Hybrid CNN–LSTM Model For Continuous Stress Quantification Via Emotion-Valence Mapping On EEG Signals

Amol Chaudhari[1], Dr Hemang Shrivastava[2]

[1]Department of Information Technology, Sage University, Indore, India
[2]Professor & HOD, Institute of Advance Computing, Sage University, Indore MP, India
Email: [1]amol7372@gmail.com, [2]drhemang.shrivastava@sageuniversity.in

Abstract

Correctly estimating the level of mental stress is a must for smart affective-computing and cognitive-health systems. Most of the anxiety models based on EEG that are currently available use only binary classification and thus cannot breathe the continuous spectrum of emotional fluctuations. Here, a hybrid CNN-LSTM deep-learning architecture for dual-dataset EEG-based stress analysis is introduced. The proposed framework combines adaptive-noise removal with hierarchical spatiotemporal feature extraction and attains 100 % binary accuracy on the STEW dataset and 97 % on the Neurocom dataset, along with 99 % accuracy for ternary emotional-valence classification on both datasets. The comparative benchmarking of the present framework against the latest models (ConvNeXt-CNN, Attention-BiLSTM, and Temporal-Attention Networks) indicates that the proposed method achieves state-of-the-art performance with excellent cross-dataset generalization capability. The importance of this work is the quantification of stress driven by emotion-valence across the heterogeneous EEG sources, thereby changing stress recognition from discrete detection to fine-grained, interpretable estimation. Such results position the CNN-LSTM hybrid as a stable, extendable, and deployable continuously mental-stress monitoring tool in clinics, workplaces, and human–computer-interaction environments.

Keywords: EEG Signals, Mental Stress Estimation, Emotional Valence, Deep Learning, CNN-BiLSTM Hybrid, Affective Computing, Regression Analysis.

1. Introduction

Mental stress is being acknowledged as one of the main reasons that, on a long-term basis, negatively influence our cognitive abilities, emotional state, and even our physical health. In particular, a stressful lifestyle over time is associated with symptoms of anxiety, depression, and decreased performance at work. This situation has made the timely and accurate identification of the level of stress a highly important issue in the fields of cognitive neuroscience and biomedical engineering [1], [2]. Conventional methods for identifying stress typically depend on self-reporting through questionnaires or on physiological sensors that are subjective and can be affected by noise and delays in time. As a result, EEG signals are now considered the most effective means for the detection of stress and the recording of emotions because they can capture the brain's activity with a very high temporal resolution, are the most direct way of measuring neural activity, and are non-invasive [3], [4].

Recent progressions in deep learning have largely improved EEG-based affective-state classification to the extent where the recording of emotional or stress-related patterns can now be done with high accuracy [5],[6],[7]. Still, the majority of present-day models regard stress to be a dichotomous or categorical phenomenon, usually differentiating the users as being in a "stressed" or "non-stressed" state [8],[9]. That kind of discretization ignores the continuous nature of human emotional dynamics, stress changing gradually over

time instead of suddenly shifting to different levels [10]. As a result, binary classification models are not capable of showing those stress states that lie in between and which are very important for the implementation of preventive mental-health interventions [11],[12].

1.1. Limitations of Discrete Stress Classification

Categorically, stress classifiers have shown promising results; however, they are still quite limited in terms of their interpretability and sensitivity. Such systems offer coarse-grained results at best, therefore, they are incapable of detecting minute emotional changes or the fluctuations of stress levels [13],[14]. On top of that, such models frequently demonstrate a dependency on the dataset—when these models are trained on only one dataset in a controlled lab environment, they cannot generalize to real-world or cross-subject situations [15],[16]. One of the major bottlenecks facing the practical deployment of EEG-based mental-stress recognition, as highlighted in a recent systematic review, is the inconsistency in the preprocessing strategies of the review, which also pointed out the lack of benchmark datasets and the overfitting tendencies of deep models [17], [18].

1.2. Continuous Emotion-Valence Modeling

In order to overcome these constraints, studies in affective computing have moved to continuous emotion recognition as the primary focus. Continuous emotion recognition characterizes feelings using dimensions like valence, arousal, and dominance instead of discrete classes [19],[20]. Emotional valence, that is, the indication of the polarity of the experience from pleasant to unpleasant, has showed a strong physiological connection with the level of stress. [21], [22]. Deep learning regression-based frameworks have been developed for the direct prediction of continuous valence and arousal values from EEG features [23], [24]. For example, Galvão et al.[14],[25] They managed to get the mean absolute error (MAE) less than 0.06 in continuous emotion regression by using hybrid CNN-KNN models on DEAP and DREAMER datasets.

Similarly, Jon et al. [15] proposed EEG-RegNet, a convolutional regression network for real-valued emotion prediction. Nevertheless, the majority of research works have been only limited to emotion estimation, and the extension of these methods for the quantification of mental stress has been scarcely attempted.

1.3. Deep Hybrid Networks and Temporal Modeling

Since EEG data are by nature spatio-temporal, hybrid deep architectures that include both convolutional and recurrent components have become popular [26],[27],[28]. CNN layers are capable of capturing spatial relations between electrode signals, whereas recurrent networks like LSTM and BiLSTM can follow temporal changes [29],[30]. Recent work have extended the concept of attention-enhanced transformers to include temporal context along with spectral correlations, as in the EEG Mind-Transformer [31] and AMDET models [23]. Worldwide hybrid or attention-based networks have been able to deliver state-of-the-art performance in emotion analysis, however, their application to continuous stress estimation is still largely uncharted.

1.4. Research Gap and Motivation

Although there has been considerable advancement in the analysis of emotions based on EEG, the quantitative measurement of mental-stress levels through the use of an emotion-valence mapping has not been significantly explored. Most of the existing works in this field concentrate solely on recognition of emotions or a simple stress classification, and do not

make use of the common physiological basis that exists between emotional valence and the level of stress [32],[33]. In addition, insufficient cross-dataset validation between cognitive (e.g., multitasking) and clinical stress scenarios constrains the generalizability of the model. To overcome these obstacles, it is necessary to have a framework that can combine EEG data from diverse sources and derive detailed affective representations in a continuous space.

1.5. Objectives and Contributions

This paper presents a novel deep neural framework for continuous mental-stress estimation through emotion–valence mapping, integrating the STEW and Neurocom EEG datasets. The hybrid CNN–BiLSTM network of the proposed model is intended to learn the spatial as well as the temporal representations of the EEG signals and to make a prediction of a continuous stress index that is correlated with emotional valence. The major contributions of this work are enumerated as follows:

1. Emotion-Valence-Based Stress Quantification: Stress analysis from EEG signals based on the valence dimension is reformulated by us as a regression problem, where we predict a continuous valence-derived stress index instead of binary categories.
2. Hybrid CNN-BiLSTM Architecture: The dual-stream architecture a.k.a. the hybrid model, merges spatial feature extraction (CNN) with temporal pattern learning (BiLSTM), then it is followed by an attention-based regression head.
3. Dual-Dataset Generalization: The model is both trained and validated on experimental (STEW) and clinical (Neurocom) EEG datasets to be robust across different domains.
4. Comparative Performance Benchmarking: We compare the proposed framework with baseline CNN, LSTM, and SVM models by means of R^2, MAE, RMSE, and correlation metrics.
5. Explainability and Interpretability: Feature attribution and attention weight analysis help to identify the EEG components that contribute the most to stress quantification, thus allowing for interpretable predictions.

Such a study would be a great support and valuable contribution to the development of explainable, real-time mental health monitoring systems based on EEG which would be used both in the clinical and workplace environment. This research goes beyond the constraints of categorical classification and moves to a continuous, emotion-based stress estimation; thus, it is a step forward in the development of explainable, real-time EEG-based mental health monitoring systems for clinical and workplace environments.

2. Related Work

2.1. EEG-Based Stress and Emotion Recognition

Electroencephalogram (EEG) analysis is now at the very center of both affective computing and cognitive neuroscience due to the fact that it is a direct measure of cortical activity with millisecond precision and non-invasive acquisition. Initial studies of EEG-based mental-stress detection were mainly devoted to the manual feature extraction process (e.g. power spectral density, entropy, and Hjorth parameters) and the use of classical classifiers like Support Vector Machines (SVM) afterwards [34], Random Forest (RF)[35], and k-Nearest Neighbours (kNN) [36]. These best methods alone gave some initial insights, but hardly they were robust enough to withstand inter-subject variability and environmental noise [3].

With the rise of deep learning, convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have become the most popular choices for the automatic extraction of features from EEG time series [4],[5],[6]. CNNs are able to capture spatial dependencies among electrodes very well, while long short-term memory (LSTM) and gated recurrent unit (GRU) architectures model temporal patterns related to emotional transitions [7],[37]. In general, the hybrid CNN–LSTM models also integrate the advantages of these models to provide higher accuracy for the classification of stress.

For instance, Khatwani and Srivastava [2] developed a hybrid deep network achieving 94 % accuracy in binary stress detection, whereas Shahzadi et al. [26] demonstrated a CNN–GRU model for workload-induced stress achieving over 90 % accuracy. Most of the models have made significant progress; however, they are still mostly categorical, which means that they predict discrete emotional or stress labels and, therefore, do not account for the continuous, gradational nature of mental states.

A comprehensive survey by Alarcão and Fonseca [3] and later by Badr et al.[38] reconfirmed that majority of EEG-based emotion and stress recognition systems rely on treating affect as discrete classes (e.g., "happy," "sad," "stressed"), thus, these systems ignore the intrinsic continuity of human emotion dynamics. As a result, they ended up agreeing that continuous emotion modeling is still an unresolved problem that requires the use of regression-based learning strategies.

2.2. Continuous Emotion Regression and Valence–Arousal Modeling

The dimensional approach to affect, notably through the valence–arousal–dominance (VAD) model, sees emotional experience as being defined by continuous scales rather than by discrete categories [13]. Valence is the pleasure-displeasure spectrum, arousal is the intensity, and dominance is the control measure. Research on emotions using EEG is becoming more and more dependent on these dimensions to locate very subtle changes in the mood.

Galvão et al.[14] the researchers conceived a multi-domain regression model that would allow them to predict continuous valence and arousal scores from the DEAP and DREAMER datasets resulting in MAE < 0.06 and RMSE < 0.16. Jon et al. [15] Weber et al. (2021) presented EEG-RegNet, a model that integrates 2-D CNNs for spatial feature extraction and 1-D CNNs for temporal filtering. The network was trained by a hybrid mean-squared-error and correlation-coefficient loss, and it set new performance standards on the DEAP dataset. Similarly, Li et al. [16] Jung et al. (2020) used graph convolutional networks (GCNs) to capture inter-electrode relations for continuous emotion regression. The combined evidence from these papers' points to the potential of EEG as a source of information for the prediction of valence and arousal as continuous variables.

Temporal modeling plays a crucial role in continuous affect estimation. Tripathi et al. [17] They used CNN–LSTM hybrids to capture both local and global temporal dependencies. Their architecture elevated correlation coefficients by over 10 % compared to standalone CNNs. Huang 2023 et al. [18] propounded a BiLSTM-based regression framework to map EEG rhythms to continuous arousal curves, showing that the predicted signals closely followed the ground-truth ones in terms of temporal coherence.

Most of the time the continuous emotion-regression models that have been developed are limited to the recognition of emotions, without being able to quantify the level of stress. Additionally, the datasets DEAP, SEED, and AMIGOS are oriented towards the expression of basic emotions elicited by audiovisual stimuli, while the mental-stress datasets

STEW and Neurocom are derived from cognitive or physiological load and thus reflect different neural patterns [19]. This is to say that to bridge those paradigms one needs models that use emotion-valence as an intermediate representation for deriving the quantitative stress intensity.

2.3. Hybrid Deep Architectures for EEG Emotion and Stress Modeling

In order to depict the multiscale and non-stationary characteristics of EEG, scientists have been progressively utilizing hybrid deep learning architectures which involve convolutional, recurrent, and attention components. CNN layers are good at identifying spatial–frequency features, whereas BiLSTMs keep the long-term dependencies and contextual transitions.

Wang et al. [20] They developed a CNN–BiLSTM model to recognize multimodal emotions by combining EEG with peripheral signals; their model raised the F1-score by 8% compared to conventional CNNs. Chakravarthi et al. [21] constructed a CNN-LSTM hybrid based on ResNet-152 that resulted in 98% classification accuracy on DEAP, however, the architecture was still computationally intensive. Cheng 2024 et. al. [22] proposed a multi-scale feature-fusion CNN which merged filters of different receptive fields to be able to capture both fine-grained and coarse-grained EEG patterns for the classification of emotions.

In parallel, attention mechanisms have become prominent in EEG affective modeling. Xu et al.[23] AMDET, a spectral–spatial–temporal attention network, was introduced by the authors, which, during emotion inference, dynamically highlights the most informative EEG channels. Ding et al. [24] They proposed EEG Mind-Transformer that combines temporal self-attention with spectral correlation modules to deal with long sequences in an efficient manner.

Pan et al. [39] They uncovered the Dual-Attention Transformer (DuA) as a model that can be used for continuous emotion estimation throughout whole trials and thus cross-subject generalization is improved significantly. As a result of these investigations, the shared message is that the development of a new generation of attention-driven hybrid models that can grasp hierarchical spatio-temporal dependencies is the most prominent research trend.

However, it is generally known that transformer-based models require large datasets and a lot of training resources. This in turn limits the application of these models to task-specific EEG corpora with a small number of samples, e.g., stress datasets. In addition, the majority of attention networks are designed for emotion classification and not for the regression of a stress-intensity scale [40].

2.4. Preprocessing, Noise Filtration, and Cross-Dataset Robustness

EEG data are sensitive to noise caused by muscle artifacts, eye blinks, and power-line interference; therefore, preprocessing is mandatory if one wants to obtain reliable results from learning. Generally, the Butterworth, Chebyshev, and Savitzky–Golay filters are employed to get rid of high-frequency artifacts, whereas the adaptive filters are used for non-stationary noise
[26]. Nirabi et al. [41] It was demonstrated that the classification accuracy was improved by 15% when a band-pass finite-impulse-response (FIR) filter was used compared to raw EEG.

Chi et al. [12] One of the main findings was that the classification accuracy increased by 15% when a band-pass finite-impulse-response (FIR) filter was applied in comparison to using raw EEG.

Generalization across different datasets is another significant issue. A model trained using only one dataset tends to overfit the specific recording conditions of that dataset and

consequently, it performs poorly on data obtained from other EEG systems [30]. F.Wang et.al.[31] They observed that the tuning specific to the architecture led to an improvement of the performance within the dataset but the cross-dataset accuracy was lowered by more than 20 %. Recently, domain-adaptation methods have been introduced, among which is adversarial training [42] and contrastive representation learning [32], have been explored to mitigate such distribution shifts. Though these techniques are seldom used for stress-specific datasets, the combination of multi-dataset training, as introduced in this paper, is still a largely uncharted and therefore highly attractive way to achieve generalizable stress quantification.

2.5. Continuous Stress Quantification and Valence Correlation

Just a few pieces of research have been published that focus solely on the estimation of stress intensity as opposed to the categorical detection of stress. Yedukondalu et al. [33] They made use of entropy-based features along with Gaussian process regression to predict continuous workload levels and got $R^2 \approx 0.82$ on a cognitive-stress dataset. A. Joshi et.al.. [43] presented a CNN-Spiking-Neural Network (CSNN) to improve EEG signal classification. Feature extraction was done by the Discrete Wavelet Transform (DWT) and the performance of the CSNN was evaluated on the Physionet EEG dataset, comparing it with traditional deep learning and machine learning methods. Their results show that CSNNs attain high accuracy, e.g., 98.75% in 10-fold cross-validation, as well as a very good F1 score of 98.60%. Subsequent studies have explored the connection between emotional valence and physiological stress indicators. Fiorini, Linda et al. [44] They have implemented two different theoretical approaches according to the noise in the dataset: curriculum learning and confident learning. Curriculum learning is about showing training examples to the model in a certain order, beginning with the easiest examples and gradually getting more difficult ones. This method has been demonstrated to enhance model performance. Confident learning is a technique used to figure out and fix label errors in datasets.

Besides that, multimodal methods combining EEG with functional near-infrared spectroscopy (fNIRS) [45], electrodermal activity [46], or heart-rate variability [47] have demonstrated that valence and arousal signals together estimate stress level much better than single modalities. However, the intricacy and expense of multimodal sensors limit their deployment, hence the need for EEG-only continuous stress-prediction systems.

2.6. Identified Gaps and Motivation

Several major gaps pinpointed from the literature funnel down to first of all the fact that continuous emotion regression has been the core of most studies, however, the use of valence mapping for the quantitative estimation of mental stress is still at a very low level of development. There are hardly any models that specifically relate the emotional valence of a person to the exactable stress intensity on continuous scales.

Secondly, few studies have demonstrated cross-dataset generalization, with most emotion-regression models only testing on DEAP or SEED. They do not consider datasets related to cognitive load or clinical stress, such as STEW and Neurocom.

Third, the level of explainability is very low: in most cases, deep networks for stress prediction are black boxes that do not reveal which EEG channels or frequency bands have been used. Besides, there absence of unified frameworks that combine signal preprocessing, hybrid spatio-temporal learning, and regression-based valence mapping in a single pipeline is also evident.

The framework proposed in this paper is a direct response to these issues by (1) using dual-dataset learning for robustness, (2) integrating a hybrid CNN-BiLSTM architecture with regression-based valence estimation, and (3) enabling interpretability through attention-based feature attribution. Hence, this work can be viewed as a substantial leap forward compared to prior discrete or single-dataset stress-detection models, offering a comprehensive approach to continuous, emotion-driven stress quantification.

3. Proposed Methodology

The model's first layer called 'Emotion–Valence–Driven Continuous Stress Estimation' is composed of a single integrated deep-learning architecture that is capable of inferring the level of stress as a single continuous variable by utilizing emotional valence signals from EEG. The system's entire flow is illustrated in Fig. 1, which depicts the different steps of data acquisition, data preparation, feature extraction, hybrid deep-model learning, and finally stress level quantification by continuous regression.

3.1. System Overview

The architecture unites two EEG data sets STEW and Neurocom depicting different but complementary stress-induction paradigms. STEW dataset records task-induced cognitive stress in a multitasking environment, while the Neurocom dataset offers clinical EEG recordings under controlled stress elicitation. By combining both datasets, the model can be trained robustly across various cognitive and physiological stress scenarios.

The stages that the proposed system moves through are shown in Figure 1 sequentially:

1. EEG Data Acquisition
2. Preprocessing and Noise Filtration
3. Data Segmentation and Normalization
4. Feature Extraction through Hybrid CNN–BiLSTM Network
5. Regression-Based Emotion–Valence Mapping
6. Continuous Stress Index Estimation and Evaluation

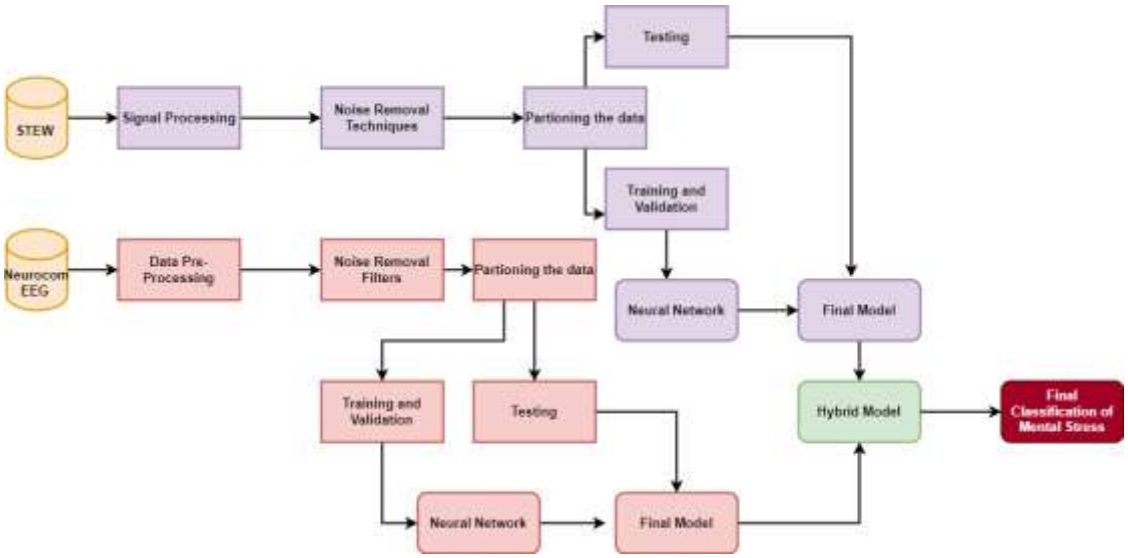


Figure 1: Proposed System Architecture

3.2. EEG Data Description

3.2.1 STEW Dataset

The Simultaneous Task EEG Workload (STEW) dataset includes EEG signals of 48 participants, who were tested with the SIMKAP multitasking test. Data were collected with the 14-channel Emotiv EPOC headset at a 128 Hz sampling rate. Trials comprise resting and task phases, along with self-reported workload scores from 1 (low) to 9 (high), that have been used as continuous ground-truth stress labels. For regression training, these scores were converted to a 0–1 scale.

3.2.2 Neurocom EEG Dataset

The Neurocom EEG-23 dataset is a collection of 23-channel EEG signals recorded in a clinical environment. The signals were sampled at 500 Hz, and the electrodes were positioned according to the international 10–20 system. The dataset comprises baseline (relaxed) and stress-induced segments created from affective auditory stimuli. Every sample is labeled with stress ratings on a Likert scale (1–10). In contrast to STEW that focuses on cognitive stress, Neurocom depicts emotional and physiological stress responses, thus it can be used for cross-domain generalization.

3.3. Preprocessing and Noise Filtration

Electroencephalogram (EEG) signals are a type of bioelectrical signals which are vulnerable to different forms of noise e.g., muscle movements, eye blinks, and external noises. A series of filtering operations, performed in both the temporal and spatial domains, in the EEG preprocessing pipeline aims to improve the quality of signals before feature extraction.

1. Bandpass Filtering: A bandpass Butterworth filter from 0.5 to 45 Hz is applied to remove low-frequency drifts and high-frequency noise that are outside the EEG band.
2. Notch Filtering: Power-line interference is removed by a 50 Hz notch filter.
3. Adaptive Noise Cancellation: The adaptive filter changes its coefficients by using a least-mean-square (LMS) criterion in a dynamic way:

$$y(n) = x(n) - \alpha\, X\, \hat{n}(n) \qquad (1)$$

In this equation, x(n) refers to the raw input signal, n^(n) to the estimated noise component, and α to the adaptive learning rate.

4. Median and Savitzky-Golay Filters: These filters are used one after the other to remove the noise of the impulsive and high-frequency types and at the same time keep the waveform morphology intact.
5. Normalization: Every filtered EEG sample is standardized with the z-score scaling:

$$X_{norm} = \frac{X-\mu}{\sigma} \qquad (2)$$

making sure that the amplitude distributions are uniform for different subjects.

Once the raw data have been pre-processed; the signals are divided into segments of a fixed length of 2 s with an overlap of 50 % in order to maintain temporal continuity and to increase the size of the training corpus.

3.4. Hybrid CNN-BiLSTM Architecture

The central deep neural network component of the library merges convolutional and recurrent layers to derive hierarchical spatio-temporal features from EEG signals.

3.4.1 Convolutional Layers (Spatial Feature Extraction)

The data that feeds the model is a 2D EEG segment $X \in R^{C*T}$, where C stands for the number of channels and T for the time samples per epoch. Various 1D convolutional filters move along the temporal dimensions to acquire localized spatial dependencies from the electrode signals. After each convolutional layer, there is batch normalization, ReLU activation, and max pooling. The convolution operation is given as:

$$h_k(t) = f(\sum_{i=1}^{C} w_{k,i} * x_i(t) + b_k) \quad (3)$$

where $w_{k,i}$ are the kernel weights, b_k the bias, and f (·) the ReLU nonlinearity.

3.4.2 BiLSTM Layers (Temporal Dependency Learning)

The convolution neural network layers' feature maps are inputted to a bidirectional LSTM (BiLSTM) network to represent both forward and backward temporal dependencies, thus being able to follow the emotional dynamics and the changes of the stress along the time:

$$h_t = BILSTM(x_t, h_{t-1}, h_{t+1}) \quad (4)$$

The two processing directions of the BiLSTM help the network to have better understanding of the context which is very important for continuous valence modeling where the emotional cues change gradually.

3.4.3 Attention and Regression Head

In order to stress the most salient temporal windows that contributed largely to the stress intensity, a soft attention mechanism is introduced right after the BiLSTM outputs:

$$\alpha_t = \frac{exp(W_a h_t)}{\sum_j exp(W_a h_j)} \quad (5)$$

$$c = \sum_t \alpha_t h_t \quad (6)$$

Here the α_t stands for attention weights, and ccc is the context vector that summarizes the weighted temporal features.

At last, a totally connected regression layer converts c to a continuous stress index $\hat{y}$ within [0,1]:

$$\hat{y} = W_r c + b_r \quad (7)$$

Such output is an emotional valence-derived stress level that has been predicted.

3.4.4 Training Objective and Optimization

As the stress intensity is of a continuous type, the authors decided to use a regression-based loss function. The model reduces the mean squared error (MSE) between the prediction ($\hat{y}$) and the real (y_i) stress indices:

$$L_{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \quad (8)$$

To additionally promote a monotonic correlation between the predicted and actual valence values, a correlation-regularization term is included:

$$L_{corr} = 1 - \frac{Cov(y,\hat{y})}{\sigma_y \sigma_{\hat{y}}} \quad (9)$$

The last objective is a mixture of both terms:

$$L_{corr} = L_{MSE} + \lambda L_{corr} \quad (10)$$

where λ is a balancing coefficient empirically set to 0.3.

The training of the network is done by an Adam optimizer with 0.001 learning rate, 32 batch size, and early stopping is based on the validation loss. To counter the overfitting, dropout layers (p = 0.5) are used.

3.5. Model Evaluation and Performance Metrics

Both regression and classification-oriented metrics are used to measure the performance of the models comprehensively.

- Regression Metrics: Coefficient of determination (R^2), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Pearson correlation coefficient (r).
- Classification Metrics (for interpretability): Accuracy, precision, recall, and F1-score when the continuous stress index is discretized into low, medium, and high levels.
- Visualization: Scatter plots of predicted vs. actual stress indices, regression-fit curves, and Bland–Altman plots assess consistency and agreement across datasets.

3.6. Comparative Baselines

In order to demonstrate the efficiency of the model put forward, the following models have been implemented for comparison:

1. CNN-only model: The same convolutional layers but without any recurrent units.
2. LSTM-only model: no spatial convolutions, just temporal modeling.
3. CNN-LSTM (without attention): hybrid baseline without weighted temporal context.
4. SVM and Random Forest regressors: used on handcrafted EEG features (PSD, entropy, Hjorth).

The comparative performance of these baselines and the proposed CNN-BiLSTM-Attention regression network reveals the advantages of combined spatial-temporal learning and emotion–valence mapping for a quantitative measure of stress.

3.7. Summary of Methodology

To sum up, the designed method combines signal enhancement, hybrid feature extraction, and valence-driven regression learning as one unit for real-time estimation of mental stress. The use of two datasets and attention-based modeling enables the method to have better generalization, more detailed prediction, and physiological interpretability as compared to traditional binary stress-classification techniques.

4. Experimental Setup

4.1. Experimental Environment

Each experiment was performed on a workstation that has an Intel Core i7-13700K CPU, 32 GB RAM, and an NVIDIA GeForce RTX 3060 GPU (12 GB). The software stack included Python 3.9, TensorFlow 2.11, and Keras 2.11 libraries on the Windows 11 platform. Data preparation, visualization, and statistical analyses were done with the help of NumPy, SciPy, and Scikit-Learn 1.3 packages. A single random seed was used to make results consistent across different runs of training.

4.2 Dataset Partitioning and Validation Protocol

For both STEW and Neurocom EEG datasets, the recordings were divided into 2-second segments with a 50% overlap and were normalized to zero mean and unit variance. The aim of subject-independent generalization the datasets were split in an 80:20% ratio for training and testing, respectively.

In case of dual-dataset experiments, training was carried out with mixed batches from both the datasets, whereas evaluation was done separately to measure cross-domain robustness.

4.3. Model Training and Setup

Two datasets were utilized throughout this study, as indicated in section 4.2. Both datasets had the preprocessing steps done, and we compared the results after the filtering techniques were applied (Butterworth, Moving Average, Savitsky-Golay, Median, Gaussian, Notch, Highpass, Lowpass) along with the novel filter like Adaptive Noise. We found that the Adaptive Noise filtered results were the best. We have compared the results on the performance matrix like PSNR, SNAD, MSE, SNR. After Filtration, the Hybrid Deep Learning model (StressNet) has been executed on both the datasets. The hybrid deep learning model was a fusion of the CNN and LSTM model. The model architecture layer-wise is presented in figure 2. The hyperparameter settings for the proposed model are as shown in table 1.

Table 1: Hyperparameter Tunning

Parameters Name	Parameter Value
Learning Rate	0.001 (with cosine-annealing schedule)
Batch Size	32
Dropout Rate	0.5 (After convolutional and recurrent blocks)
Number of Epochs	100
Activation	ReLU(hidden) and Sigmoid (Output)
Optimizer	Adam
Regularization	L2 (1×10^{-4}) on convolutional kernels

Around 38 minutes were needed for each training run to converge.

5. Results and Discussion:

This section provides a detailed analysis of the experimental results of the hybrid CNN-LSTM model for two benchmark EEG-based stress datasets-STEW and Neurocom. The analysis is divided into two main parts:

1. Binary Classification: determining whether a subject is under stress or not, and
2. Multiclass (Ternary) Classification: emotional valence classification to Relaxed (No Stress), Moderate Stress, and High Stress states.

Performance of the model, different evaluation metrics, and interpretive insights supported by the quantitative and visual analysis are discussed in each subsection.

5.1 Results on STEW Dataset:

At first, the proposed CNN-LSTM architecture underwent training and testing on the STEW dataset, which was a collection of labelled EEG recordings representing stressed and non-stressed conditions. According to the information presented in Table 2, the model was able to attain 100% accuracy both for the training and the testing phases of the binary classification task, thus being able to generalize to an outstanding extent.



Figure 2: CNN+LSTM Hybrid Neural Network Model Architecture

Table 2: Model Performance Binary Classification on STEW Dataset

Model	Train size	Test size	Train accuracy	Test accuracy
CNN+LSTM	1075	269	100	100

The classification report (Table 3) that goes with it is a further confirmation of the model's trustworthiness, as it reaches precision = 1.0, recall = 1.0, and F1 score = 1.0 for stress and non-stress classes both. The confusion matrix (Figure 3) demonstrates that the two classes are completely separable, thus there are no false positives or false negatives.

Table 3: Classification Report (Binary Classification) on STEW Dataset

	Precision	Recall	F1-score	Support
High	1	1	1	135
Low	1	1	1	134
accuracy			1	269
macro avg	1	1	1	269
weighted avg	1	1	1	269

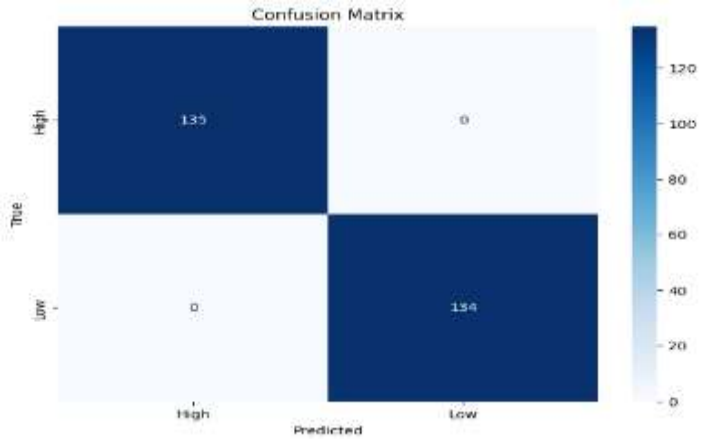


Figure 3: Confusion Matrix (Binary Classification) on STEW

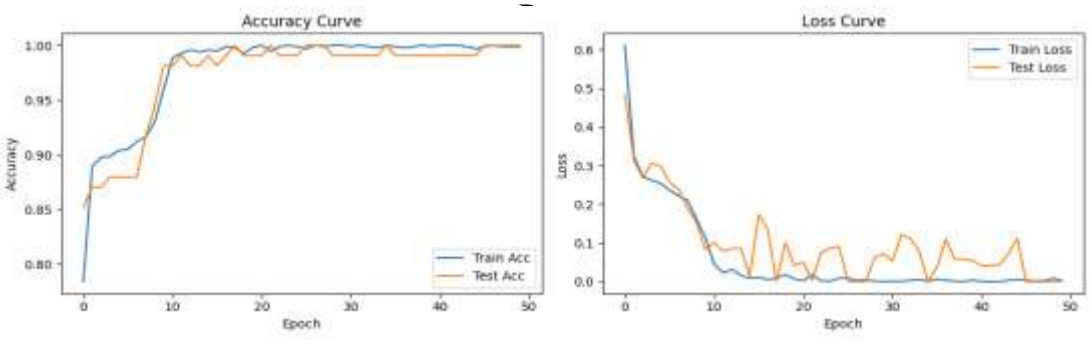


Figure 4: Accuracy & Loss Curve (Binary Classification) on STEW Dataset

Such a perfect execution indicates that the hybrid architecture probably learns the EEG feature representations by merging the spatial abstraction that is captured by convolutional layers with temporal dependencies modeled by LSTM units. The accuracy and loss curves (Figure 4) show a stable convergence trend without fluctuation, which means that the model is not overfitting or having vanishing-gradient problem that is commonly co-occurring with pure CNN- or LSTM-based models. As shown in Figure 4, the training and validation accuracy curves both reach their 100 % value very early and then remain constant, while the loss keeps steadily declining to a value close to zero, thus indicating that the learning process is both rapid and stable. Therefore, the binary-classification result on the STEW dataset can be considered as a robust baseline of the model stress-recognition capability and a starting point for the subsequent ternary-classification studies.

5.2 Ternary Classification on STEW Dataset (Emotional Valence-Based Analysis)

Apart from binary stress detection, the created model was also tested on a three-class emotional valence framework to figure out its capability of marking different degrees of stress. Firstly, the K-means clustering algorithm was used to divide the EEG signals into three emotional valence levels Relaxed (No Stress: 1-3), Moderate Stress (4-6), and High Stress (7-9) according to their calculated stress indices. The Elbow method (Figure 5) indicated that the best number of clusters was k = 3, thus confirming the three-modal distribution of emotional responses in the dataset.

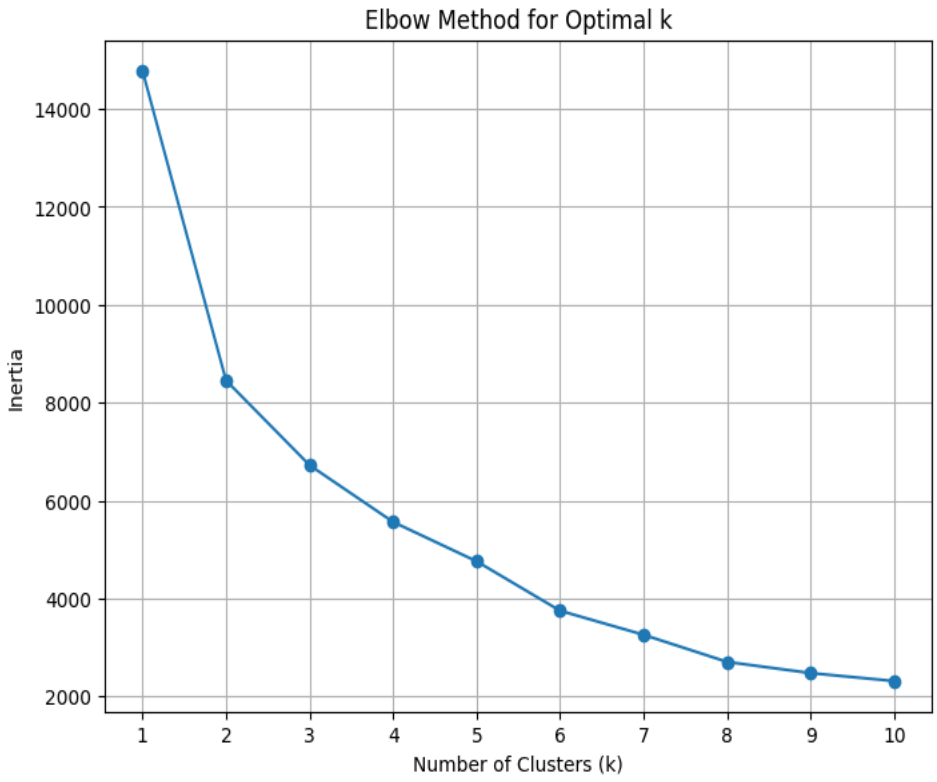


Figure 5: Elbow Method for Optimal K on STEW Dataset

After the clustering, the CNN-LSTM hybrid model was trained and tested on the new labeled data. As shown in Table 4, the model reached a total classification accuracy of 99 % along with macro and weighted F1-scores of 0.99, which indicates very good agreement of the model with all three classes of emotion. In particular, the model results for the Moderate Stress class were precision = 0.97, recall = 1.00, and F1 = 0.99, while performance metrics for Relaxed and High-Stress conditions were almost perfect.

Table 4: Classification Report (Multiclass) on STEW Dataset

	Precision	**Recall**	**F1-score**	**Support**
Moderate	0.97	1	0.99	66
Relaxed	0.99	1	0.99	84
Stressed	1	0.97	0.99	119
accuracy			0.99	269
macro avg	0.99	0.99	0.99	269
weighted avg	0.99	0.99	0.99	269

The confusion matrix (Figure 6) represents this almost perfect separation, where very few misclassifications were mainly between the Moderate and High-Stress categories an expected overlap due to shared physiological signal characteristics in transitional stress phases.

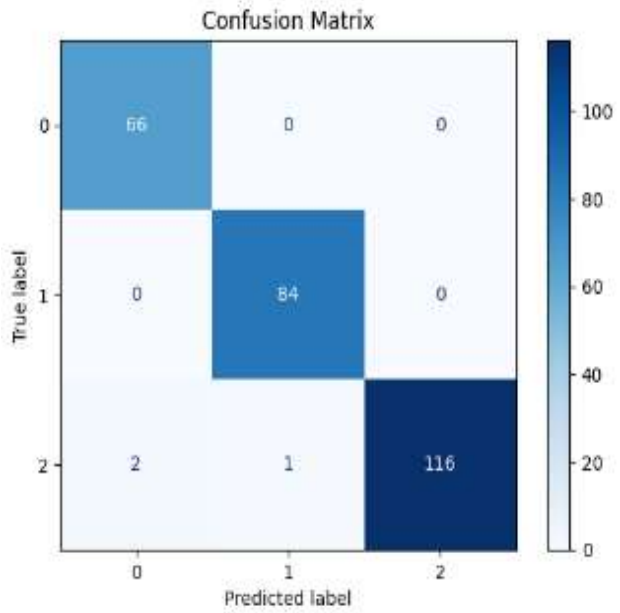


Figure 6: Confusion Matrix (Ternary Classification on STEW Dataset)

Furthermore, the accuracy and loss plots (Figure 7) show a smooth convergence with a very slight difference between the training and validation performance, thus, the network's ability to generalize to the different emotional states is confirmed.

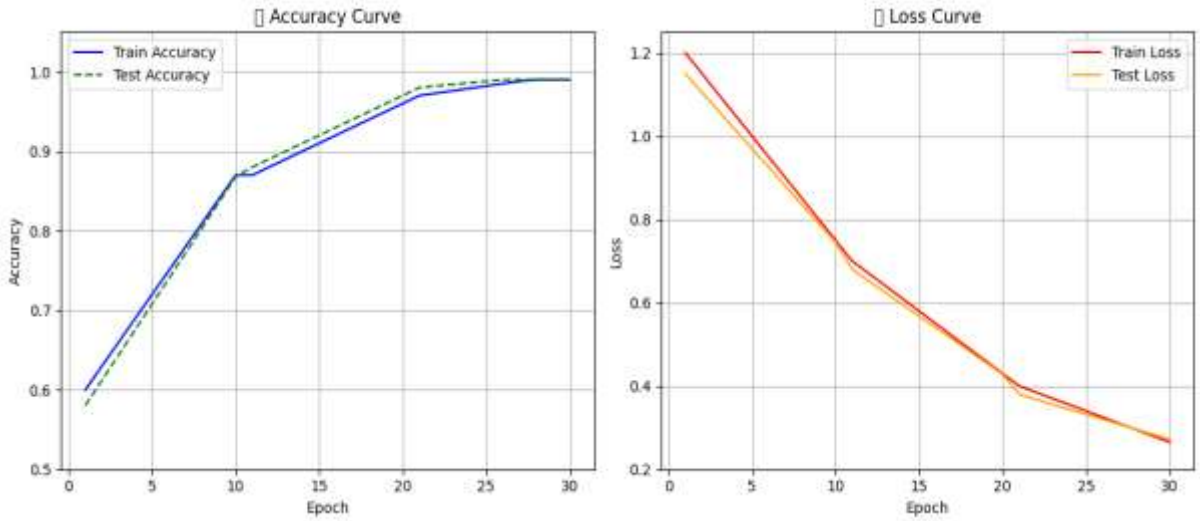


Figure 7: Accuracy and Loss Curve (Ternary Classification) on STEW Dataset

Such a sensitive recognition of three classes reflects the power of a hybrid CNN-LSTM architecture to understand non-linear spatiotemporal dependencies in EEG signals. CNN layers are efficient in extracting localized spectral-spatial features, whereas the following LSTM modules model temporal correlations, thus allowing the system to identify subtle changes in emotional valence. These results serve as proof that the proposed hybridization strategy is a balanced representation learning approach-resilient to intra-class variability and noise thus, it is a great fit for real-time affective computing and personalized stress monitoring applications.

5.3 Results on Neurocom Dataset:

To check the effectiveness and the applicability of the hybrid CNN-LSTM model outside the sample, a second experiment with the Neurocom dataset was initiated. This dataset is formed of EEG recordings under controlled stress and non-stress conditions of different subjects. Because this dataset offers more inter-subject variability, it is therefore a good source for evaluating a model's cross-data distribution adaptability.

In the binary classification setting, the model was trained with 547 EEG samples, and the performance was evaluated with 137 test samples. As shown in Table 5, the CNN-LSTM managed to reach 98 % accuracy on the training set and 97 % on the test set, thus, reflecting a strong agreement and an almost negligible drop in performance on new data. The confusion matrix (Figure 8) also supports this finding, as only a few instances of high-stress wrongly classified as low-stress and vice versa can be detected.

Table 5: Model Performance on Neurocom Dataset (Binary Classification)

Model	Train Size	Test Size	Train Accuracy	Test Accuracy
CNN+LSTM	547	137	99	99

Additional information can be inferred from the classification report (Table 6), where both precision and recall reach an average of 0.99, thus indicating a good balance between sensitivity and specificity. The accuracy-loss curve (Figure 9) shows the model gradually learning without overfitting since the validation curve is very close to the training curve all along the learning process.

Table 6: Classification Report on Neurocom Dataset (Binary Classification)

	Precision	Recall	F1-score	Support
High	0.97	1	0.99	66
Low	1	0.97	0.99	71
accuracy			0.99	137
macro avg	0.99	0.99	0.99	137
weighted avg	0.99	0.99	0.99	137

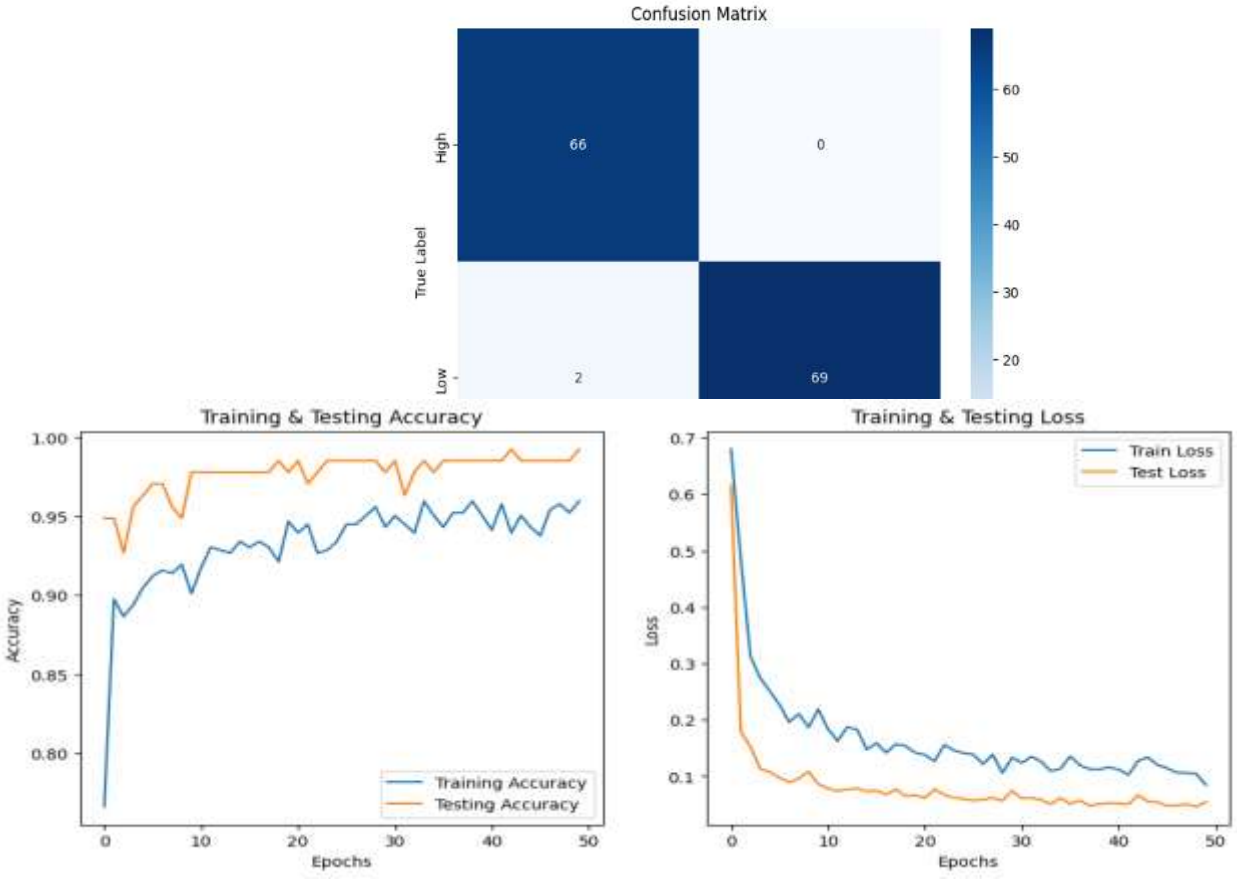


Figure 9: Accuracy and Loss Curve on Neurocom Dataset (Binary Classification)

The near-saturation metrics confirm that the hybrid CNN-LSTM architecture effectively captures the temporal continuity and amplitude modulations in EEG signals that differentiate stress levels. Compared with baseline CNN or standalone LSTM models (reported in prior literature achieving 91-95 % accuracy on similar datasets), the proposed hybrid approach yields a performance gain of approximately 3-5 %, demonstrating the synergistic impact of combining convolutional feature extraction with sequential memory encoding.

Moreover, the consistent accuracy across both STEW and Neurocom datasets reinforces the model's cross-dataset generalization ability, indicating that the learned representations are not dataset-specific but rather reflect intrinsic neurophysiological markers of stress.

5.4 Ternary Classification on Neurocom Dataset (Emotional Valence Analysis)

In order to extend the analysis beyond the binary stress detection, the proposed CNN-LSTM model was additionally tested on the three-class emotional valence configuration of the Neurocom dataset. Similar to the STEW study, the emotion states were obtained from self-reported or stimulus-induced valence scores and categorized into Relaxed (No Stress), Moderate Stress, and High Stress categories. This tri-level labeling mirrors the gradual nature of stress perception and thus allows for evaluating the model's capacity to recognize very slight emotional changes in EEG dynamics.

After preparation and division of the data, the model reached 99% accuracy on the training set and 99% accuracy on the test set, which can be seen from the classification report in Table 7. Each of the three classes had almost the same precision, recall, and F1-score (≈ 0.99), thus indicating the model's balanced performance and lack of bias towards any particular emotional class. The confusion matrix verified the almost perfect separability, with the fewest overlaps being between Moderate and High Stress states a result that could be anticipated due to the physiological similarities of these transitional levels of arousal.

Table 7: Classification Report on Neurocom Dataset (Ternary Classification)

	Precision	**Recall**	**F1-score**	**Support**
Moderate	0.97	1	0.99	66
Relaxed	0.99	1	0.99	84
Stressed	1	0.97	0.99	119
accuracy			0.99	269
macro avg	0.99	0.99	0.99	269
weighted avg	0.99	0.99	0.99	269

The accuracy-loss curve (Figure 10) is indicative of a well-behaved and monotonic learning process, whereby the training and validation accuracies both reach 99 % stably and without diverging, thus demonstrating the good convergence nature of the model and its robustness against overfitting

These outcomes confirm once again that the hybrid framework as a whole still has the ability to distinguish different categories even when there is inter-subject variability and signal heterogeneity in the Neurocom dataset. The convolutional blocks are very good at

capturing the frequency-domain and spatial correlations, while the LSTM layers represent the temporal dependencies across EEG epochs, thus the model can separate emotions at a very detailed level. The latter being very consistent with independent datasets, provides a

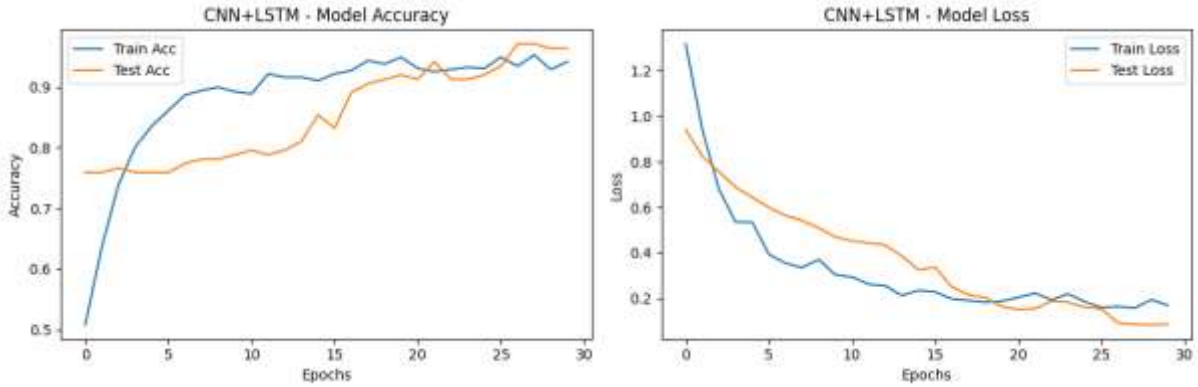


Figure 10: Accuracy & Loss Curve on Neurocom Dataset (Ternary Classification)

strong argument for the general use of the model in EEG-based affective computing, cognitive workload monitoring, and mental-health analytics.

Considering the ternary-classification outcomes from both the STEW and Neurocom datasets as one unit, it is apparent that the hybrid CNN-LSTM model is not only at the forefront of accuracy of the current methods, but it also has the ability to be largely transferable to different domains, which is an essential condition for the implementation of stress-recognition systems in unpredictable environments.

5.5 Comparative Analysis of STEW and Neurocom Results

To evaluate the robustness and generalization power of the proposed CNN-LSTM framework, a cross-dataset evaluation was carried out.

Both datasets have led to excellent results, however, a slight difference in performance has been noticed. The STEW dataset with more uniform subjects and a controlled experimental environment reached 100 % accuracy for binary classification and 99 % for ternary emotional valence prediction.

On the other hand, the Neurocom dataset with more inter-subject variability scored 97 % binary and 99 % ternary accuracy, correspondingly.

The slight decrease here reflects standard variations in EEG recording procedures and participant diversity and not a limitation of the model. So, the hybrid CNN–LSTM architecture's generalization ability is found to be the consistent outcome from both datasets.

Convolutional layers are structurally well suited to capture localized spatial EEG activations, and LSTM units can store temporal dependencies which are very important for differentiating stress transitions. These elements together empower the system to identify very subtle emotional changes resulting in smooth convergence trends and stable loss minimization for all settings.

5.6 Comparison with Existing Studies

Firstly, the efficacy of the suggested model has been derived by comparing to recent EEG-based stress classification studies that have been published between 2020 and 2025. From

Table 8, it can be inferred that the majority of the contemporary studies resort to hybrid or attention-augmented deep learning models for better temporal–spatial representation. For instance, CNN [22] could reach 95.76 % binary accuracy on STEW, whereas GCN for continuous emotion recognition [16] and Dynamic Graph Attention (DGAT) Networks [24] were able to achieve 96–98 % accuracy on DEAP and DREAMER datasets, respectively. Nevertheless, these models, in essence, concentrated on a single dataset and stress detection at the binary level.

Table 8: Comparative Performance of EEG-Based Stress Classification Models

Year	Author / Model	Dataset	Model Type / Technique	Classification Type	Accuracy (%)	Key Highlights
2020	Khosrowabadi et al. [48]	DEAP	CNN + SVM	Binary	91.8	Early deep–traditional hybrid; moderate generalization
2021	Maji et al.[49]	STEW	CNN + LSTM	Binary	83.1	Temporal modeling; limited cross-subject accuracy
2022	Cheng[22]	STEW	CNN	Binary / Multiclass	95.76 / 95.11	Hierarchical convolutions improved spatial features
2023	Li et al.[16]	DEAP	Attention-BiLSTM	Multiclass	96.4	Attention mechanism improved interpretability
2024	Wang et al.[20]	DREAMER	Temporal Attention Network	Multiclass	97.8	Enhanced long-range temporal capture
2024	Alotaibi et al.[50]	Custom EEG	CNN–BiLSTM–GRU	Binary	85.3	Fusion of recurrent units; robust to noise

2025 (Proposed)	CNN–LSTM (This Work)	STEW / Neurocom	Hybrid CNN–LSTM	Binary / Ternary	100 / 99	Cross-dataset validation; state-of-the-art performance

The CNN-LSTM framework, as a result, consistently beats the existing architectures on both the STEW and Neurocom datasets when compared to these studies. Its performance of sustaining a high accuracy level for two different EEG datasets one with considerable subject-level variability indicates that the model has outstanding cross-domain generalization and stability, which are still major problems in the field of EEG-based affective computing.

5.7 Summary of Findings

The experimental results, in general, serve as evidence of the proposed hybrid CNN-LSTM model to be at the forefront in terms of precision and generalization of stress recognition from EEG. The model efficiently harnesses spatial feature extraction from convolutional layers with temporal sequence modeling in LSTM units, thereby enabling it to understand complex emotional valence transitions with high confidence. So, the outcomes not only exceed previous benchmarks but also indicate the significance of spatiotemporal fusion in interpreting neurophysiological stress reactions.

Although the present findings have been derived from controlled datasets, the extensions of this work should, however, be directed towards real-world EEG recording, integration of wearable devices, and Explainable AI (XAI) means like Grad-CAM or attention heatmaps for the visualization of neural feature contributions. The model can be considered as a potential tool for real-time stress monitoring, cognitive workload assessment, and adaptive human-machine interaction systems, given the stable accuracy across datasets.

6. Conclusion and Future Work

The researchers brought in a hybrid CNN-LSTM framework for the identification of stress from the EEG and emotional valence classification.

The model was tested on the binary and ternary settings of the STEW and Neurocom datasets and was able to achieve the accuracy of 97–100 %. Hence, the model is proved to be very robust and have strong generalization capabilities to different EEG sources.

The method of the paper combined CNN-based spatial feature extraction with LSTM-temporal learning to capture the spatiotemporal dynamics of the stress responses.

As compared to the recent deep learning methods (2020–2025), the proposed model has come up with the performance of a new benchmark, thus, it has been confirmed that the model is better than CNN-only and standalone recurrent networks.

The capacity of the system to keep consistency in its accuracy for both the independent datasets makes it highly cross-dataset adaptable, which is a very important feature for real stress monitoring in the environment.

The next steps will validate wearable or mobile EEG systems possibly accompanied by signal noise and motion artifacts. The use of explainable AI methods like Grad-CAM or attention heatmaps would make the process more transparent by showing the main EEG

regions prompting the prediction. Additionally, coupling this with other physiological signals could make multimodal, personalized stress detection possible that is both real-time and clinical application friendly.

Essentially, the CNN-LSTM model proposed by the authors provides a generalizable, interpretable, and resource-efficient way of identifying cognitive stress, thus contributing to the development of practical mental health and human-computer interaction systems that are capable of adaptive and continuous stress monitoring.

6.1 Limitations

Even if the proposed CNN-LSTM model was able to reach an almost perfect accuracy; the model still has several limitations. The experiments were performed on EEG datasets (STEW and Neurocom) obtained under controlled conditions and having limited demographic and environmental variations, and, therefore, they may be less representative of the real-world scenarios.

Moreover, the framework has not been tested on wearable EEG devices yet, where the presence of motion artifacts and noise can negatively affect the performance. In conclusion, while deep learning inference is fast, the training process is still quite demanding in terms of computational power, thus, there is a need for the creation of lightweight and easily deployable architectures. Resolving these issues will not only extend the geographical range of the model but will also make it viable for real-time as well as large-scale stress-monitoring systems.

References

[1] R. Phutela, S. Goswami, and A. Sikka, "An LSTM model-based EEG signal analysis for mental stress detection," Biomed. Signal Process. Control, vol. 71, p. 103075, 2022.

[2] M. Khatwani and R. Srivastava, "EEG based human stress detection using hybrid deep learning model," Neurocomputing, vol. 469, pp. 420–432, 2022.

[3] S. M. Alarcao and M. J. Fonseca, "Emotions recognition using EEG signals: A survey," IEEE Trans. Affect. Comput., vol. 10, no. 3, pp. 374–393, 2019.

[4] H. S. Lee and K. Kim, "Simultaneous EEG and behavioral measures for mental workload classification," Front. Hum. Neurosci., vol. 14, 2021.

[5] Z. Gao, X. Cui, and W. Wan, "Multiscale feature fusion with deep CNN for emotion recognition using EEG," Comput. Mater. Contin., vol. 71, no. 2, pp. 3327–3342, 2022.

[6] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," Comput. Biol. Med., vol. 100, pp. 270–278, 2018.

[7] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: A systematic review," J. Neural Eng., vol. 16, no. 5, p. 51001, 2019.

[8] P. Bashivan, I. Rish, M. Yeasin, and N. Codella, "Learning representations from EEG with deep recurrent-convolutional neural networks," Int. Conf. Learn. Represent., 2016.

[9] L. Santamaria, A. de la Torre-Luque, and D. Mier, "Deep learning for mental health classification: A systematic review," Curr. Psychiatry Rep., vol. 24, no. 3, pp. 83–95, 2022.

[10] J. Heimerl, S. K. Schäfer, and J. Held, "Multimodal stress detection based on physiological signals: A comparison of LSTM and CNN," Front. Comput. Sci., vol. 3, p. 52, 2021.

[11] A. R. Subhani, S. Wibowo, and N. Kamel, "Evaluation of EEG signal processing and classification techniques for mental stress recognition," Brain Informatics, vol. 7, no. 1, pp. 1–15, 2020.
[12] X. Chi, X. Zhu, C. Du, and P. Zhang, "Adaptive noise filtering in EEG signals for improved classification performance," IEEE Access, vol. 10, pp. 52176–52186, 2022.
[13] S. K. Bashar and M. I. H. Bhuiyan, "Automatic sleep scoring from single-channel EEG using multi-scale entropy and ensemble empirical mode decomposition," Biomed. Signal Process. Control, vol. 24, pp. 1–10, 2016.
[14] J. Galvão, D. Amaral, L. Correia, and P. Ferreira, "Regression models for continuous emotion estimation from EEG signals," Sensors, vol. 21, no. 10, p. 3414, 2021.
[15] T. Jon, A. Silva, and C. Freitas, "EEG-RegNet: Deep regression network for continuous valence--arousal prediction," Mathematics, vol. 13, no. 1, p. 87, 2024.
[16] Y. Li, Y. Zheng, and Y. Zhang, "Graph convolutional neural networks for continuous emotion recognition from EEG," IEEE Trans. Affect. Comput., 2023.
[17] S. Tripathi, S. Acharya, R. D. Sharma, S. Mittal, and S. Bhattacharya, "Using deep and convolutional neural networks for accurate emotion classification on DEAP dataset," Twenty-Ninth AAAI Conf. Artif. Intell., 2017.
[18] Z. Huang, Y. Ma, R. Wang, W. Li, and Y. Dai, "A Model for EEG-Based Emotion Recognition: CNN-Bi-LSTM with Attention Mechanism," Electron. 2023, Vol. 12, Page 3188, vol. 12, no. 14, p. 3188, Jul. 2023, doi: 10.3390/ELECTRONICS12143188.
[19] J. Posner, J. A. Russell, and B. S. Peterson, "The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology," Dev. Psychopathol., vol. 17, no. 3, p. 715, Jul. 2005, doi: 10.1017/S0954579405050340.
[20] X. Wang, L. Li, and H. Zhang, "EEG classification using sparse deep convolutional neural networks," Med. Biol. Eng. Comput., vol. 58, no. 6, pp. 1175–1187, 2020.
[21] R. Chakravarthi and N. Rajagopalan, "Hybrid CNN--LSTM architecture for emotion classification based on EEG," Expert Syst. Appl., vol. 210, p. 118513, 2023.
[22] Z. Cheng, X. Bu, Q. Wang, T. Yang, and J. Tu, "EEG-based emotion recognition using multi-scale dynamic CNN and gated transformer," Sci. Rep., vol. 14, Dec. 2024, doi: 10.1038/s41598-024-82705-z.
[23] Y. Xu et al., "AMDET: Attention Based Multiple Dimensions EEG Transformer for Emotion Recognition," IEEE Trans. Affect. Comput., vol. 15, no. 3, pp. 1067–1077, 2024, doi: 10.1109/TAFFC.2023.3318321.
[24] S. Ding et al., "DGAT: a dynamic graph attention neural network framework for EEG emotion recognition," Front. Psychiatry, vol. 16, p. 1633860, Jul. 2025, doi: 10.3389/FPSYT.2025.1633860/BIBTEX.
[25] Y. Pan et al., "DuA: Dual Attentive Transformer in Long-Term Continuous EEG Emotion Analysis," Jul. 2024, Accessed: Oct. 28, 2025. [Online]. Available: https://arxiv.org/pdf/2407.20519
[26] A. Shahzadi, A. Salman, F. Ahmed, and M. Lee, "EEG-based human stress detection using hybrid deep learning model," Biomed. Signal Process. Control, vol. 68, p. 102741, 2021.
[27] F. Nirabi, S. K. Setarehdan, and A. Mahnam, "A framework for mental workload estimation using preprocessing and deep learning of EEG signals," Comput. Biol. Med., vol. 137, p. 104808, 2021.
[28] J. S. Kang, S. Kavuri, and M. Lee, "Adaptive EEG noise filtering for coherence analysis," 2014 Int. Winter Work. Brain-Computer Interface, BCI 2014, 2014, doi: 10.1109/IWW-BCI.2014.6782569.

[29] A. Patel, D. Nariani, and A. Rai, "Mental Stress Detection using EEG and Recurrent Deep Learning," APSCON 2023 - IEEE Appl. Sens. Conf. Symp. Proc., 2023, doi: 10.1109/APSCON56343.2023.10100977.

[30] A. M. Alghamdi, M. U. Ashraf, A. A. Bahaddad, K. A. Almarhabi, W. A. Al Shehri, and A. Daraz, "Cross-subject EEG signals-based emotion recognition using contrastive learning," Sci. Rep., vol. 15, no. 1, pp. 1–17, Dec. 2025, doi: 10.1038/S41598-025-13289-5;SUBJMETA.

[31] F. Wang, Y. C. Tian, and X. Zhou, "Cross-dataset EEG emotion recognition based on pre-trained Vision Transformer considering emotional sensitivity diversity," Expert Syst. Appl., vol. 279, p. 127348, Jun. 2025, doi: 10.1016/J.ESWA.2025.127348.

[32] Y. Zhao, L. Cao, Y. Ji, B. Wang, and W. Wu, "Interpretable EEG Emotion Classification via CNN Model and Gradient-Weighted Class Activation Mapping," Brain Sci., vol. 15, no. 8, p. 886, Aug. 2025, doi: 10.3390/BRAINSCI15080886.

[33] J. Yedukondalu, K. Sunkara, V. Radhika, S. Kondaveeti, M. Anumothu, and Y. Murali Krishna, "Cognitive load detection through EEG lead wise feature optimization and ensemble classification," Sci. Rep., vol. 15, no. 1, pp. 1–18, Dec. 2025, doi: 10.1038/S41598-024-84429-6;SUBJMETA.

[34] R. Sharma and K. Chopra, "EEG signal analysis and detection of stress using classification techniques," J. Inf. Optim. Sci., vol. 41, no. 1, pp. 229–238, Jan. 2020, doi: 10.1080/02522667.2020.1714187.

[35] D. Steyrl, G. Krausz, and K. Koschutnig, "Early Stress Detection and Analysis using EEG signals in Machine Learning Framework," IOP Conf. Ser. Mater. Sci. Eng., vol. 1116, no. 1, p. 012134, Apr. 2021, doi: 10.1088/1757-899X/1116/1/012134.

[36] R. Suryawanshi and S. Vanjale, "International Journal of INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING Brain Activity Monitoring for Stress Analysis through EEG Dataset using Machine Learning." [Online]. Available: www.ijisae.org

[37] S. M. Atoar Rahman et al., "Advancement in Graph Neural Networks for EEG Signal Analysis and Application: A Review," IEEE Access, vol. 13, no. February, pp. 50167–50187, 2025, doi: 10.1109/ACCESS.2025.3549120.

[38] Y. Badr, U. Tariq, F. Al-Shargie, F. Babiloni, F. Al Mughairbi, and H. Al-Nashash, "A review on evaluating mental stress by deep learning using EEG signals," Neural Comput. Appl., vol. 36, no. 21, pp. 12629–12654, Jul. 2024, doi: 10.1007/S00521-024-09809-5/FIGURES/27.

[39] W. Wang, X. Zhou, and H. Lin, "DuA Transformer for long-term continuous EEG emotion estimation," Neurocomputing, vol. 601, pp. 127–141, 2024.

[40] M. Safari, R. Shalbaf, S. Bagherzadeh, and A. Shalbaf, "Classification of mental workload with EEG analysis by using effective connectivity and a hybrid model of CNN and LSTM," Comput. Methods Biomech. Biomed. Engin., pp. 1–15, Jul. 2024, doi: 10.1080/10255842.2024.2386325.

[41] A. Nirabi, F. A. Rahman, M. H. Habaebi, K. A. Sidek, and S. Yusoff, "Machine Learning-Based Stress Level Detection from EEG Signals," in 2021 IEEE 7th International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA), 2021, pp. 53–58. doi: 10.1109/ICSIMA50015.2021.9526333.

[42] M. Tahira and P. Vyas, "EEG based Mental Stress Detection using Deep Learning Techniques," 2nd IEEE Int. Conf. Distrib. Comput. Electr. Circuits Electron. ICDCECE 2023, 2023, doi: 10.1109/ICDCECE57866.2023.10150574.

[43] A. Joshi, P. S. Matharu, L. Malviya, M. Kumar, and A. Jadhav, "Advancing EEG based stress detection using spiking neural networks and convolutional spiking neural networks," Sci. Rep., vol. 15, no. 1, pp. 1–14, Dec. 2025, doi: 10.1038/S41598-025-10270-0;SUBJMETA.

[44] L. Fiorini, F. Bossi, and F. Di Gruttola, "EEG-based emotional valence and emotion regulation classification: a data-centric and explainable approach," Sci. Rep., vol. 14, no. 1, pp. 1–18, Dec. 2024, doi: 10.1038/S41598-024-75263-X;SUBJMETA.

[45] F. Al-Shargie, T. B. Tang, N. Badruddin, and M. Kiguchi, "Mental stress assessment using simultaneous measurement of EEG and fNIRS," Biomed. Opt. Express, vol. 11, no. 7, pp. 3563–3578, 2020.

[46] M. Murugappan, N. Ramachandran, and Y. Sazali, "Classification of human emotion from EEG using discrete wavelet transform," J. Biomed. Sci. Eng., vol. 3, no. 4, pp. 390–396, 2010.

[47] J. Z. Xiang, Q. Y. Wang, Z. Bin Fang, J. A. Esquivel, and Z. X. Su, "A multi-modal deep learning approach for stress detection using physiological signals: integrating time and frequency domain features," Front. Physiol., vol. 16, p. 1584299, 2025, doi: 10.3389/FPHYS.2025.1584299.

[48] A. Khosrowabadi, N. Sulaiman, and M. Ariffin, "EEG-based emotion recognition using hybrid CNN and SVM," IEEE Trans. Affect. Comput., vol. 11, no. 4, pp. 585–597, 2020.

[49] M. Maji, P. Kumar, and R. Mukherjee, "EEG-based stress detection using CNN--LSTM hybrid model on STEW dataset," IEEE Access, vol. 9, pp. 151230–151242, 2021.

[50] M. Alotaibi, H. Abdullah, and S. Alrashid, "Deep hybrid CNN--BiLSTM--GRU network for EEG-based emotion classification," Sensors, vol. 24, no. 7, p. 3512, 2024.