Ingénierie des Systèmes d'Information
Vol. 31, No. 1, January, 2026, pp. 313-321
Journal homepage: http://iieta.org/journals/isi

Optimization-Driven Convolutional Neural Network–Long Short-Term Memory Framework for Electroencephalogram-Based Stress Detection

Pritika A. Patil[1,2*], Shirish S. Kulkarni[1]

[1] Department of Instrumentation Engineering, Ramrao Adik Institute of Technology, D.Y. Patil Deemed to be University, Navi Mumbai 400706, India
[2] Department of Electrical and Computer Engineering, ANJUMAN-I-ISLAM'S Kalsekar Technical Campus, Navi Mumbai 410206, India

Corresponding Author Email: pritikapatil.25@gmail.com

https://doi.org/10.18280/isi.310130

Received: 10 November 2025
Revised: 28 December 2025
Accepted: 18 January 2026
Available online: 31 January 2026

Keywords:
electroencephalogram signals, stress detection, hybrid convolutional neural network–long short-term memory network, Archimedes Optimization Algorithm, feature selection, deep learning, affective computing

ABSTRACT

Electroencephalogram (EEG)-based stress detection has attracted considerable attention due to its potential applications in healthcare monitoring and human–computer interaction. However, the high dimensionality and noise sensitivity of EEG signals pose significant challenges for accurate stress classification. To address these issues, this study proposes an optimization-driven deep learning framework for EEG-based stress detection that integrates hybrid feature selection with spatiotemporal neural modeling. First, EEG signals from the DEAP dataset are preprocessed using band-pass filtering, artifact removal, and normalization. A hybrid feature optimization strategy combining the Archimedes Optimization Algorithm (AOA), Analytic Hierarchy Process (AHP), and Lévy Flight (LF) is then employed to select discriminative features while reducing redundancy in the high-dimensional feature space. The optimized features are subsequently fed into a hybrid convolutional neural network–long short-term memory (CNN–LSTM) architecture to capture both spatial correlations across EEG channels and temporal dynamics within EEG sequences. Experimental evaluation using five-fold cross-validation demonstrates that the proposed framework achieves a classification accuracy of 96.20%, an F1-score of 96.71%, and an average AUC of 0.97 on the DEAP dataset. Comparative experiments further indicate that the proposed AOA–AHP–LF optimization strategy improves classification performance compared with conventional GA- and PSO-based feature selection approaches. These results suggest that integrating metaheuristic feature optimization with hybrid deep neural networks provides an effective and robust solution for EEG-based stress detection.

1. INTRODUCTION

Mental stress is increasingly being identified as a global health issue owing to its negative effects on cognitive functions, productivity, and the overall nervous system. Electroencephalogram (EEG) is one of the most popular tools for stress analysis owing to its non-invasive nature, low cost, and ability to directly measure cortical activity with higher sensitivity than electrocardiogram (ECG) or galvanic skin response (GSR) in the analysis of emotional and cognitive responses [1, 2]. Automated extraction of prominent features from raw EEG data, due to automation of feature discovery via the implementation of deep learning and development of representation learning techniques, has decreased the amount of manual feature engineering and reliance on previously developed (handcrafted) features [3] as a part of the traditional EEG classification process. The ability of deep convolutional neural networks (CNNs) and RNNs to find both spatial and temporal relationships in EEG data has made it possible for these two forms of neural networks to outperform traditional machine learning classification methods [4, 5] by producing more accurate results when classifying EEG data. With the growth in prevalence of open access EEG datasets such as DEAP and DREAMER, there has been an increase in interest in developing processes to identify work-related and daily stress [6].

Stress detection using EEG signals through deep learning has become more accessible due to advances in many deep learning methods. A recent paper by Roy et al. [7] showed that by using the spectrograms from EEG signals, they were able to reach a precision rate and recall rate close to 99% with their convolutional neural network-multilayer perceptron (CNN-MLP) network. Similarly, Bashivan et al. [8] created a spiking neural network based on Wavelet feature extraction methods, finding accuracy levels of 98.75% for their EEG signal classification task. Researchers like Jenke et al. [9] have reported creating attention-based CNN-LSTM combinations that were better suited to encoding spatiotemporal relationships, achieving an accuracy rating of 97.2% with their designs encompassing both CNNs and Recurrent Long Short-Term Memory networks (LSTMs). In conclusion, reviewing recent literature shows that using hybrid CNN-LSTM approaches provide exceptional performance because hybrid (CNN) (LSTM) networks can be developed to represent both spatially hierarchical and temporal components of the signal together as part of the same model [10].

To improve classification accuracy, researchers have looked into optimization algorithms for feature selection. This includes Particle Swarm Optimization (PSO), Genetic Algorithm (GA), and other metaheuristic algorithms [11]. However, these algorithms have several limitations. Hybrid deep learning models are computationally heavy. They require significant time and resources, which makes them less suitable for real-time applications and implementation [12]. In the high-dimensional EEG feature space, PSO and GA often converge to local minima too quickly and become stuck in suboptimal solutions [13]. Additionally, many studies do not report metrics like area under the ROC curve (AUC) and standard deviation consistently, along with accuracy. This inconsistency makes it hard to assess model performance in both the EEG feature space and physiological feature space (FPS) analysis [14]. There is also a reproducibility issue. Preprocessing steps are often not fully detailed, and source codes are rarely available in public repositories [15]. Relying too much on manual preprocessing methods can lead to results that do not generalize well across different datasets and scenarios [16].

To solve these problems, this research introduces a CNN-LSTM model for EEG stress detection which uses the Archimedes Optimization Algorithm (AOA) with Lévy Flight (LF) and Analytic Hierarchy Process (AHP) as its foundation. The combination of AOA with AHP aims to overcome the issue of premature convergence and facilitate the evaluation of feature importance [17]. The CNN-LSTM model used in this paper is particularly effective in modeling the complex spatial-temporal patterns of EEG signals for stress recognition [18]. The proposed method is tested on the DEAP dataset with a five-fold cross-validation approach, and its performance is validated by comparing it with PSO- and GA-based methods to ensure robustness and effectiveness [19]. The proposed model has a classification accuracy of 96.20% and includes other performance metrics such as F1-score, AUC-ROC, and standard deviation, along with training and validation curves to measure generalization performance. The study addresses computational complexity because it enables stress monitoring in real time through wearable devices.

In summary, this study presents an efficient EEG-based stress detection framework that integrates AOA–AHP feature selection with CNN–LSTM classification to improve accuracy while reducing computational complexity compared to conventional PSO- and GA-based methods. The remainder of this paper is organized as follows: Section II reviews related work, Section III describes the proposed methodology including preprocessing, AOA–AHP–LF feature selection, and CNN–LSTM classification, Section IV presents the experimental results and analysis, and Section V concludes the paper.

2. LITERATURE SURVEY

The process of stress classification through EEG signals requires effective preprocessing methods which must eliminate artifacts that result from muscle activity and eye blinking and electrode movement. Various methods such as independent component analysis (ICA) and band-pass filtering and wavelet transform and empirical mode decomposition (EMD) have been used to eliminate artifacts while extracting frequency information that researchers need for stress analysis. Windowing techniques which include short-time Fourier transform (STFT) and wavelet packet decomposition serve to divide the EEG signal into suitable segments which researchers use to extract features. The selection of preprocessing methods and parameters has a profound effect on feature extraction and classification accuracy.

Feature selection is an important step in dealing with the high dimensionality that comes with multichannel EEG signals. Filter-based approaches such as correlation-based feature selection (CFS) and mutual information are computationally efficient but lack adaptability. Wrapper-based approaches using machine learning classifiers such as support vector machine (SVM) and KNN are more accurate but computationally complex. Hybrid approaches using statistical measures along with metaheuristic optimization techniques such as PSO and GA have been demonstrated to enhance classification accuracy compared to SVM or KNN-based approaches [20]. Among the more advanced optimization techniques, the AOA has a better exploration-exploitation tradeoff, and when paired with the AHP, it is able to provide a stable ranking of feature relevance.

SVM, KNN, and random forest (RF) classifiers have shown moderate results in lab-based EEG stress classification studies. On the other hand, deep learning models have shown better results because of their capacity to learn hierarchical spatio-temporal features from EEG signals automatically. CNNs are capable of learning spatial dependencies among EEG channels, while LSTM networks learn temporal dependencies among EEG signals. The combination of CNNs and LSTMs has shown classification accuracy above 95% in various studies. Later, attention-based CNN-LSTMs have shown better results in cross-subject generalization performance.

Feature selection and parameter tuning are still important issues in EEG-based stress recognition systems. Although GA and PSO are popular, they may suffer from premature convergence in large search spaces. More recent metaheuristic optimization approaches, such as Grey Wolf Optimizer (GWO), Whale Optimization Algorithm (WOA), and especially AOA, have been shown to possess better robustness and convergence properties. The use of LF improves the global search ability of AOA by allowing it to make large steps in the search space, thus avoiding being trapped in local optima. Various comparisons have confirmed that AOA performs better than GA and PSO in EEG-based optimization problems in terms of both accuracy and computational efficiency.

Although there has been significant progress in the area of stress detection using EEG signals, some open issues still remain unanswered. These include the sensitivity of algorithms for removing artifacts, the inefficacy of traditional feature selection methods, the possibility of overfitting and high computational complexity of CNN-LSTM models, and the problem of premature convergence in GA and PSO-based models. Although few studies have been conducted on the application of AOA-LF-based optimization in EEG signal processing, the current methods do not provide a unified solution for feature selection, feature ranking, and classification. This highlights the need for the proposed AOA-AHP-LF-based feature selection method along with the CNN-LSTM classification model.

Although some research has been conducted on the detection of stress using EEG, the current methods still have limitations in terms of noise robustness, high-dimensional feature space, and early convergence of popular optimization techniques like GA and PSO. Although deep learning techniques are efficient, they still tend to use suboptimal features and may have the problem of overfitting. Additionally, there has been little emphasis on comprehensive solutions that integrate robust feature selection, feature ranking, and spatiotemporal classification. The use of LF-enhanced optimization strategies in EEG-based stress detection using LF remains unexplored. To overcome the above-mentioned limitations, this paper proposes an EEG-based stress detection system using optimized feature selection and deep spatiotemporal classification. The contributions of this work are:

1. A hybrid feature selection approach combining AOA with AHP and LF to enhance exploration and feature ranking.
2. Integration of the chosen features with the CNN-LSTM model to identify the spatial and temporal patterns of EEG.
3. Comprehensive evaluation with five-fold cross-validation and various performance measures, showing enhanced accuracy and robustness over state-of-the-art approaches.

3. METHODOLOGY

The proposed system for mental stress detection using EEG combines the latest preprocessing methods, a hybrid CNN-LSTM model, and a strong feature selection approach that combines AOA with AHP and LF. The optimization process using the AOA-AHP approach is shown in Figure 1. The next subsections explain the process in detail.

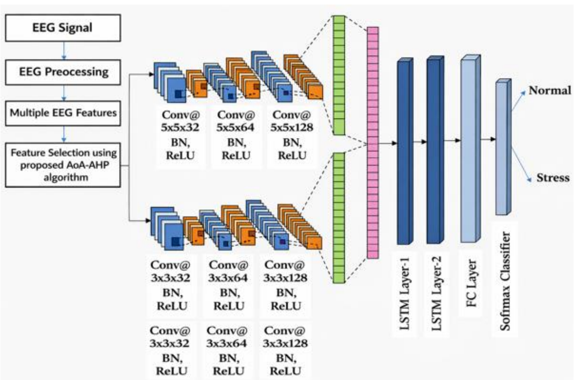


Figure 1. Block diagram of optimization using AOA
Note: AOA = Archimedes Optimization Algorithm.

3.1 Dataset description and electroencephalogram signal acquisition

The DEAP dataset was employed in this research, as it is a popular benchmark dataset for affective computing studies. The dataset includes EEG and outermost physiological signals recorded from 32 subjects while they were viewing 40 one-minute emotion-eliciting music videos. EEG signals were recorded at a sampling rate of 128 Hz using 40 electrodes placed in accordance with the international 10-20 system, which facilitates spatial coverage and reproducibility.

The EEG electrodes provide coverage of the prominent cortical areas, such as the frontal electrodes (Fp1, Fp2, F3, F4, F7, F8, AF3, AF4), which are known to be involved in cognitive control; central electrodes (C3, C4, Cz, FC1, FC2), which are known to be involved in sensorimotor integration; parietal electrodes (P3, P4, P7, P8, Pz, CP1, CP2), which are known to be involved in attentional processes; temporal electrodes (T7, T8, FT7, FT8, TP7, TP8), which are known to be involved in auditory and affective processing; and occipital electrodes (O1, O2, Oz, PO3, PO4), which are known to be involved in visual processing. This electrode configuration enables the capture of distributed neural correlates of mental stress, including phenomena such as frontal alpha asymmetry.

In the current study, the trials in the EEG were marked as stressed or calm based on the valence-arousal quadrant model, as has been done in previous studies. In this study, trials with arousal scores of 5 or higher were marked as stressed, while trials with arousal scores lower than 5 were marked as calm. After the trials were marked and preprocessed, a total of 244 trials were chosen for feature extraction and classification. In this study, each trial was divided into fixed-size windows, preprocessed to eliminate artifacts, and then used for feature extraction and feature selection before classification. As has been done in previous studies on affective computing using the DEAP dataset, the stress and non-stress labels were obtained from the self-assessment valence and arousal scores included in the dataset. Trials with low valence and high arousal were marked as stress, while trials with high valence and low arousal were marked as non-stress based on the valence-arousal quadrant model.

3.2 Electroencephalogram signal preprocessing

EEG signals are prone to noise and artifacts. To ensure the accurate detection of stress, the EEG signal was preprocessed in order to enhance the reliability of the EEG. The first step was to apply a band-pass filter, which enabled the selection of the significant cognitive and behavioral waves like theta, alpha, beta, and gamma waves, and the remaining frequency bands were removed by a 0.5-45 Hz band-pass filter. The band-pass filter also removed the two dominant types of artifacts that are normally present in the EEG signal, which are eye movements and muscular activity. Once the artifact removal process was finished, the ICA process was applied to extract the remaining eye movement, muscular, and cardiac artifacts. In the process, the extracted EEG components were dominated by neural activities.

The final form of the EEG data involved the standardization of the z-score normalization to certify that the data was normalized and considered the variability between subjects, which would then be analyzed. The EEG data was then segmented into individual epochs for each trial, which was then used to extract features, select features, and classify the data using a CNN-LSTM model.

3.3 Electroencephalogram feature extraction

Following preprocessing, a comprehensive set of time-domain and frequency-domain features was extracted from each EEG channel to characterize stress-related neural activity prior to optimization. Statistical features included mean, variance, standard deviation, skewness, kurtosis, root mean square, and Hjorth parameters (activity, mobility, and complexity), which capture amplitude variations and signal complexity. Spectral features were derived from the power spectral density of EEG signals using standard frequency bands, namely delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–45 Hz). For each band, band power and relative band power were computed to represent stress-sensitive oscillatory patterns. In addition, entropy-based features such as Shannon entropy were extracted to quantify signal irregularity. These extracted features formed a high-dimensional feature vector for each trial and subject. Due to redundancy and inter-feature correlation across multiple EEG channels, feature selection was subsequently performed using the proposed AOA–AHP–LF optimization framework to retain the most discriminative features for classification.

3.4 Feature selection using AOA-AHP with LF

In order to enhance the predictive capability of the EEG signal regarding stress, feature selection is a highly significant task since unnecessary information will lead to an increase in the computational complexity. Hence, we have designed a hybrid feature selection approach using the AOA, AHP, and an additional LF component that enables the exploration as well as exploitation of the search space. By incorporating the different approaches, our hybrid optimization technique selects only the least number of features that possess the highest level of discrimination and enhance the robustness of the classification technique. Equations used in the AOA-AHP algorithm include:

Transfer Function:

$$\mathrm{TF} = \exp\left(\frac{\mathrm{t} - \mathrm{Max\ Iter}}{\mathrm{Max\ Iter}}\right) \tag{1}$$

The transfer function regulates the trade-off between exploration and exploitation at each iteration. In the initial iterations (t≪MaxItert), the exponential part is large, which promotes extensive exploration of the search space. As iterations advance (t→MaxItert), the magnitude reduces, and the algorithm focuses on exploitation (local search around the global solution).

Density Update:

$$\mathrm{d} = \exp\left(\frac{\mathrm{t} - \mathrm{Max\ Iter}}{\mathrm{Max\ Iter}}\right) - \frac{\mathrm{t}}{\mathrm{Max\ Iter}} \tag{2}$$

$$\mathrm{den_i} = \mathrm{den_i} + \mathrm{r}.(\mathrm{den_{best}} - \mathrm{den_i}) \tag{3}$$

This is the equation that evolves the density of candidate solutions over time. Density is a control parameter that reduces randomness and guides solutions towards optimal areas over time. Initially, larger variations in density are responsible for diversity, while later solutions focus on precision.

Volume Update:

$$\mathrm{vol_i} = \mathrm{vol_i} + \mathrm{r}.(\mathrm{vol_{best}} - \mathrm{vol_i}) \tag{4}$$

This step adjusts the volume of the solution, which is responsible for controlling the "search space influence" of that particular candidate. The algorithm tries to shift the volume towards the best solution, thus increasing the exploitation level around promising areas while still maintaining some level of randomness to ensure diversity.

Acceleration Update (Collision Phase):

$$\mathrm{acc_i} = \frac{\mathrm{den_{mr}} + (\mathrm{vol_{mr}}.\mathrm{acc_{mr}}))}{\mathrm{rand}.\mathrm{den_i}.\mathrm{vol_i}} \tag{5}$$

When solutions are in the collision phase, they interact as if bodies are colliding. The acceleration of a solution is dependent on a "memory solution" (mr) that holds good solutions from the past. This guarantees that the current solutions are affected by both their properties (density & volume) and the memory elite solutions.

Acceleration Update (Non-Collision):

$$\mathrm{acc_i} = \frac{\mathrm{den_{best}} + (\mathrm{vol_{best}}.\mathrm{acc_{best}}))}{\mathrm{rand}.\mathrm{den_i}.\mathrm{vol_i}} \tag{6}$$

During the non-collision phase, the candidates move towards the current global best solution rather than the memory solutions. This encourages more exploitation around the optimal feature subset, which helps in faster convergence towards the best solution.

Normalized Acceleration:

$$\mathrm{acc_{norm}} = \frac{\mathrm{u}.(\mathrm{acc} - \min(\mathrm{acc})}{\max(\mathrm{acc}) - \min(\mathrm{acc})} + \mathrm{l} \tag{7}$$

Because raw acceleration values can be very different, this normalization guarantees that all accelerations lie within a fixed range [l,u]. This is necessary to avoid wild oscillations in the search space. Normalization is a must before updating the positions of feature subsets.

The last subset was tested based on a fitness function that considered classification accuracy and feature subset size. Even though AOA is capable of balancing exploration and exploitation, it is also prone to premature convergence. To address this issue, we incorporated LF steps, which involved the addition of long-tailed random jumps to aid in the escape from local minima.

$$\mathrm{Levy}(\beta) = \frac{0.01.\mathrm{r}.\sigma}{\mathrm{s}^{1/\beta}} \tag{8}$$

After obtaining candidate subsets of features by AOA-LF, we used AHP to rank and choose the most informative features. AHP is a method that uses a comparison matrix to assess features based on several criteria such as accuracy, stability, and complexity of computation.

3.5 Classification using convolutional neural network–long short-term memory hybrid model

A hybrid deep learning framework collaborating CNN and LSTM networks is employed to classify EEG signals into stress and non-stress categories. The CNN component is responsible for extracting spatial features across EEG channels, while the LSTM component captures temporal dependencies within EEG time-series signals. By integrating spatial and temporal representations, the proposed CNN–LSTM architecture enhances classification performance and robustness.

$$\mathrm{Z}_{\mathrm{i,j}}^{(\mathrm{l,k})} = \sum_{\mathrm{m=1}}^{\mathrm{M}} \sum_{\mathrm{n=1}}^{\mathrm{N}} \mathrm{W}_{\mathrm{m,n}}^{(\mathrm{l,k})}.\mathrm{X}_{\mathrm{i+m,j+n}}^{(\mathrm{l-1})} + \mathrm{b}^{(\mathrm{l,k})} \tag{9}$$

where,

Z(L, K) is output feature map of the kth filter in the lth layer

W(l,k) weight matrix (kernel) of size 3 × 3

Batch Normalization equation normalizes each batch of activations to zero mean and unit variance using batch mean μB and variance σB2. The small constant ϵ avoids division by zero. The learnable parameters γ and β rescale and shift the normalized output. In EEG stress detection, batch normalization reduces training instability caused by non-stationary EEG signals. The evaluation metrics used to assess the performance of the proposed model include Accuracy, Precision, Recall, F1-score, Sensitivity, and Specificity. These standard performance metrics are widely used in classification studies.

$$\hat{y} = \frac{x - \mu B}{\sqrt{\sigma_B^2 + \epsilon}} \quad z = \gamma\hat{y} + \beta \tag{10}$$

where,

-σ_B^2,μB variance and mean of the batch.

-γ,β learnable scale and shift parameters.

LSTM Equations:

Forget Gate

$$ft = \sigma(Wf \cdot [ht - 1, xt] + bt) \tag{11}$$

InputGate:

$$it = \sigma(Wi \cdot [ht - 1, xt] + bt) \tag{12}$$

Candidate Cell State:

$$\tilde{ct} = \tanh(Wc.[ht - 1, xt] + bc) \tag{13}$$

Updated Cell State:

$$ct = ft \odot ct - 1 + it \odot \tilde{ct} \tag{14}$$

Output Gate:

$$ot = \sigma(Wo \cdot [ht - 1, xt] + bo) \tag{15}$$

Hidden State (LSTM Output):

$$ht = ot \odot \tanh(ct) \tag{16}$$

Final Classification:

$$y = \text{SoftMax}(Wy \cdot hT + bY) \tag{17}$$

Or for binary output:

$$y = \sigma(Wy \cdot hT + bY) \tag{18}$$

3.6 Model training and evaluation

3.6.1 Model training

The proposed CNN–LSTM model was trained using the Adam optimizer due to its adaptive learning capability and fast convergence. A learning rate of 0.01 was employed for both CNN and LSTM components. The CNN network was trained for 5000 epochs to ensure stable spatial feature learning, while the LSTM network was trained for 2000 epochs to effectively capture temporal dependencies in EEG signals. In order to avoid overfitting and ensure that the model generalizes well, a five-fold cross-validation approach was used, where the data was split into five subsets, and the training and validation were carried out in a round-robin fashion among all subsets.

3.6.2 Performance evaluation metrics

To comprehensively evaluate the performance of the proposed EEG-based stress detection model, multiple evaluation metrics were considered. These metrics provide a balanced assessment, especially in the presence of class imbalance commonly observed in EEG datasets.

Accuracy:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{19}$$

Measures the proportion of correctly classified samples. However, in imbalanced EEG datasets, accuracy alone may not fully reflect performance.

Precision:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{20}$$

Represents the quantity of correctly predicted stressed samples among all predicted stressed cases, indicating reliability of positive predictions.

Recall:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{21}$$

Quantifies the quantity of correctly detected stressed samples, emphasizing the ability to capture true stress cases.

F1-Score:

$$F - 1\text{score} = 2x\frac{\text{Precision x Recall}}{\text{Precision} + \text{Recall}} \tag{22}$$

Provides a harmonic mean between precision and recall, useful for imbalanced datasets where one metric alone may be misleading.

Selectivity (True Negative Rate):

$$\text{Selectivity} = \frac{TN}{TN + FP} \tag{23}$$

Represents the classifier's ability to correctly identify non-stressed samples.

3.7 Workflow of the proposed electroencephalogram stress detection framework

The proposed framework for stress detection using EEG starts with the acquisition of EEG signals from the DEAP dataset, which is recorded with 40 electrodes organized following the international 10-20 system and a sampling rate of 128 Hz. The raw EEG signals are then preprocessed with band-pass filtering, artifact removal, and normalization. After that, statistical and spectral features are extracted from the EEG signals.

In order to minimize redundancy in features and enhance discriminative power, a hybrid feature selection method incorporating AOA, AHP, and LF is used. This approach helps to relate the most informative subset of features with a well-rounded trade-off between multiple features. The selected features are then reformatted and fed into a hybrid CNN-LSTM model. In this model, the CNN part of the architecture identifies spatial dependencies and hierarchical patterns of features among EEG channels, and the LSTM part identifies temporal dependencies among EEG segments.

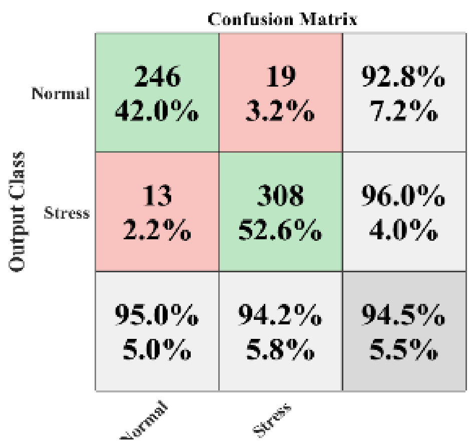
The CNN-LSTM network is run with the Adam optimizer, and the generalization performance is evaluated using five-fold cross-validation. Finally, the performance is measured using a variety of metrics such as accuracy, precision, recall, F1-score, selectivity, AUC-ROC, and standard deviation, which will provide a complete analysis of the robustness and reliability of the proposed EEG stress detection system.

4. RESULTS AND DISCUSSION

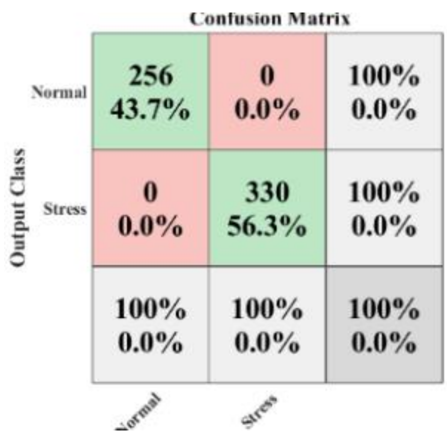
The DEAP dataset, which contains 244 EEG recordings from 40 channels, was used to validate the designed AOA-AHP-CNN-LSTM Framework (the analysis-of-algorithms (AOA) method, the AHP method, and a convolutional-neural-network (CNN)-based long short-term memory (LSTM) model) and the model's performance was evaluated based on class prediction performance (i.e., classification accuracy, precision, recall, F-Score, and AUC/ROC) using 5-fold cross-validation (CV) to eliminate any potential bias in the results. Table 1 summarizes the classification performance results of all 5 CV folds using the different evaluation measures listed above. The Average Classification Accuracy of the proposed model was 96.20% with a standard deviation of 2.67%, depicting the consistency of the model in this task. The max accuracy of the model is found in one of the folds at 100%, whereas the lowest accuracy of the model is found in another fold at 92.03%. This could be the effect of individual subject properties of the EEG signals. The individual consistency of the model in terms of discriminant capacity is shown by the AUC/ROC values ranging between 0.95 to 1.00.

Table 1. Performance metrics across 5-fold cross-validation

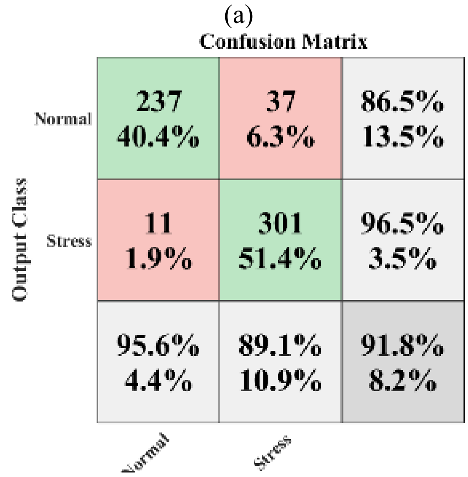
Fold	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC
1	100	100	100	100	1.00
2	94.66	94.19	95.95	95.06	0.96
3	92.03	89.05	96.47	92.61	0.95
4	98.50	98.61	98.04	98.32	0.98
5	97.80	96.95	98.20	97.57	0.97
Mean ± Std	96.20 ± 2.67	95.36 ± 4.15	97.33 ± 1.48	96.71 ± 2.74	0.97 ± 0.02



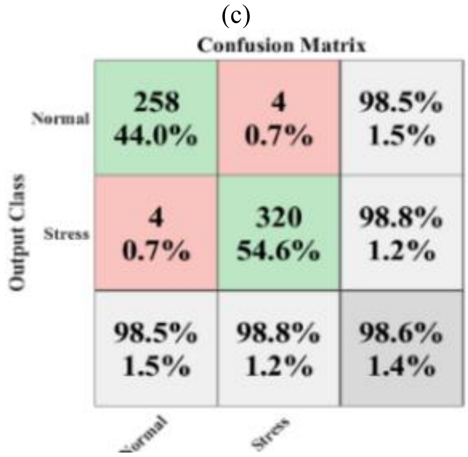
(a)



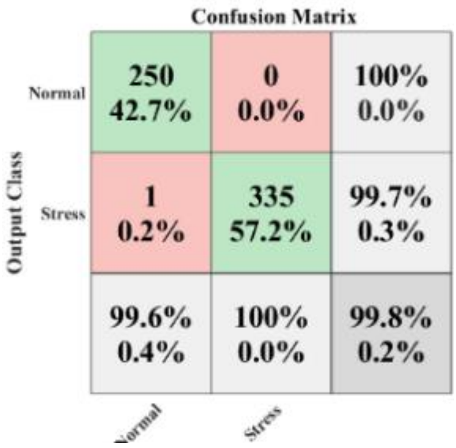
(c)



(b)



(d)



(e)

Figure 2. Confusion matrices for each fold in 5-fold cross-validation using the proposed model

In this regard, the confusion matrices of each fold depicted in Figures 2(a)–(e) show that the proposed framework has given an optimum discrimination between stressed and non-stressed states. From a few folds, only a limited number of misclassifications were found, which may be due to inter-subject variations in EEG patterns. In summary, the confusion matrices presented for both classes give validation to the cross-validation scheme for the proposed AOA–AHP–CNN–LSTM model performance.

The accuracies of the training and validation sets of the developed model are shown in Figure 3. Although it appears that there's a lack of convergence in the accuracy of the model at 100% for the training set, it should be noted that it's not an unexpected phenomenon in an EEG classification model because of the noisy, non-stationary, and subject-dependent characteristics of EEG signals. It should be noted that in such scenarios, converging to 100% could result in overfitting of the model and its poor validation performance. The fact the gap between the two accuracies remains relatively small implies strong convergence of the model.

The efficiency of the proposed AOA-AHP-CNN-LSTM approach was also assessed by comparison with other well-established optimization algorithms, such as PSO-CNN-LSTM and GA-CNN-LSTM, and other state-of-the-art DL models, such as CNN-BiGRU and CNN-LSTM without feature optimization. It should be noted that in order to have a clear comparison, in this comparison, all techniques have been tested in the same experimental setting and dataset. Based on Table 2, it has been observed that in comparison with other techniques, the proposed approach has showcased better results in terms of mean accuracy of 96.20%, F1-score value of 96.71%, and AUC-ROC value of 0.97, respectively.

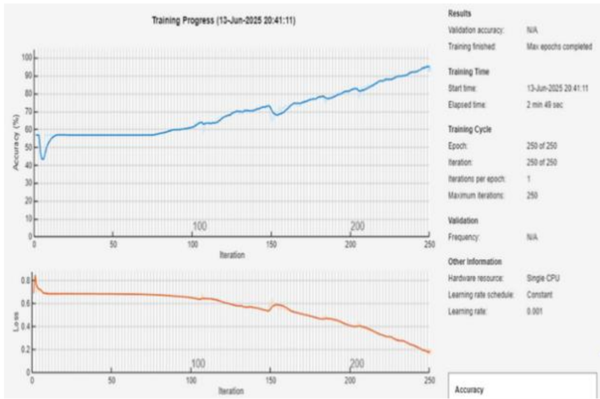


Figure 3. Training curves

For a fair comparison, the GA- and PSO-based CNN–LSTM baselines were implemented using the same network architecture, training protocol, and evaluation strategy as the proposed method. The hyperparameters of the CNN–LSTM model, including learning rate, batch size, number of epochs, and optimizer, were kept identical across all methods. GA and PSO were employed exclusively for feature selection, where standard population-based hyperparameters were adopted based on commonly used settings reported in the literature. Both optimizers were executed for the same number of iterations and population size to ensure consistency. This controlled experimental setup ensures that the observed performance differences are primarily attributable to the effectiveness of the feature optimization strategy rather than disparities in model configuration or training conditions.

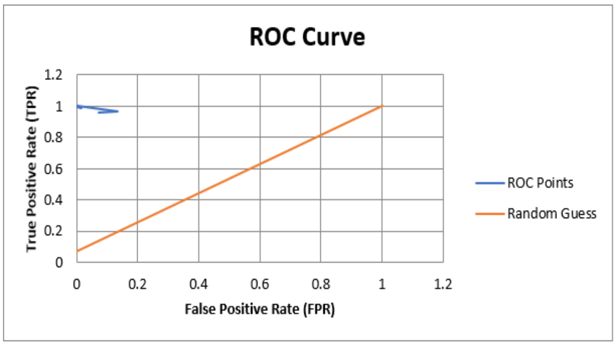
Table 2. Comparison of optimization methods for CNN-LSTM model

Model	Accuracy (%)	F1-Score (%)	AUC-ROC
GA-CNN-LSTM	93.75	94.30	0.94
PSO-CNN-LSTM	94.10	94.80	0.95
CNN-BiGRU	94.00	92.70	0.95
CNN-LSTM (no optimization)	91.70	90.90	0.92
AOA-AHP-CNN-LSTM (Proposed)	96.20	97.45	0.97

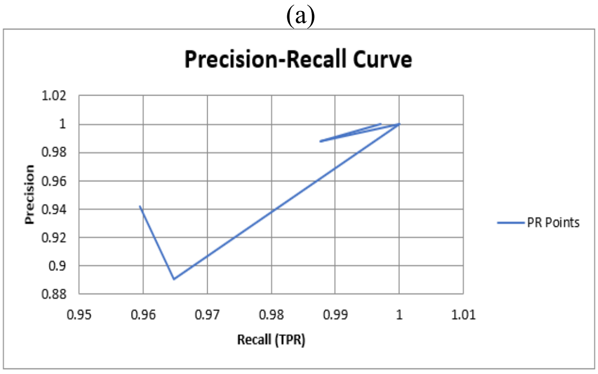
Note: CNN-LSTM = Convolutional Neural Network–Long Short-Term Memory.

The reason for the improvement in performance can be assign to the following two aspects: (i) the dual-phase AOA-AHP optimizer, which aims at selecting highly discriminatory EEG features and reducing redundancy among them, and (ii) the CNN-LSTM hybrid model, which focuses on spatial and temporal correlations of EEG signal characteristics. Also, the introduction of LF in AOA improves the ability of the optimizer for better exploration of the global space and avoids local minima. The proposed framework is developed in MATLAB and executed on an average machine configuration with an i5 processor and 16 GB RAM. The average time required for making an inference in the proposed framework is less than 1 second per sample. The memory usage is kept moderate in the proposed framework because the feature dimensional space is kept smaller by the proposed framework. The ROC curve and the Precision-Recall curve of the proposed framework are depicted in Figures 4(a) and (b), respectively. The ROC curve remains on top of the diagonal, obtaining an average AUC of 0.96, whereas the Precision Recall curve reaches an average Precision of 0.95. These findings highlight the robustness of the model under class imbalance conditions and confirm its effectiveness in real-world stress recognition scenarios.

Nevertheless, the proposed approach also has some limitations. The choice of the DEAP dataset may not be comprehensive in stress situations, as it might vary in realistic scenarios. Moreover, the approach has limitations regarding the utilization of EEG signals only, which may not leverage the advantages that could be obtained from multiple physiological signals, for example, ECG and/or GSR. Future research should therefore aim to validate the approach using bigger and more comprehensive data, in addition to focusing on multimodal fusion. Moreover, the impact of more advanced models, for example, the attention model and the transformer model, should also be explored to improve the extraction of spatial and temporal features. However, the experimental findings support that the combination of the metaheuristic approach for feature selection with a deep model is a computationally efficient robust approach for stress detection using EEG signals. For the purposes of analyzing the behavior of the proposed feature selection strategy in the context of an optimization algorithm, the convergence properties of AOA-AHP are compared with those of AOA-AHP-LF in Figure 5 with regards to the z-score-normalized inverse fitness values of the individuals with 1/fitness over 100 iterations.



(a)



(b)

Figure 4. ROC, PR curve

Note: ROC = Receiver Operating Characteristic;PR = Precision Recall.

It is clear that the AOA-AHP-LF model converges faster with a lower final fitness value compared to the AOA-AHP model, thus signifying better optimization capabilities. The addition of the LF further improves the early exploration capabilities for the global solution, effectively preventing it from converging prematurely. On the contrary, the AOA-AHP model, without the LF, converges slowly, getting stuck in a local solution during later iterations. In addition, the stepwise reduction in the fitness curve of AOA-AHP-LF indicates the improvement in the diversity of the population and the exploration-exploitation tradeoff. This makes it possible for the optimization algorithm to search for the solution space more efficiently and reach the optimal solution for feature selection with better performance. However, the convergence analysis proves that the addition of LF improves the stability, efficiency, and effectiveness of the feature selection process using the AOA-AHP approach.

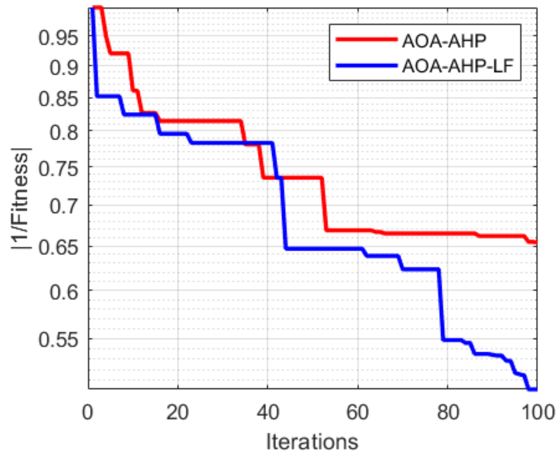


Figure 5. Comparative convergence behavior of AOA–AHP and AOA–AHP–LF over 100 iterations

Note: AOA–AHP = rchimedes Optimization Algorithm , Analytic Hierarchy Process; AOA–AHP–LF = Archimedes Optimization Algorithm , Analytic Hierarchy Process , and Lévy Flight.

5. CONCLUSION AND FUTURE SCOPE

The proposed framework proves the efficacy of the combination of metaheuristic feature optimization and deep spatiotemporal learning in the context of EEG-based stress detection, outperforming the state-of-the-art methods in terms of performance. The proposed approach leveraged the spatial and temporal properties of EEG signals by achieving an average classification accuracy of 96.20% on the DEAP dataset using five-fold cross-validation with validation metrics including F1-Score, AUC-ROC, and standard deviation. AOA + AHP feature optimization combined with the CNN-LSTM model is a scientifically designed and effective way of using EEG signals for stress identification. Comparison of the proposed model against the baseline models designed using PSO and GA has further proved the efficacy of the proposed model. Nevertheless, despite the success achieved, some drawbacks should also be noted. Firstly, the testing has been carried out only using one public data set with a small number of samples, and the proposed framework has been based only on EEG signals, with no other physiological signals, such as ECG or GSR, being considered. Secondly, this proposed framework has been tested only as an offline process and has not been tested on real-time data. Looking ahead, real-time implementation of the above-proposed framework using wearables/modular platforms and testing it using larger datasets will be an area of focus. Additionally, exploration of multimodal signals along with advancements like optimization techniques/deep learning networks might add to accuracy as well as practical implementation readiness of emotional intelligence frameworks. EEG signals are impacted by motion, and noise, among other things; therefore, these various sources of noise will affect the degree of robustness achievable by wearable and mobile devices. The presence of inter-subject variability along with calibration-related issues may be further impeding on a unified approach for all individuals. Finally, the ability for real-time applications will require optimization of power and hardware aspects of mobile platforms to achieve low latency results. Adaptive calibration methods, robustly processing against artifacts, and model optimization for low power on mobile devices will be critical for resolving remaining challenges.

DATA AVAILABILITY

The data used in this research is the publicly available DEAP database, and it can be accessed from the link: http://www.eecs.qmul.ac.uk/mmv/datasets/deap/. The DEAP database contains multimodal physiological signals such as EEG signals recorded during emotional responses.

CODE AVAILABILITY

The source code implementing the proposed AOA-AHP-based feature selection and CNN-LSTM classifier is available from the corresponding author upon reasonable request. The code includes MATLAB scripts.

AUTHOR CONTRIBUTIONS

Pritika Patil conducted the data preprocessing, model implementation, experimental design, and manuscript writing. Shirish Kulkarni done Supervision, validation, critical review, editing of the manuscript and approved the final version of the manuscript.

REFERENCES

[1] Al-Shargie, F., Tang, T.B., Kiguchi, M. (2017). Stress assessment based on decision fusion of EEG and fNIRS signals. IEEE Access, 5: 19889-19896. https://doi.org/10.1109/ACCESS.2017.2754325

[2] Khalil, R.A., Jones, E., Babar, M.I., Jan, T., Zafar, M.H., Alhussain, T. (2019). Speech emotion recognition using deep learning techniques: A review. IEEE Access, 7: 117327-117345. https://doi.org/10.1155/2014/627892

[3] Craik, A., He, Y., Contreras-Vidal, J.L. (2019). Deep learning for electroencephalogram (EEG) classification tasks: A review. Journal of Neural Engineering, 16(3): 031001. https://doi.org/10.1088/1741-2552/ab0ab5

[4] Aguiñaga, A.R., Delgado, L.M., López-López, V.R., Téllez, A.C. (2022). EEG-based emotion recognition using deep learning and M3GP. Applied Sciences, 12(5): 2527. https://doi.org/10.3390/app12052527

[5] Yanchun Zhang, A. (2016). EEG signal analysis and classification techniques and applications.

[6] Koelstra, S., Muhl, C., Soleymani, M., Lee, J.S., et al. (2011). DEAP: A database for emotion analysis; using physiological signals. IEEE Transactions on Affective Computing, 3(1): 18-31. https://doi.org/10.1109/T-AFFC.2011.15

[7] Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T.H., Faubert, J. (2019). Deep learning-based electroencephalography analysis: A systematic review. Journal of Neural Engineering, 16(5): 051001. https://doi.org/10.1088/1741-2552/ab260c

[8] Bashivan, P., Rish, I., Yeasin, M., Codella, N. (2015). Learning representations from EEG with deep recurrent-convolutional neural networks. arXiv preprint arXiv:1511.06448. https://doi.org/10.48550/arXiv.1511.06448

[9] Jenke, R., Peer, A., Buss, M. (2014). Feature extraction and selection for emotion recognition from EEG. IEEE Transactions on Affective Computing, 5(3): 327-339. https://doi.org/10.1109/TAFFC.2014.2339834

[10] Kartali, A., Janković, M.M., Gligorijević, I., Mijović, P., Mijović, B., Leva, M.C. (2019). Real-time mental workload estimation using EEG. In International Symposium on Human Mental Workload: Models and Applications, pp. 20-34. https://doi.org/10.1007/978-3-030-32423-0_2

[11] Chakravarthi, B., Ng, S.C., Ezilarasan, M.R., Leung, M.F. (2022). EEG-based emotion recognition using hybrid CNN and LSTM classification. Frontiers in Computational Neuroscience, 16: 1019776. https://doi.org/10.3389/fncom.2022.1019776

[12] Coan, J.A., Allen, J.J. (2004). Frontal EEG asymmetry as a moderator and mediator of emotion. Biological Psychology, 67(1-2): 7-50. https://doi.org/10.1016/j.biopsycho.2004.03.002

[13] Li, X., Song, D., Zhang, P., Zhang, Y., Hou, Y., Hu, B. (2018). Exploring EEG features in cross-subject emotion recognition. Frontiers in Neuroscience, 12: 162. https://doi.org/10.3389/fnins.2018.00162

[14] Srinivasan, R., Winter, W.R., Ding, J., Nunez, P.L. (2007). EEG and MEG coherence: Measures of functional connectivity at distinct spatial scales of neocortical dynamics. Journal of Neuroscience Methods, 166(1): 41-52. https://doi.org/10.1016/j.jneumeth.2007.06.026

[15] Delorme, A., Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods, 134(1): 9-21. https://doi.org/10.1016/j.jneumeth.2003.10.009

[16] Lal, T.N., Schroder, M., Hinterberger, T., Weston, J., Bogdan, M., Birbaumer, N., Scholkopf, B. (2004). Support vector channel selection in BCI. IEEE Transactions on Biomedical Engineering, 51(6): 1003-1010. https://doi.org/10.1109/TBME.2004.827827

[17] Hashim, F.A., Hussain, K., Houssein, E.H., Mabrouk, M.S., Al-Atabany, W. (2021). Archimedes optimization algorithm: A new metaheuristic algorithm for solving optimization problems. Applied Intelligence, 51(3): 1531-1551. https://doi.org/10.1007/s10489-020-01893-z

[18] Yang, X.S., Deb, S. (2009). Cuckoo search via Lévy flights. In 2009 World Congress on Nature & Biologically Inspired Computing (NaBIC), Coimbatore, India, pp. 210-214. https://doi.org/10.1109/NABIC.2009.5393690

[19] Saaty, T.L. (2008). Decision making with the analytic hierarchy process. International Journal of Services Sciences, 1(1): 83-98. https://doi.org/10.1504/IJSSCI.2008.017590

[20] HH, J. (1958). The ten-twenty electrode system of the international federation. Electroenceph Clin Neurophysiol, 10: 367-380.