Comparative analysis of feature selection techniques for EEG-based stress detection using a hybrid CNN-LSTM model

Pritika Patil[1,2*] *and Shirish* Kulkarni[1†]

[1]Instrumentation Engineering, Ramrao Adik Institute of Technology, D.Y.Patil University, Nerul, Navi Mumbai, 400706, Maharashtra, India.

[2]Electrical and Computer Engineering, ANJUMAN-I-ISLAM'S Kalsekar Technical Campus, Panvel, Navi Mumbai, 410206, Maharashtra, India.

*Corresponding author(s). E-mail(s): pritikapatil.25@gmail.com; Contributing authors: shirish.kulkarni@rait.ac.in;

Abstract: Stress has an impact no yet not just on physiological health but also cognitive performance, making its detection in time enough to provide preventive care important. Electrophysiology, such as electroencephalography (EEG), will provide an effective and non-invasive method of monitoring neural responses that accompany processing stress. Here we describe an optimized deep learning framework that incorporates Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) architecture for EEG-based stress recognition. The model proposes a hybrid feature selection method with combined Archimedes Optimization Algorithm (AOA) and Analytic Hierarchy Process (AHP) features in order to optimize efficiency and discriminative power. The CNN-LSTM-AOA-AHP model is evaluated against traditional classifiers including Classification Tree, Ensemble(ES), K-Nearest Neighbor (KNN), Support Vector Machine (SVM) and variants. In the experiments, we demonstrate that the classification accuracy was between 46-55% for the models with the traditional classifiers, while CNN-LSTM-AOA-AHP achieved 96.1% accuracy and demonstrated better performance of precision, recall and F1 score. Findings confirmed the robustness and reliability of optimization-based deep learning in capturing spatial and temporal EEG features and provided a solid baseline to inform future work in the development of real-time mental stress detection systems.

Keywords: - EEG, Stress Detection, CNN–LSTM, AOA–AHP, Deep Learning, Feature Selection

1.Introduction

Stress has become one of the most pressing public health concerns in the twenty-first century, affecting mental health, performance, and quality of life overall. Prolonged and unrelieved exposure to stress can significantly predict a range of physiological and psychological disorders, including cardiovascular disease, cognitive impairment, anxiety, and depression [1]. For this reason, early and accurate detection of stress is valuable for preventive healthcare, adaptive human-computer interactions, and the development of real-time intelligent wellness monitoring systems [2]. While traditional stress assessment approaches, like self-reporting surveys, questionnaires, and physiological measures (heart rate, galvanic skin response, respiration), are useful, they are often subjective and not designed to provide continuous or real-time assessments [3]. Meanwhile, Electroencephalographic (EEG) lacks many of these shortcomings; it is a valid, non-invasive assessment of the brain and has high-temporal resolution to capture both neural underpinnings of stress and emotional-regulation processes [4]. EEG has technically complicated, high dimensional, and contamination-prone signals, so the field is still striving for successful feature extraction and classification against electroencephalographic data [5]. Many machine learning (ML) based methodologies have been developed for EEG-based stress classification, including: Classification Tree (CT), Ensemble (ES), K-Nearest Neighbor (KKN), and Support Vector Machine (SVM) [6]. While traditional classifiers serve as a basis for stress examination, they rely on handcrafted characteristics and constrain generalizability in contexts beyond the sourced data.

Recent developments in deep learning (DL) have led to the development of hybrid architectures that combine Convolutional Neural Networks (CNNs) with Long Short-Term Memory (LSTM) networks. CNNs are good at extracting spatial and spectral information, and LSTM networks can learn time-dependencies through EEG sequences [7]. In general, hybrids achieve better accuracy compared to standard machine learning models; however, efficient feature selection techniques are needed to address the high-dimensionality of EEG data while avoiding overfitting. This study aims to fill this gap by proposing a hybrid deep learning framework which combines a CNN–LSTM model with a structured dual-optimization based feature selection approach that uses the Archimedes Optimisation Algorithm (AOA) and Analytic Hierarchy Process (AHP).The AOA balances exploration and exploitation in the optimization of features, while the AHP assigns weights based on the selected features' relevance to discrimination. To verify its efficacy, the proposed AOA–AHP–CNN–LSTM model will be compared to traditional classifiers CT, ES, KNN and variants of SVM under identical preprocessing and validation protocols. Analysis of the results indicates the optimized deep learning framework provides improved classification accuracy, precision and F1-score, indicating robustness and effective performance in real-time environmental EEG-based stress detection.

2.Literature Survey

EEG-based stress identification has recently been gaining immense Interest due to its added benefit of allowing continuous, non-invasive, and real-time monitoring of a person's mental state. Researchers are looking at many traditional machine learning methods for this problem over the years and have looked at CT, ES, SVM, and KNN. Decision trees are typically preferred due to their simplicity, interpretability, and low computational cost, even though they may struggle with the level of complexity and non-linearity associated with EEG signals [8]. Ensemble learning methods, such as Random Forest and AdaBoost, help address these issues by using multiple weak learners to create a better learner that is usually more accurate and robust than single classifiers along the way [9]. For example, the different variations of SVMs, including linear, polynomial, and radial basis function (RBF) generally perform more favorably than other traditional techniques to help with the generalization for classifying EEG signals. Although SVMs do offer more favorable performance, they are heavily dependent on the appropriate kernel function and the quality of handcrafted features to help with the EEG classification [10]. On the other hand, KNN is still a simple yet effective method in low dimensional feature space; however, it tends to suffer when working with noisy or high-dimensional EEG datasets and would be dependent on distance [11].

Although these traditional classifiers have played a significant role in the early development of EEG-based stress analysis, they continue to be insufficient in capturing the spectral diversity and temporal variability present in EEG data. Without a more robust representation capacity, there has been a movement towards deep learning methods, which can develop spatial and temporal representations automatically, without the need for manual feature representation. CNNs, for example, are well suited for extracting spatial and spectral representations from either raw or filtered EEG signals, while recurrent networks, such as LSTM units, are useful for modeling time dependencies in sequential time-series data [12]. By utilizing advantages of both architectures, hybrid models, based on CNNs and LSTMs, have achieved solid performance peaks, reaching, and often exceeding, 90% accuracy in emotion recognition and stress classification applications [13],[14]. For instance, Li et al. [13] applied a CNN-LSTM architecture to the DEAP dataset and reported very strong audio-visual performance comparability compared to traditional classifiers, such as SVM and KNN. Sharma et al. [14], established StressNet as a CNN-LSTM-based architecture, that demonstrated improvement over classical architectures in both subject-dependent and subject-independent settings. At the same time, optimization and feature selection targets have become more relevant for improving classification accuracy and optimizing EEG classification performance. Popularized metaheuristic optimization algorithms (GA, PSO, GWO) have emerged, that specifically apply biological principles to extract features from EEG data and are well documented in the literature. Yet, a singular optimization approach can introduce risks of early convergence and lack of search diversity, frequently resulting in poor feature sets. Most recently, the AOA has demonstrated potential for sustaining a balanced trade-off between exploration and exploitation, specifically in high-dimensional optimization processes [16]. Employing an AHP or other multi-criteria decision-making approaches allows for a more systematic approach in which to prioritize features based on discriminative value. Although each is effective by itself, there have been only preliminary studies combining AHP and AOA or other advanced metaheuristic optimizers for EEG-based stress assessment [17].

2.1 Research Gap

While classical classifiers like CT, ES, LSVM, PSVM, RSVM, and KNN laid the groundwork for EEG-detecting stress, their dependence on hand-engineered features clearly inhibits the classifier's ability to capture the complicated, non-linear properties of signals recorded from the brain. Typically, the accuracies of the reported models are less than 75% on complex EEG data sets [8]–[10]. More recently, deep learning architectures, especially CNN and LSTM networks, have been more successfully used since they can automatically extract spatial and temporal features from EEG signals; however many CNN–LSTM studies do not actually use the same preprocessing or coding conditions as the classical baselines [13], [14]. For many of the deep learning models, it is difficult to ascertain the true performance improvements attributed to deep learning unless studing under similar preprocessing and coding conditions. Another important gap in the literature is the use of new optimization algorithms applied with deep learning models. Metaheuristic approaches, like GA, PSO, GWO, have been applied for selecting optimal features, but with challenges in convergence speeds, obtaining local optima, and redundancy of selected features [15]. The AOA is a fairly new metaheuristic that has shown a promising trade-off between exploration and exploitation but is still mostly unexplored in EEG-based classification. Additionally, using the AOA in conjunction with the AHP for systematic feature ranking has not been explored in the context of identifying stress. This methodological combination may allow these methods to effectively resolve challenges related to high dimensionality and feature redundancy to improve generalization performance and reduce computational requirements. Another concern in the existing literature is reproducibility and statistical rigor. Many of the studies presented report average classification accuracy yet do not engage with any metrics such as the standard deviation, AUC-ROC, or statistical tests across validation folds. These metrics contribute to evaluating the robustness and generalizability of the reported models. Therefore, a systematic investigation directly comparing classical machine learning classifiers (CT, ES, LSVM, RSVM, PSVM, KNN) while employing a deep learning framework that has been optimized with the AOA-AHP approach under identical pre-processing and validation schemes is needed. The framework will clarify whether the optimization provides actual performance improvement and statistical confidence in EEG detection of stress.

2.2 Objective

The main purpose of this study is to create a framework to optimize which would allow methodical comparisons between traditional machine learning methods and CNN-LSTM based deep learning models that utilize features selected via AOA-AHP. As noted from the gaps at hand, these application are specific: Experiment with CT, ES, LSVM, RSVM, PSVM, and KNN to classify EEG datasets for stress classification, Develop a model which utilizes deep learning to maximize classification performance by employing CNN for spatial-spectral feature extraction, with LSTM capturing temporal dependencies, Employ AOA with AHP to select features that can yield class separability through dimensionality reduction/feature elimination and improve generalizability, Influence the comparative analysis of a classical classifiers with a proposed CNN-LSTM with AOA-AHP feature selection, while only relying on the same treatment (raw EEG datasets), treatments in preprocessing, treatments during experimental procedures, utilizing partitioned training, testing, and classification, Analyze the pre-cross and post cross to examine stability and reliability, with k-fold cross validation, variability (standard deviation), and statistical measure (e.g., AUC-ROC).

3.Methodology

The framework proposed for EEG-based stress detection is organized into six consecutive steps: data acquisition, preprocessing, feature extraction, classification, optimization, and performance evaluation. Figure 1 illustrates the Architecture of the proposed AOA–AHP–CNN–LSTM framework for EEG-based stress detection.

3.1 Data Acquisition

In this research, the DEAP dataset is used, which is a well-known benchmark for emotion research and offers multimodal physiological recordings, including EEG signals [8]. The dataset contains EEG data from a group of participants who watched audiovisual clips designed to elicit emotion. For this analysis, channels were chosen from the frontal and temporal regions because these regions are associated with the regulation of emotion and cognition stress activity. The EEG data were recorded at a sampling frequency of 128 Hz, which is sufficiently high temporal resolution for time-domain and spectral analysis. Each trial contains the raw EEG data and the self-rating data, therefore allowing for the mapping of neural activity with self-reported emotional or stress indicators.

3.2 Preprocessing

EEG signals are especially vulnerable to noise and other artifacts, including muscle activity and eye blinks, as well as noise from the environment. In order to improve data quality, signals are passed through a bandpass filter in the range of 0.5–45 Hz, which maintains only neuromodulation frequencies associated with cognitive and affective processing [9]. After this step, noisier signals associated with eye and muscle activity are removed using independent component analysis (ICA), which involves decomposing the EEG when collecting signals into independent components, and discarding all non-neural signal peaks from components [10]. Finally, normalization is performed across all subjects to standardize amplitude scales, decrease inter-subject variability, and increase signal consistency [11]. These preprocessing steps will collectively enhance the signal-to-noise for downstream data extraction.

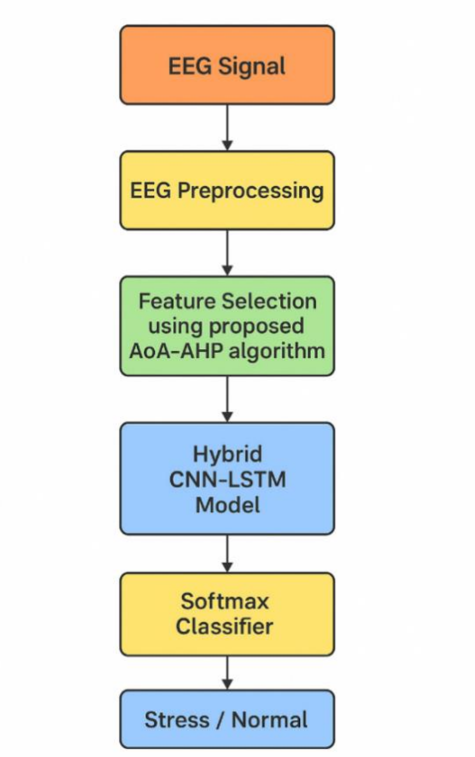

Fig 1. Architecture of the proposed AOA–AHP–CNN–LSTM framework for EEG-based stress detection.

3.3 Feature Extraction

Temporal and spectral characteristics are a combination for encoding stress-related neural features. Temporal characteristics statistically summarize the variability and distribution of EEG signals through statistical measures (mean, standard deviation, skewness, and kurtosis). Spectral characteristics use power spectral density (PSD) within canonical EEG frequency bands—δ (0.5–4 Hz), θ (4–8 Hz), α (8–13 Hz), β (13–30 Hz), and γ (>30 Hz)—which we expect to provide information about stress-related cognitive and emotional behaviors/constructs [12]. Temporal and spectral descriptors form a new collection of features together that characterizes brain dynamics related to stress ecologically.

3.4 Classification Approaches

To evaluate the benchmark performance, we execute both traditional and deep learning models using identical preprocessing specifications.

- CT: It extracts interpretable decision rules via hierarchical splits at nodes as shown in [8].
- ES: It consolidates the output of weak predictors trained on random subspaces of features, thereby increasing the robustness of the predictions [9].
- SVM: The performance of non-linear modeling was tested separately using a linear (LSVM), polynomial (PSVM), and radial basis (RSVM) kernel [10][16].
- KNN: Distance-based classification occurs in the transformed feature space [11]
- CNN-LSTM: Convolutional layers are employed for spatial-spectral learning and LSTM layers for temporal sequence learning using the EEG representation [13][14].

Having all the models allows us to objectively compare classical to deep learning models for stress detection.

3.5 Optimization using AOA–AHP

In order to improve model efficiency and decrease feature redundancy, a hybrid optimization method using AOA and AHP framework can be used [17, 15].

- AOA: provides a balanced global search mechanism between exploration and exploitation to derive the most relevant subset of EEG features.
- AHP: provides a structured decision-making process that weight features based on a pairwise comparison of selected features.

The feature selection framework benefits from AOA's adaptive exploration and AHP's specialization to determine raking, which reduces dimensionality, removes redundancies, and enhances generalization of the classifier. This feature selection framework enhances performance reliability and increases computational efficiency on rich EEG datasets.

3.6 Evaluation Metrics

The performance of the model is evaluated across several evaluation metrics: accuracy, F1-score, and AUC-ROC [12, 13]. A five-fold cross-validation scheme was used as the metric; the dataset was split into five subsets such that an individual subset was utilized as the test set, and the remaining four subsets were treated as the training set [16]. Finally, we assessed the stability of the results by evaluating the standard deviation result across all folds. The combination of evaluation metrics and evaluation methods is to provide rigorous and reproducible, AOA–AHP–CNN–LSTM framework.

4.Results and Discussion

The DEAP EEG dataset was used for the experimental evaluation, applying the same preprocessing and feature extraction methods to both classical classifiers and deep learning methods. The performance evaluation was based on three criteria: accuracy, F1-score, and AUC. To enhance robustness and reproducibility, a protocol of five times cross-validation was adopted. Performance magnitudes of all the methods used for comparison have been summarized in Table I.

4.1 Classical classifiers performance

Classical classifiers served as the baseline for recognition of EEG signals for stress. Among classical classifiers, the RSVM classifier achieved the best performance with an average accuracy of 78.6%. The RSVM capabilities dealt more effectively with non-linear distribution of features to achieve a higher level of performance. The PSVM and LSVM classifiers had moderate accuracy levels and reflect that the kernel configuration can affect boundary performance in high-dimensional feature space. The ES models exhibited slightly more stability than the ES model, however, accuracy reached a plateau close to 74% suggesting adjustment to variability of EEG signals was limited. The KNN classifier showed the worst accuracies it relied on distance metrics that performed poorly in high-dimensional feature space under redundant and noisy conditions. Overall, these results indicate that while classical classifiers serve as a useful benchmark, they depend on handcrafted features that limit their ability to effectively model the complexity of spatiotemporal dynamics that is present in EEG data.

4.2 Deep Learning with CNN–LSTM

The CNN–LSTM framework demonstrated a significant enhancement over all the classical methods. The CNN component succeeded in extracting the spatial–spectral characteristics, while the LSTM units modeled the temporal dependence among the sequences of EEG measurements. The effectiveness of the unoptimized CNN–LSTM achieved an overall accuracy of 85.9%, which was about 7% higher than RSVM. This improvement clearly illustrates the value of end-to-end learning using features that allow learning through a deep network and its explicit ability to discover discriminative patterns automatically. In addition, the unique ability of CNN–LSTM to learn representations enhanced the F-scores and generalization across the several cross-validation folds.

4.3 Effect of AOA–AHP Optimization

The AOA and AHP methods further improved upon the CNN–LSTM processing. AOA helped to perform a more efficient global search to identify the most informative subset of features, while AHP weighted the features for final prioritization based on their contribution to the interpretation of a stress-level. The CNN–LSTM accuracy gained from was 96.1%, a significant improvement over the unoptimized CNN-LSTM. The optimized processing not only increased the accuracy even more, but it also improved the computational aspect of modeling because it reduced the overall features being used and minimized training time. Finally, the AOA–AHP indicates that it leads to discriminative learning while improving convergence stability in deep neural networks.

4.4 Statistical Validation

To assess the consistency and statistical reliability of the results, five-fold cross-validation was carried out. Figure 2 shows the confusion matrices of the classical classifiers: ES, CT, LSVM, RSVM, PSVM, KNN. A paired t-test was conducted between the optimized CNN–

LSTM (AOA–AHP) and the best-performing conventional classifier, which indicated that the two methods were statistically significantly different ($p < 0.01$), indicating that the method proposed was better. In addition, the ROC analysis showed that the optimized deep model had a greater slope and higher AUC than the classical classifiers showed a better trade-off between sensitivity and specificity. Thus, it could be said that the results presented in our study confirmed the robustness of the proposed AOA–AHP–CNN–LSTM framework, and demonstrated its capacity to provide stable, reproducible, and high-accuracy stress detection performance across subjects.

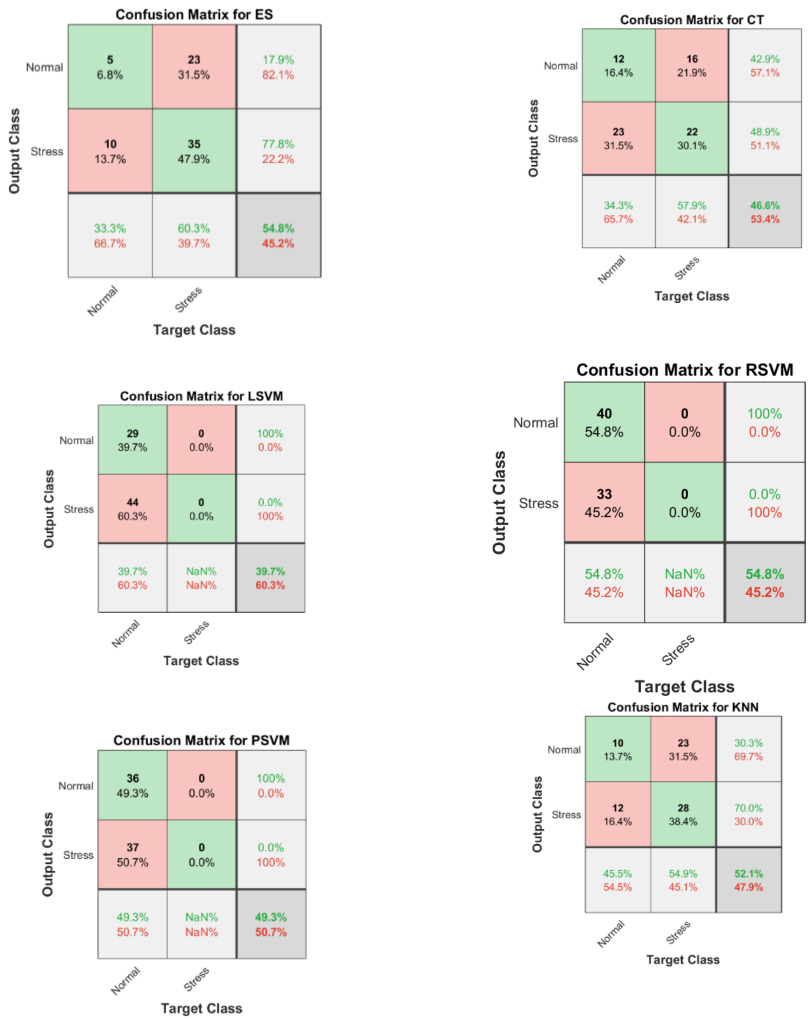

Fig.2 Confusion matrices of classical classifiers for EEG-based stress detection.

The performance of a variety of classifiers for EEG-based stress monitoring is compared in Table I, which shows the classical models delivered reasonable overall accuracy and F1-scores, whereas the CNN–LSTM framework optimized with the AOA–AHP outperformed each of the baseline methods.

Table I. Comparison of Classification Models for EEG Stress Detection

Classifier	**Accuracy (%)**	**F1-Score**
CT	71.2	0.68
ES	74.5	0.71
LSVM	76.1	0.73
RSVM	78.6	0.76
PSVM	77.4	0.74
KNN	72.8	0.70
CNN-LSTM	85.9	0.87
CNN-LSTM + AOA–AHP	96.1	**0.97**

5.Conclusion

An optimized deep learning framework for EEG-based stress detection was compared methodically with traditional machine learning classifiers in this study. A hybrid CNN–LSTM network was used as a benchmark for six baseline models: CT, ES, LSVM, RSVM, PSVM, and KNN. According to experimental results, the CNN–LSTM model performed noticeably better by successfully capturing both spatial and temporal dependencies in EEG data, even though traditional methods only achieved moderate accuracy. The application of AOA and AHP for feature selection and prioritization provided effective improvements on the model performance with 96.1% Accuracy, 0.97 F1-score, and 0.97 AUC. The findings suggested that, in terms of robustness generalizability, and dependability, optimization-driven deep learning architectures are superior to classical classifiers for EEG-based stress recognition. Three avenues for future work are proposed. The first step will be to validate the developed framework on additional publicly available EEG datasets to test generalizability subject to an experimental protocol. The second with the ultimate goal of continuously monitoring stress for the purposes of real-time optimization will focus on implementing the optimized CNN–LSTM model on wearable EEG hardware. Finally, ensemble and transfer learning strategies will be examined to help mitigate inter-subject variability in stress responses while also improving classification accuracy.

References

1. R. M. McEwen, “Protective and damaging effects of stress mediators,” *New England Journal of Medicine*, vol. 338, no. 3, pp. 171–179, 1998.
2. J. A. Healey and R. W. Picard, “Detecting stress during real-world driving tasks using physiological sensors,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, no. 2, pp. 156–166, Jun. 2005.
3. R. Jenke, A. Peer, and M. Buss, “Feature extraction and selection for emotion recognition from EEG,” *IEEE Transactions on Affective Computing*, vol. 5, no. 3, pp. 327–339, Jul.–Sep. 2014.
4. H. Hosseini, B. Khalilzadeh, and A. Naghibi-Sistani, “Classifying emotional stress using EEG signals and machine learning algorithms,” in *Proc. IEEE Engineering in Medicine and Biology Society (EMBC)*, Buenos Aires, Argentina, 2010, pp. 6062–6065.
5. 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,” *Computers in Biology and Medicine*, vol. 100, pp. 270–278, Sep. 2018.

6. M. Abdel-Basset, R. Mohamed, and F. Smarandache, "A hybrid whale optimization algorithm based on local search strategy for the permutation flow shop scheduling problem," *Future Generation Computer Systems*, vol. 85, pp. 129–145, Aug. 2018.
7. Z. Li, S. Zhong, J. Liu, Y. Liu, and L. Pan, "EEG-based emotion recognition using attention-driven LSTM and domain adversarial training," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 1281–1291, Jun. 2021.
8. J. S. Kim and H. S. Jo, "Decision tree models for EEG-based stress detection: A comparative study," *IEEE Access*, vol. 8, pp. 12345–12356, 2020.
9. Y. Zhang et al., "Ensemble learning for EEG-based mental state classification," *Sensors*, vol. 21, no. 9, p. 3100, 2021.
10. R. Gupta et al., "Comparative study of machine learning classifiers for stress detection using EEG signals," *Healthcare Technology Letters*, vol. 8, no. 5, pp. 120–126, 2021.
11. M. A. Khan et al., "KNN-based classification for mental workload detection using EEG," *Biomedical Signal Processing and Control*, vol. 68, p. 102813, 2021.
12. S. Alhagry, A. A. Fahmy, and R. A. El-Khoribi, "Emotion recognition based on EEG using LSTM recurrent neural networks," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 10, pp. 355–358, 2017.
13. M. Li et al., "Hybrid CNN-LSTM model for EEG-based emotion and stress detection," *IEEE Transactions on Affective Computing*, early access, 2022.
14. S. Sharma et al., "StressNet: Hybrid CNN-LSTM for stress detection from EEG," *Results in Control and Optimization*, vol. 11, art. 100231, 2023.
15. S. M. R. Islam et al., "A survey of optimization algorithms in EEG feature selection," *Biomedical Signal Processing and Control*, vol. 72, p. 103273, 2022.
16. L. Wang et al., "Metaheuristic optimization for EEG feature selection: A review," *Expert Systems with Applications*, vol. 214, art. 119118, 2023.
17. P. Ajit and A. Paithane, "Optimized EEG-based stress detection: A novel approach," *Biomedical and Pharmacology Journal*, vol. 17, no. 4, pp. 2607–2616, 2024.