the unpredictability of the wavelet coefficients since higher values mean more disorder [20], providing comprehensive coverage of stress-indicative signals.

A summary of the feature extraction algorithms used in the present work is shown in Table 2.

Table 2. Feature extraction algorithms

Feature	Equation
Time Domain Features	
Maximum	$\mathrm{Max}(X_i), (X_i \epsilon C)$
Minimum	$\mathrm{Min}(X_i), (X_i \epsilon C)$
Median	$Median = sort(x)_{\frac{N+1}{2}}$
Mean	$Mean = \frac{1}{N}\sum_{i=1}^{N} X_i$
Variance (Var)	$Var = \frac{1}{N}\sum_{i=1}^{N}(X_i - mean)^2$
Root mean square (RMS)	$RMS = \sqrt{mean(x^2)}$
No linear energy (Nl)	$Nl = \sum_{i=1}^{N-1} x_i^2 - x_{i+1} * x_{i-1}$
Peak-to-Peak (p-p)	$p - p = \lvert max(x) - min(x)\rvert$
Energy (Eng)	$Eng = \sum_{n=-\infty}^{\infty} \lvert X(n)\rvert^2$
Average power (Avg)	$Avg = \lim_{N\to\infty} \frac{1}{2N}\sum_{n=-N}^{N} \lvert X(n)\rvert^2$
Line length (Ll)	$\mathrm{Ll} = \sum_{i=1}^{N-1} \lvert x_i - x_{i-1}\rvert$
No linear energy (Nl)	$Nl = \sum_{i=1}^{N-1} {X_i}^2 - X_{i+1} * X_{i-1}$
First differences (d1)	$d1 = \frac{1}{N-1}\sum_{i=1}^{N-1} \lvert x_{i+1} - x_i\rvert$
Second differences (d2)	$d2 = \frac{1}{N-2}\sum_{i=1}^{N-2} \lvert x_{i+2} - x_i$
Area under the curve (AUC)	$\Delta x = \frac{b-a}{n}$ $\int_a^b \mathrm{f(x)dx} \approx \frac{\Delta x}{2}[f(x_0 + 2f(x_1) \dots 2f(x_{n-1}) + f(x_1)$
Kurtosis (k)	$k = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^4}{\left(\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2\right)^2}$
Skewness (skew)	$skew = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^3}{(N-1)\sigma^3}$
Hjorth activity (σ_1)	$\boldsymbol{\sigma_1} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2$

	$\sigma_2 = \text{var}(x')$
Hjorth mobility	$\text{Mobility} = \sqrt{\frac{\sigma_1}{\sigma_2}}$
Hjorth complexity	$\text{Complexity} = \frac{\text{Mobility}(x')}{\text{Mobility}(x)}$
Petrosian Fractal Dimension (PFD)	$\text{PFD} = \frac{log_{10}N}{log_{10}N + log_{10}(\frac{N}{N} + 0.4N_\delta)}$
Hurst Exponent	$\text{Hurst} = \frac{R(T)}{S(T)}(23) = \frac{max(X(t,T)) - \min(X(t,T))}{\sqrt{(\frac{1}{t})\sum_{t=1}^{T}[x(t) - \bar{x}]^2}}$
Frequency Domain Features	
Power Spectral Bands	$Theta = power(x, f\epsilon[4hz - 8hz])$ Alpha=$power(x, f\epsilon[8hz - 12hz])$ $Low_{Beta} = power(x, f\epsilon\{12Hz - 16Hz]$ $High_{Beta} = power(x, f\epsilon[12hz - 25hz])$ $Gamma = power(x, f\epsilon[25hz - 45hz])$
Spectral Entropy (H)	$H = -\frac{1}{\log(k)}\sum_{i=1}^{k} RIR_i log RIR_i$
Frequency – Time Domain Features	
Wavelet_energy (w)	$w(x_j) = \sqrt{\frac{1}{N}\sum_{i=1}^{N_i}(x_i)^2[i]}$
Wavelet variance (w_{var})	$w_{var} = var(w(x_j))$
Wavelet standard deviation (w_{std})	$Wstd = \text{std}(w(x_j))$
Wavelet absolute mean (wabs_mean)	$Wabs\_mean = abs(mean\left(w(x_j)\right))$
Wavelet entropy (W_{ent})	$w_{ent} = -\sum_{j=1}^{N} d_j log(d_j)$ $whered_j = \frac{\lvert w(x_j)\rvert}{\sum_{j=1}^{N} w(x_j)}$

3.4 Feature selection with Principal Component Analysis

Principal Component Analysis (PCA) is an unsupervised dimensionality-reduction technique that linearly transforms the original feature space into a new set of orthogonal axes considered as the principal components, ordered by the amount of variance they explain. By projecting high-dimensional data onto the first principal components, PCA both reduces computational burden and mitigates overfitting by discarding directions in which the data have little variation. In the context of EEG feature vectors, where each channel may contribute dozens of time-, frequency-, and time-frequency–domain descriptors, PCA allows us to concentrate on the combinations of features that capture the greatest signal variability while suppressing noise and redundancy.

We applied PCA after concatenating all extracted features for each 60-second epoch, before feeding the data into our 1D-CNN. The number of retained components was chosen to explain at least 95% of the total variance; this method reduced feature dimensionality by more than 70%, greatly accelerating network training without sacrificing classification performance.

3.5 Classification algorithms

The classification method used was CNN, which was chosen due to its suitability for complex, multi-dimensional data. Multiple CNN configurations were tested with a 1D dimension for several layer compositions. The experiments were structured to determine which feature combinations and CNN configurations yield optimal accuracy in stress detection.

4. RESULTS

The dataset that gave the best results within a 5-fold cross-validation framework had the combination of PCA for feature selection and RandomOverSampler. This PCA + RandomOverSampler consistently produced the best overall accuracy when used in different CNN configurations.

Moreover, experimental results demonstrated promising performance across different CNN configurations. The best-performing CNN model configurations achieved an average accuracy of 90.2%, indicating significant potential for practical applications in real-world stress detection. This configuration is presented in Table 3.

An essential finding demonstrated in Table 4 was the benefit of combining features from multiple signal domains (time, frequency, and time-frequency), which delivers a more robust and discriminative input vector than single-domain features. This multi-domain feature integration approach outperformed traditional single-domain methods commonly reported in the literature. This way, the best results were obtained when features of all the domains were combined, producing a 90.20% accuracy using a 5-fold cross-validation. The architecture that yielded the best results is shown in Table 3. This Convolutional Neural Network (CNN) architecture follows a structured pattern of convolution, batch, normalization, pooling, and dense layers to process the input vector that includes features extracted in the time, frequency,

time-frequency, and combined domains of each labeled set of EEG data corresponding to a 60-second interval. The average processing time was in the millisecond range.

The configuration presents three 1D convolutional layers, with 32, 64, and 128 filters, respectively. Each convolutional layer is followed by batch normalization to stabilize learning and enhance convergence. Also, each convolutional block includes average pooling with a pool size of 2 to reduce the dimensionality while retaining key features. Finally, the tensors are flattened to transition into fully connected layers with 1024 neurons and rectified linear unit (ReLU) activation, 512 neurons with hyperbolic tangent (tanh) activation, 256 neurons with Tanh activation, and 128 neurons with ReLU activation. Each dense layer is followed by a dropout with a rate of 0.5 to prevent overfitting. The output dense layer has three neurons and a softmax activation with a probability distribution over three classes for stress, calm, and other.

PCs (95% var.) indicates the number of principal components retained after PCA so that at least 95% of the original feature variance is preserved. In Table 4, the Time Domain set required 10 components, the Frequency Domain 12, the Time-Frequency Domain 15, and the full-concatenated feature vector 20, to reach a 95% variance threshold. Figure 5 shows the pipeline for the multi-domain feature extraction and classification process.

Table 3. CNN classifier

Block	Layer (Keras)	Activation	Notes
Convolutional block 1	Conv1D(32, kernel_size = 2, strides = 1, padding = 'same')	tanh	Tensor shape (input→output): [batch, 20, 1] → [batch, 20, 32] → [batch, 10, 32]. Input: 20-dimensional feature vector from PCA. Followed by Batch Normalization (BN) and Average Pooling (AvgPool1D, pool_size=2).
Convolutional block 2	Conv1D(64, kernel_size = 2, strides = 1, padding = 'same')	tanh	Tensor shape (input→output): [batch, 10, 32] → [batch, 10, 64] → [batch, 5, 64]. Followed by BN and AvgPool1D(pool_size=2).
Convolutional block 3	Conv1D(128, kernel_size = 2, strides = 1, padding = 'same')	tanh	Tensor shape (input→output): [batch, 5, 64] → [batch, 5, 128] → [batch, 2, 128]. Followed by BN and AvgPool1D(pool_size=2).
Flatten	Flatten()	–	Tensor shape (input→output): [batch, 2, 128] → [batch, 256]. After the three convolution + average pooling blocks, the 1D feature map has a length 2 and 128 channels, so the Flatten layer produces a 256-dimensional vector that is fed into the fully connected block.
Fully connected block	Dense(1024) + Dropout(0.5)	ReLU	Tensor shape (input→output): [batch, 256] → [batch, 1024].
	Dense(512) + Dropout(0.5)	tanh	Tensor shape (input→output): [batch, 1024] → [batch, 512].
	Dense(256) + Dropout(0.5)	tanh	Tensor shape (input→output): [batch, 512] → [batch, 256].
	Dense(128) + Dropout(0.5)	ReLU	Tensor shape (input→output): [batch, 256] → [batch, 128].
Output layer	Dense(3)	softmax	Tensor shape (input→output): [batch, 128] → [batch, 3]. 3 classes.

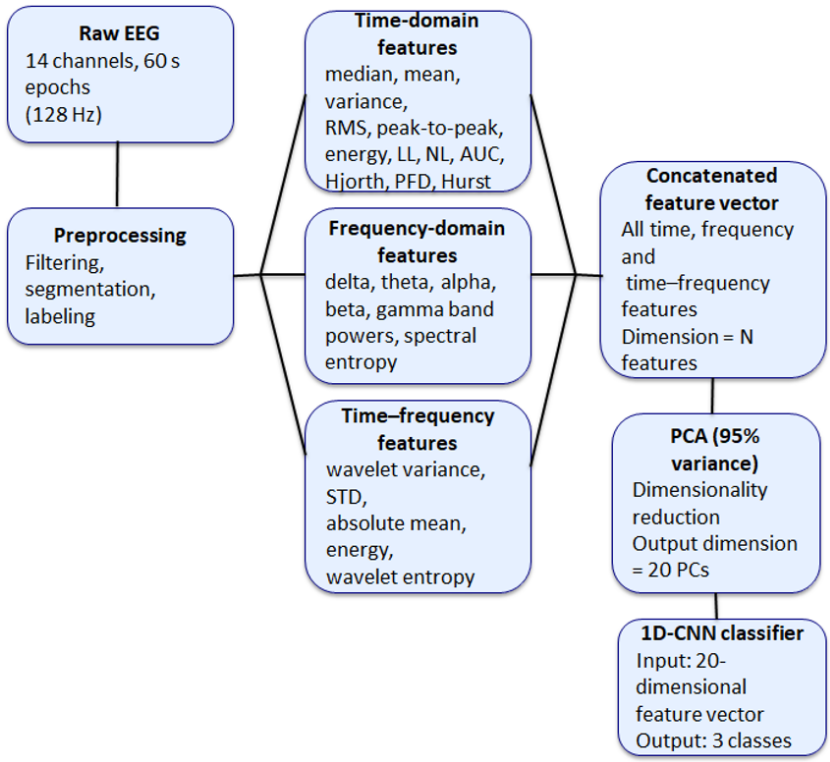

Figure 5. Pipeline for feature extraction and classification

Table 4. Number of principal components, feature domains, and accuracy

Feature Domain	#PCs (95% var.)	Mean Acc.	Std. Dev.
Time only	10	0.835	0.03
Frequency Only	12	0.864	0.04
Time-Frequency	15	0.873	0.02
Combined Domains	20	0.902	0.02

5. CONCLUSIONS AND FUTURE WORK

This study demonstrates the feasibility of emotional stress recognition using EEG signals analyzed through CNNs within the valence-arousal emotional framework. Key contributions include defining clear quadrant boundaries for stress identification, validating arithmetic problems under time pressure as effective stress-inducing stimuli, utilization of class balancing via several methods, dimensionality reduction of features through PCA, and exploring multi-domain feature extraction techniques.

Unlike previous works that, when using deep networks, rely solely on raw EEG data and end-to-end characteristics extracted within the CNN algorithm, our approach leverages pre-extracted features in different domains: time, frequency, time-frequency, and a combination of all features. This methodology, compared with an input vector with raw data, enables faster training, better interpretability, and the potential for real-time implementation due to faster processing. The results are compared, and the conclusion is that the best accuracy is obtained when features of the three domains are combined, outperforming those using single-domain inputs, indicating that stress manifests simultaneously across multiple EEG signal characteristics.

CNN architectures effectively handled complex EEG data with the optimal configurations.

The limitation of this work includes the modest sample size of 11 participants. However, an important strength is that the cohort is heterogeneous with male and female individuals from a broad age range, from 26 to 79 years old. This demographic enriches the EEGstress1-epn dataset and increases applicability in real-world populations, where stress may affect people of all ages and backgrounds differently.

New research directions include exploring advanced deep learning techniques, such as transformer-based neural networks, to handle EEG temporal dynamics better. Another possible expansion is the integration of other physiological measures (skin conductance, heart rate variability) to validate and improve stress classification accuracy further.

These findings open avenues for real-time applications in mental health monitoring, occupational safety, academic performance improvement, and enhancing human-computer interaction technologies.

ACKNOWLEDGMENT

This work was supported by Escuela Politécnica Nacional, Dirección de Investigación in Quito, Ecuador.

REFERENCES

[1] Caviola, S., Carey, E., Mammarella, I.C., Szucs, D. (2017). Stress, time pressure, strategy selection and math anxiety in mathematics: A review of the literature. Frontiers in Psychology, 8: 1488. https://doi.org/10.3389/fpsyg.2017.01488

[2] Zanetti, M., Mizumoto, T., Faes, L., Fornaser, A., De Cecco, M., Maule, L., Valente, M., Nollo, G. (2021). Multilevel assessment of mental stress via network physiology paradigm using consumer wearable devices. Journal of Ambient Intelligence and Humanized Computing, 12(4): 4409-4418. https://doi.org/10.1007/s12652-019-01571-0

[3] Li, Z. (2023). AI-assisted emotion recognition: Impacts on mental health education and learning motivation. International Journal of Emerging Technologies in Learning, 18(24): 34-48. https://doi.org/10.3991/ijet.v18i24.45645

[4] Hikmawati, E., Alamsyah, N. (2024). Supervised learning for emotional prediction and feature importance analysis using SHAP on social media user data. Ingénierie des Systèmes d'Information, 29(6): 2345-2356. https://doi.org/10.18280/isi.290622

[5] Khan, S.S., Sudan, J.S., Pathak, A., Pandit, R., Rane, P., Kumawat, A.K. (2024). A review of EEG artifact removal methods for brain-computer interface applications. Ingénierie des Systèmes d'Information, 29(1): 247-252. https://doi.org/10.18280/isi.290124

[6] Martínez-Rodrigo, A., García-Martínez, B., Huerta, Á., Alcaraz, R. (2021). Detection of negative stress through spectral features of electroencephalographic recordings and a convolutional neural network. Sensors, 21(9): 3050. https://doi.org/10.3390/s21093050

[7] Pathapati, S., Nalini, N.J. (2025). Hybrid optimized techniques for EEG based mild cognitive impairment detection using time domain feature extraction. Ingénierie des Systèmes d'Information, 30(1): 231-242. https://doi.org/10.18280/isi.300120

[8] Kumari, S., Khare, V., Arora, P. (2024). Optimizing seizure detection: A comparative study of SVM, CNN, and RNN-LSTM. International Journal of Computational Methods and Experimental Measurements, 12(4): 369-378. https://doi.org/10.18280/ijcmem.120405

[9] Vos, G., Ebrahimpour, M., van Eijk, L., Sarnyai, Z., Azghadi, M.R. (2025). Stress monitoring using low-cost electroencephalogram devices: A systematic literature review. International Journal of Medical Informatics, 198: 105859. https://doi.org/10.1016/j.ijmedinf.2025.105859

[10] Giannakakis, G., Grigoriadis, D., Giannakaki, K., Simantiraki, O., Roniotis, A., Tsiknakis, M. (2019). Review on psychological stress detection using biosignals. IEEE Transactions on Affective Computing, 13(1): 440-460. https://doi.org/10.1109/TAFFC.2019.2927337

[11] Arsalan, A., Majid, M. (2021). Human stress classification during public speaking using physiological signals. Computers in Biology and Medicine, 133: 104377. https://doi.org/10.1016/j.compbiomed.2021.104377

[12] Zhuang, J.B. Meng, P.C. (2023). Convolutional Neural Network-Assisted scattering inversion in diverse noise environments. Acadlore Transactions on AI and Machine Learning, 2(3): 170-181. https://doi.org/10.56578/ataiml020305

[13] Naqvi, S.F., Ali, S.S. A., Yahya, N., Yasin, M.A., et al.

(2020). Real-time stress assessment using sliding window based convolutional neural network. Sensors, 20(16), 4400. https://doi.org/10.3390/s20164400

[14] Pange, S.M., Pawa, V.R. (2025). Machine learning based framework for depression diagnosis using EEG and ECG Signals. Ingénierie des Systèmes d'Information, 30(5): 1251-1257. https://doi.org/10.18280/isi.300512

[15] Deshwandikar, K., Bhosale, N., Bhagwat, R., Chavan, P. (2025). Stress detection using electroencephalogram and deep learning. Mathematical Modelling of Engineering Problems, 12(5): 1812-1820. https://doi.org/10.18280/mmep.120535

[16] Jaloli, M., Choudhary, D., Cescon, M. (2020). Neurological status classification using Convolutional Neural Network. IFAC-PapersOnLine, 53(5): 409-414. https://doi.org/10.1016/j.ifacol.2021.04.193

[17] Álvarez-Jiménez, M., Calle-Jimenez, T., Hernández-Álvarez, M. (2024). A comprehensive evaluation of features and simple machine learning algorithms for electroencephalographic-based emotion recognition. Applied Sciences, 14(6): 2228. https://doi.org/10.3390/app14062228

[18] Bastos-Filho, T.F., Ferreira, A., Atencio, A.C., Arjunan, S., Kumar, D. (2012). Evaluation of feature extraction techniques in emotional state recognition. In 2012 4th International Conference on Intelligent Human Computer Interaction (IHCI), Kharagpur, India, pp. 1-6. https://doi.org/10.1109/IHCI.2012.6481860

[19] Hag, A., Handayani, D., Altalhi, M., Pillai, T., Mantoro, T., Kit, M.H., Al-Shargie, F. (2021). Enhancing EEG-based mental stress state recognition using an improved hybrid feature selection algorithm. Sensors, 21(24): 8370. https://doi.org/10.3390/s21248370

[20] Gao, Y., Wang, X., Potter, T., Zhang, J., Zhang, Y. (2020). Single-trial EEG emotion recognition using Granger Causality/Transfer Entropy analysis. Journal of Neuroscience Methods, 346: 108904. https://doi.org/10.1016/j.jneumeth.2020.108904