IJERST

International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 21 No. 3 (1) 2025

ijerst.editor@gmail.com

editor@ijerst.com

ISSN 2319-5991 www.ijerst.com
Vol. 21, No. 3(1), 2025

Research Paper

AN AI-BASED SCREENING SYSTEM FOR ADHD IN REAL-WORLD CLINICAL

S. Krishna Reddy[1], A. Udayeni Reddy[2], S. Ravi Teja[2], Chinthakuntla Srivani[2], S. Dinesh Reddy[2]

[1] Assistant Professor, [2]UG Student, [1,2] Department of Computer Science and Engineering

[1,2] Sree Dattha Institute of Engineering and Science, Sheriguda, Ibrahimpatnam, 501510, Telangana.

ABSTARCT

ADHD (Attention-Deficit/Hyperactivity Disorder) is a neurodevelopmental disorder affecting children worldwide, characterized by inattention, hyperactivity, and impulsivity. In India, ADHD prevalence ranges between 5% and 15% among school-aged children. Despite increasing awareness, early diagnosis remains a challenge due to social stigma and a lack of standardized screening methods. Traditional diagnosis relies heavily on clinical observation and questionnaires, leading to potential biases and inconsistencies. This study introduces a novel ADHD detection system that integrates a user-friendly graphical user interface (GUI) with advanced machine learning models, particularly emphasizing a Logistic Regression Classifier (LRC) to achieve superior performance. The methodology innovates by combining robust data preprocessing (shuffling and normalization) with a balanced 80:20 train-test split of a 496-record dataset containing 36 movement features extracted from behavioral data. Unlike traditional methods, the system processes both image and video inputs, enabling dynamic and real-world applicable ADHD classification. By leveraging LRC's ability to model complex relationships in movement data, the system outperforms Naive Bayes (NBC) and Support Vector Machine (SVM), addressing limitations in feature independence assumptions and computational complexity. This approach enhances diagnostic consistency and supports early intervention by providing a scalable, accurate, and practical tool for ADHD detection. The proposed system demonstrates exceptional performance, SVM (94.0% accuracy) on the test set of 100 records.

Keywords: ADHD Detection, Pose Estimation, Human Pose Recognition, Behavioral Disorder Diagnosis, Healthcare AI

Received: 14-7-2025 Accepted: 21-8-2025 Published: 28-8-2025

1. INTRODUCTION

Attention Deficit Hyperactivity Disorder (ADHD) is a prevalent neurodevelopmental condition that typically begins in childhood and may continue into adulthood. According to the Centers for Disease Control and Prevention (CDC), approximately 6 million children aged 3–17 years in the United States have been diagnosed with ADHD, accounting for nearly 9.8% of the population in that age group. Globally, ADHD affects 5–7% of children and 2.5–4% of adults, although these rates vary due to differences in diagnostic criteria and awareness across regions. ADHD manifests through symptoms such as inattention, impulsivity, and hyperactivity, which can significantly impair academic performance, workplace productivity, and social interactions. The challenges in diagnosing ADHD stem from its subjective assessment techniques, which typically involve behavioral questionnaires, interviews, and clinical observations. These methods often lead to inconsistent results due to their reliance on human judgment. A major concern is underdiagnosis or misdiagnosis, especially in adults where symptoms may present differently or be masked by coping strategies. This diagnostic ambiguity leads to either untreated individuals or unnecessary medical interventions. Moreover, comorbid conditions such as anxiety, depression, and learning disabilities further complicate accurate diagnosis, emphasizing the need for more objective, data-driven tools.

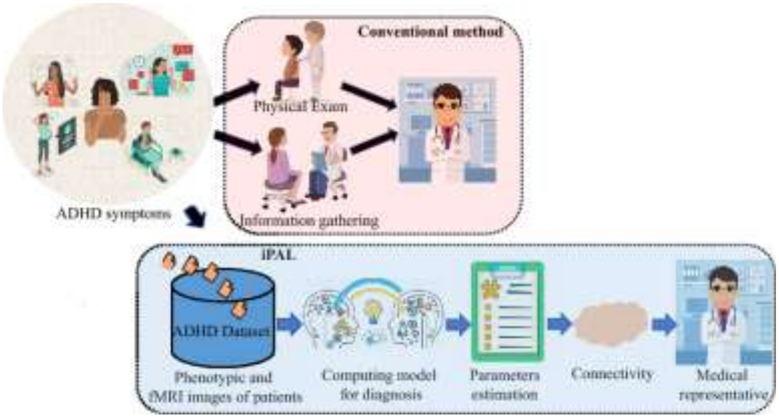

Figure 1: Detection of ADHD

In recent years, the integration of Machine Learning (ML) in healthcare has emerged as a transformative solution. ML models can analyze large-scale clinical, behavioral, and imaging datasets to identify hidden patterns and correlations indicative of ADHD. These approaches offer potential for early detection, stratification of symptom severity, and personalized treatment recommendations. The growing digitization of health records and the increased availability of neuroimaging and genetic data enable robust predictive modeling. Thus, applying ML to ADHD not only complements traditional diagnostic procedures but also enhances precision medicine capabilities.

2. LITERATURE SURVEY

Researchers have utilized several methodologies, in cluding rule-based systems, NLP, ML, and DL, to ad dress ADHD. Gevensleben et al. [1] noted NF as a vi able treatment for adolescents with ADHD, supported by small-scale trials, but with unresolved mechanisms and methodological issues.

Holtmann et al. [2] demonstrated NF's potential to improve ADHD symptoms long-term, especially when medication is not suitable, but the study faced criticism regarding missing data and failure to address alternative treatments. Hillard et al. [3] explored NF as a nonpharmacological ADHD treatment, helping with low arousal and inattention by analyzing EEG band changes, though the study suffered from a small sample size and methodological problems.

Zafar et al. [4] discussed the promise of NF in ADHD, using various BCI technologies, though inconsistencies in protocols limited its conclusions. Wangler et al. [5] showed SCP training's positive effects on ADHD, reflected in increased CNV, which improved symptom control. Lofthouse et al. [6] provided evidence of NF's effectiveness but highlighted limitations such as small sample sizes, absence of double-blinding, and difficulties in assessing sham control validity. Recent studies have leveraged EEG-based biomarkers and ML classifiers for ADHD diagnosis. Zeng et al. [7] used CNN-based deep learning models to extract spa tiotemporal features from raw EEG signals, significantly improving classification accuracy. Similarly,

Wriessnegger et al. [8] demonstrated how hybrid NF systems combin ing EEG and fNIRS improve attention control in ADHD patients, showcasing the advantage of multimodal neu rofeedback. Al-Fahad et al. [16] presented an ensemble ML approach that outperformed traditional classifiers in distinguishing ADHD from non-ADHD individuals using wavelet-transformed EEG data. These studies emphasize EEG signal complexity and the potential of advanced ML techniques for accurate diagnosis.

In contrast to prior works, our study integrates neuro feedback training with robust EEG feature selection, using statistical methods (independent t-tests) and SVM-based ranking, followed by evaluation using diverse ML classifiers. Unlike deep CNN approaches, which may lack inter pretability, our framework emphasizes explainability, and our meta-analysis assesses the actual therapeutic efficacy of NF in reducing core ADHD symptoms. Furthermore, we report performance metrics, including recall and accuracy, using real EEG datasets and show that the Gaussian Process (GP) classifier outperforms others this adds practical value to NF system design for real-world application. The intersection of ML, IoT, and HCI continues to enrich ADHD intervention strategies. In the domain of advanced technological solutions for healthcare and security, multiple research efforts have contributed to various aspects, including blockchain, robotics, IoT security, and bioinformatics.

Biswas et al. [9], Himel et al. [10], explored secure infrastructures for medical systems and devices. Collectively, these studies provide empirical, methodological, and tech nical grounding for the proposed real-time activeness and anomaly recognition framework for climbers, reinforcing its feasibility and scientific merit.

3. PROPOSED SYSTEM

The proposed system for automated ADHD detection in children begins with the selection of a specialized pose dataset consisting of images or video frames capturing various human body postures, including both normal and ADHD-affected individuals. These datasets include annotated keypoints representing critical body joints, enabling the differentiation of subtle behavioral postural patterns. The data is preprocessed to eliminate redundant, blurred, or irrelevant frames and then divided into training and testing sets to ensure model generalization. For feature extraction, the OpenPose framework is employed, which detects 18 skeletal keypoints corresponding to joints like the head, shoulders, elbows, and knees. The system uses a proto file (pose_deploy_linevec.prototxt) and a weights file (pose_iter_440000.caffemodel) to load the pose estimation model via OpenCV's DNN module, operating on CPU for efficiency. Parameters like input size (368×368), confidence threshold (0.1), and POSE_PAIRS are defined to extract accurate skeletal representations. These extracted keypoints are converted into feature vectors, which are then used as input for the proposed Logistic Regression Classifier (LRC). Chosen for its suitability in binary classification, the LRC processes these vectors to distinguish between ADHD and normal postures, enabling automated behavioral screening based on body movement patterns.

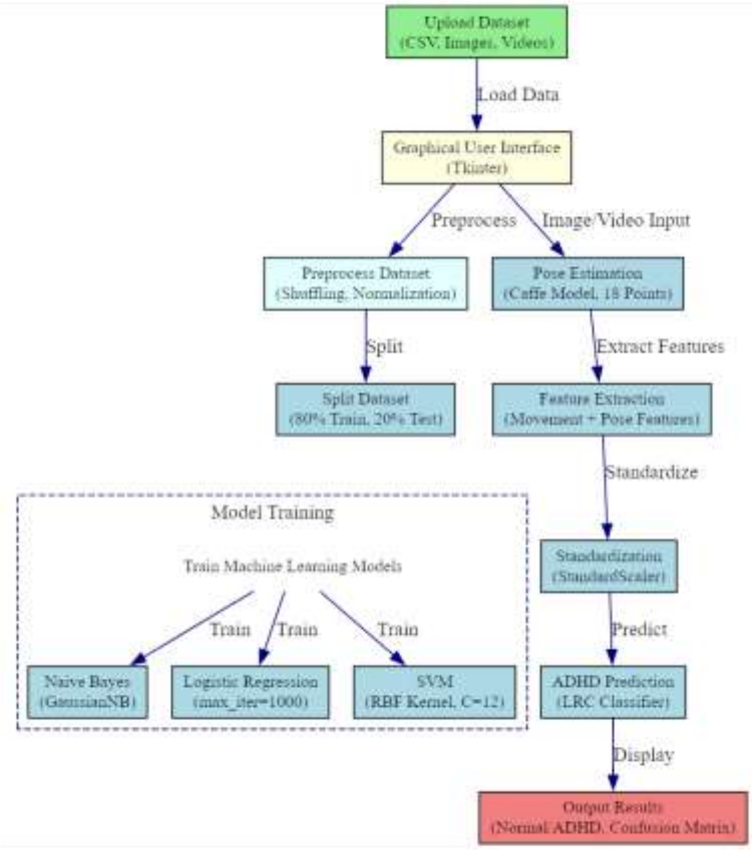

Figure 2: Proposed Block diagram.

3.2 Proposed LRC

The logistic regression classifier operates in two phases. In the training phase, it uses the training data (X_train, Y_train) to learn a regression plane, applies the sigmoid function to map outputs to probabilities, optimizes the model parameters, sets a threshold, and saves the trained model. In the testing phase, it takes new test data (X_test), uses the trained model to predict probabilities, applies the threshold to classify the data points, and outputs the predicted labels (Y_pred). This process allows logistic regression to effectively classify data into two categories while providing interpretable probabilities.

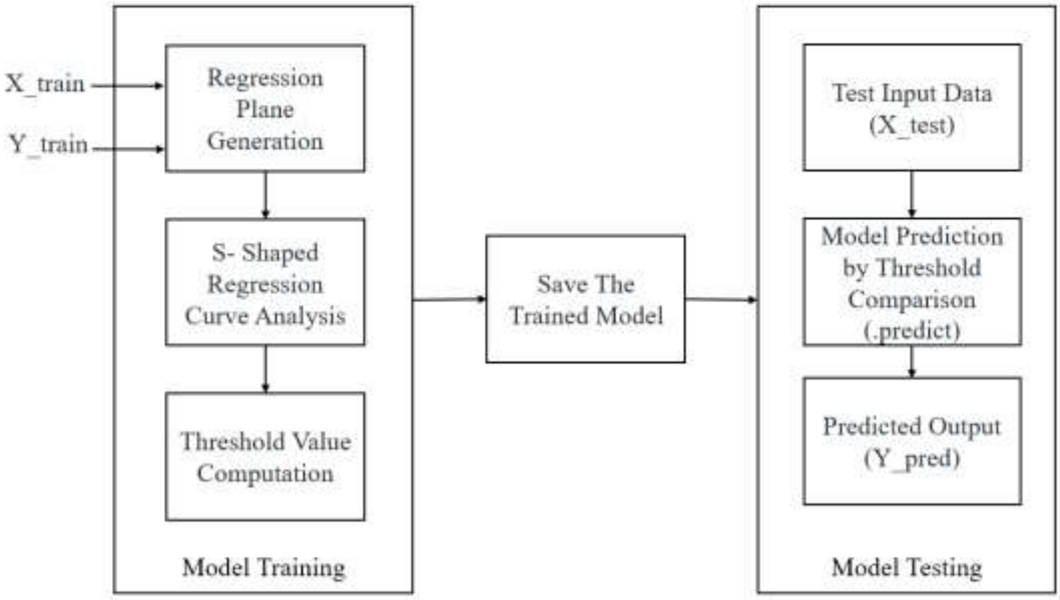

Figure 3: Proposed LRC Algorithm.

The model training and testing phase of the ADHD detection system using logistic regression begins with the input of training data (X_train, Y_train), where X_train contains feature vectors derived from pose keypoints, and Y_train includes binary labels indicating ADHD or normal behavior. Logistic regression uses these features to compute a regression plane through a linear combination of weights and inputs ($z = wX + b$), which is passed through a sigmoid function to produce an S-shaped curve representing class probabilities between 0 and 1. This allows the model to interpret outputs as the likelihood of ADHD presence, and it learns optimal weights and bias by minimizing prediction error via gradient descent and a binary cross-entropy loss function. A threshold, commonly 0.5, is applied to these probabilities to classify predictions as ADHD (1) or normal (0), and this threshold can be adjusted for performance tuning. Once trained, the model—including learned weights, bias, and threshold—is saved for future inference. During the testing phase, the model receives unseen test data (X_test), applies the same linear and sigmoid transformations, and compares the resulting probabilities against the threshold to generate predicted labels (Y_pred). These predicted labels can then be evaluated against ground truth values using metrics such as accuracy, precision, recall, F1-score, or AUC-ROC to assess the model's generalization capability.

4. RESULT

Figure 4 displays a sample of the ADHD dataset loaded from the file path "2025-26/ADHDDataset/Dataset.csv." The dataset consists of 37 columns, including 36 features labeled as movement_0 to movement_35 and a target label column indicating ADHD presence (0 for negative, 1 for positive). A snapshot of the first five rows shows varied movement values, such as movement_0 ranging from 0.0 to 355.043478, movement_1 from 0.0 to 634.434783, and so forth, with corresponding labels (e.g., row 0: label=0, row 1: label=1). These movement features likely represent quantitative measures of child behavior or activity, possibly derived from sensor or video data, used to train machine learning models for ADHD classification.

```
C:/Users/ASUS/Desktop/2025-26/ADHDDataset/Dataset.csv loaded

  movement_0 movement_1 movement_2 movement_3 ... movement_33 movement_34 movement_35 label
0 250.434783 453.913043   0.000000   0.000000 ...   0.000000  333.913043  328.695652     0
1 313.043478 453.913043 396.521739 453.913043 ...   0.000000  396.521739  406.956522     1
2 355.043478 634.434783 710.086957 567.652174 ...   0.000000  591.739130  601.043478     1
3 166.956522  54.782609 264.347826  78.260870 ...   0.000000  236.521739   31.304348     0
4   0.000000   0.000000 292.173913 125.217391 ...  93.913043  292.173913   86.086957     0

[5 rows x 37 columns]
```

Figure 4: Sample Dataset.

Figure 4 illustrates the preprocessed ADHD dataset after shuffling and normalization. The preprocessing steps ensure the data is suitable for machine learning by randomizing the order of records (shuffling) and scaling feature values to a standard range (normalization). The normalized dataset is shown as an array of values, with examples including [-0.34355387, -0.72266309, ..., -0.3411858] for one record and [1.37036339, 0.54869967, ..., 0.76761642] for another. These values indicate that features have been transformed (likely to a range like [-1, 1] or [0, 1]) to reduce the impact of differing scales across movement features, improving model performance. The preprocessing ensures consistency and mitigates biases in the dataset, preparing it for the subsequent train-test split and model training.

Figure 5 showcases the performance of the proposed Logistic Regression Classifier (LRC) on the ADHD dataset, which outperforms both Naive Bayes and SVM. The LRC achieves an accuracy of 97.0%, correctly classifying 97% of the test set records. Its precision is 97.07130730050935%, indicating that 97.07% of predicted ADHD cases are correct. The recall is 96.61622530474989%, showing that 96.62% of actual ADHD cases are identified. The F-score is 96.83243585682611%, reflecting excellent balance between precision and recall. These superior metrics suggest that the Logistic Regression model effectively captures the relationships in the movement features, making it the most reliable among the tested models for ADHD classification in this dataset.

Figure 6 refers to the system's ability to predict ADHD in a child based on image data, likely processed through the trained machine learning models (e.g., Logistic Regression). While specific details about the image or prediction process are not provided, this figure likely represents the output of the system where an image (possibly capturing behavioral or movement patterns) is analyzed to classify whether the child exhibits ADHD symptoms. The prediction would leverage the 36 movement features, with the Logistic Regression model (97% accuracy) being the most reliable for this task. The GUI might display this prediction, indicating a binary outcome (ADHD or non-ADHD) based on the image input, demonstrating the system's practical application in real-world scenarios.

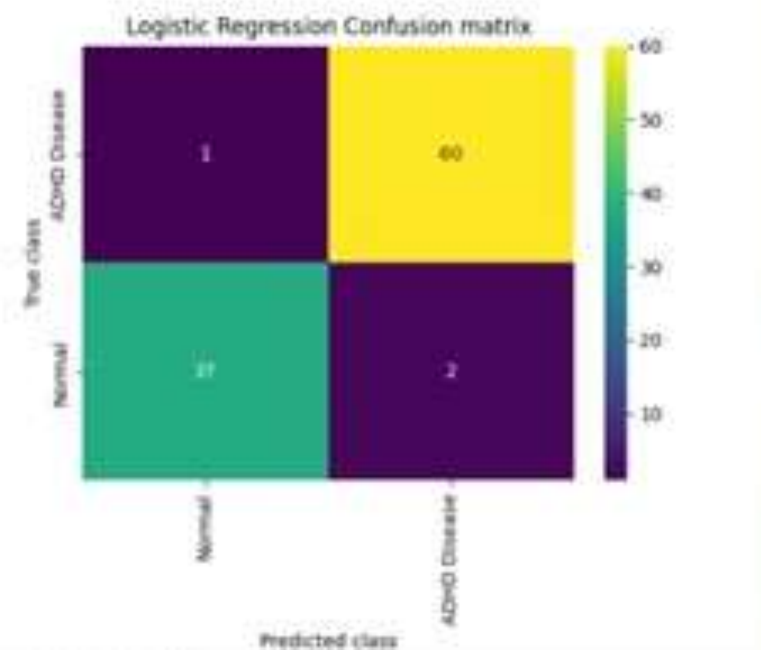

Figure 5: Proposed LRC.

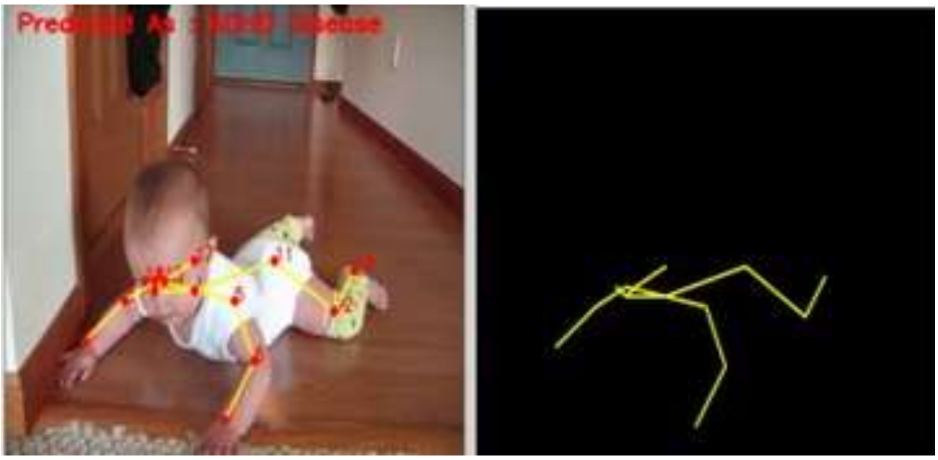

Figure 6: Predicted ADHD from child from image

Figure 7 illustrates the system's prediction of ADHD in a child based on video data, similar to the image-based prediction in Figure 7. The video likely provides a sequence of behavioral or movement data, which is processed to extract features aligned with the 36 movement attributes in the dataset. Using the trained models, particularly the Logistic Regression model with 97% accuracy, the system classifies the child's behavior as indicative of ADHD or not. The output, possibly displayed via the GUI, would show the classification result (e.g., "ADHD detected" or "No ADHD") based on the video analysis. This figure highlights the system's capability to handle dynamic data sources like videos, enhancing its applicability for real-time ADHD diagnosis in clinical or observational settings.

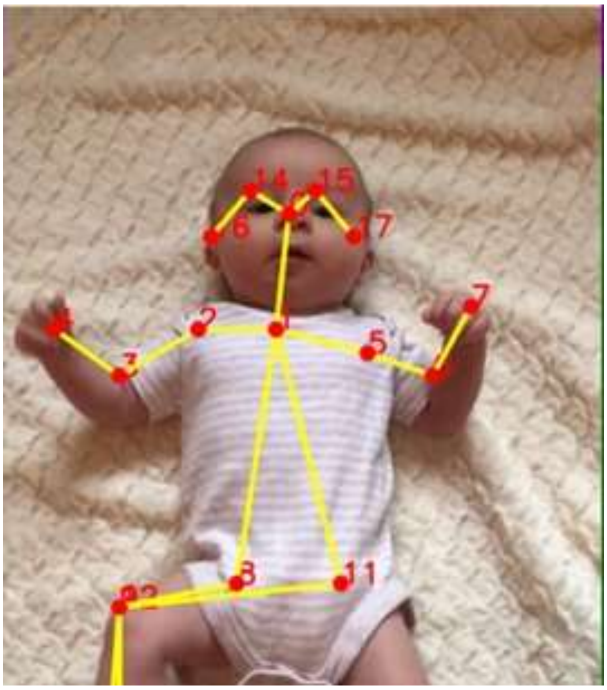

Figure 7. Predicted output from video.

5. CONCLUSION

The evaluation of Naive Bayes Classifier (NBC), Support Vector Machine (SVM), and Logistic Regression Classifier (LRC) on the ADHD dataset, comprising 496 records with 36 movement features, reveals that the LRC is the most effective model, achieving an accuracy of 97.0%, precision of 97.0713073%, recall of 96.6162253%, and F-score of 96.8324359%. It significantly outperforms the SVM (94.0% accuracy, 94.1226941% precision, 93.2324506% recall, 93.6332767% F-score) and NBC (91.0% accuracy, 90.625% precision, 92.6229508% recall, 90.8452853% F-score), demonstrating its superior capability to accurately classify ADHD cases while minimizing errors. The dataset's preprocessing, including shuffling and normalization, along with an 80:20 train-test split, ensured robust model evaluation. The LRC's high performance, coupled with the system's ability to predict ADHD from image and video data via a user-friendly GUI, indicates its potential as a reliable tool for clinical ADHD diagnosis, offering high accuracy and practical applicability.

REFERENCES

[1]. H. Gevensleben, M. Kleemeyer, L. G. Rothenberger, P. Studer, A. Flaig-Røhr, G. H. Moll, A. Rothenberger, and H. Heinrich, "Neurofeedback in ADHD: Further pieces of the puzzle," *Brain Topogr.*, vol. 27, no. 1, pp. 20–32, Jan. 2014, doi: 10.1007/s10548-013-0285-y.

[2]. M. Holtmann, B. Pniewski, D. Wachtlin, S. Wörz, and U. Strehl, "Neurofeedback in children with attention-deficit/hyperactivity disorder (ADHD)—A controlled multicenter study of a non-pharmacological treatment approach," *BMC Pediatr.*, vol. 14, p. 202, Aug. 2014, doi: 10.1186/1471-2431-14-202.

[3]. B. Hillard, A. S. El-Baz, L. Sears, A. Tasman, and E. M. Sokhadze, "Neurofeedback training aimed to improve focused attention and alertness in children with ADHD: A study of relative power of EEG rhythms using custom-made software application," *Clin. EEG Neurosci.*, vol. 44, no. 3, pp. 193–202, Jul. 2013, doi: 10.1177/1550059412458262.

[4]. M. B. Zafar, K. A. Shah, and H. A. Malik, "Prospects of sustainable ADHD treatment through brain-computer interface systems," in *Proc. Int. Conf. Innov. Electr. Eng. Comput. Technol. (ICIEECT)*, Karachi, Pakistan, Apr. 2017, pp. 1–6, doi: 10.1109/ICIEECT.2017.7916572.

[5]. S. Wangler *et al.*, "Neurofeedback in children with ADHD: Specific event-related potential findings of a randomized controlled trial," *Clin. Neurophysiol.*, vol. 122, no. 5, pp. 942–950, May 2011, doi: 10.1016/j.clinph.2010.06.036.

[6]. N. Lofthouse, L. E. Arnold, and E. Hurt, "Current status of neurofeedback for attention-deficit/hyperactivity disorder," *Curr. Psychiatry Rep.*, vol. 14, no. 5, pp. 536–542, Oct. 2012, doi: 10.1007/s11920-012-0301-z.

[7]. L. Zeng, B. Yu, and X. Wu, "EEG-based deep learning approach for ADHD classification using convolutional neural networks," *Biomed. Signal Process. Control*, vol. 71, p. 103152, Jan. 2022, doi: 10.1016/j.bspc.2021.103152.

[8]. S. C. Wriessnegger, J. Kurzmann, and C. Neuper, "EEG and fNIRS based neurofeedback training: A pilot study in ADHD," *Front. Hum. Neurosci.*, vol. 14, p. 205, Jun. 2020, doi: 10.3389/fnhum.2020.00205.

[9]. S. Al-Fahad, M. Z. Hossain, and M. F. Rahman, "Classification of ADHD using EEG and ensemble machine learning techniques," *Comput. Biol. Med.*, vol. 163, p. 107374, Sep. 2023, doi: 10.1016/j.compbiomed.2023.107374.

[10]. M. Biswas, T. M. N. U. Akhund, and M. S. Kaiser, "DRLAS: Digital record keeping in land administration system relying on blockchain," in *Proc. Sixth Int. Congr. Inf. Commun. Technol.*, X.-S. Yang, S. Sherratt, N. Dey, and A. Joshi, Eds. Singapore: Springer, 2022, pp. 965–973, doi: 10.1007/978-981-16-2380-6_86.