EEG-Based Assessment of Suicidality Risk: An Integrated Framework with Self-Adaptive Chaotic Cuckoo Search and AttentionBiSqueezeNet

B. S. Anjan Kumar
Department of Electronics and Instrumentation Engineering, Bangalore Institute of Engineering, Visvesvaraya Technological University, Belagavi, India
anjan.cmp@gmail.com (corresponding author)

H. N. Suresh
Department of Electronics and Communication, AMC Engineering College, Bengaluru, Visvesvaraya Technological University, Belagavi, India
drsuresh.narayanagowda@amceducation.in

S. Ranjitha
Software Engineering Department, Microland Eco Space, Bellanduru, Bengaluru, India
ranjithaneha@gmail.com

Received: 22 August 2025 | Revised: 14 September 2025, 4 October 2025, 19 October 2025, and 12 November 2025 | Accepted: 15 November 2025

 DOI: https://doi.org/10.48084/etasr.14247

ABSTRACT

This study proposes a novel integrated framework for electroencephalogram (EEG)-based suicide risk prediction designed to overcome key challenges, including signal noise, high dimensionality, and model interpretability. Our framework incorporates a robust preprocessing pipeline combining a modified Finite Impulse Response (FIR) filter and Fast Independent Component Analysis (Fast-ICA) for artifact removal. Then, a comprehensive suite of features spanning time, frequency, time-frequency, and functional connectivity domains is extracted, while dimensionality reduction is performed using a modified t-Distributed Stochastic Neighbor Embedding (t-SNE) technique. At the core of the framework is AttentionBiSqueezeNet, a hybrid Deep Learning (DL) architecture that combines Convolutional Neural Network (CNN) with an attention mechanism, a Bidirectional Long Short-Term Memory (Bi-LSTM), and a SqueezeNet backbone. To further enhance the capture of complex temporal dependencies, the activation functions of the Bi-LSTM are optimized using a Self-Adaptive Chaotic Cuckoo Search (SA-CCS) algorithm. The model was evaluated on a publicly available EEG dataset comprising 2,758 subjects and achieved an accuracy of 98.40%, precision of 98.87%, and sensitivity of 97.61%, outperforming all baseline methods. These results demonstrate that the proposed framework provides a robust, interpretable, and high-performing solution for early suicide risk detection using EEG data.

Keywords-Electroencephalogram (EEG); Suicide Risk Prediction; FastICA; Modified t-SNE; Attention Mechanism; Bi-LSTM; Self-Adaptive Chaotic Cuckoo Search (SA-CCS)

I. INTRODUCTION

Suicide represents a critical global public health crisis, claiming over 700,000 lives annually and standing as a leading cause of death, particularly among adolescents and young adults [1]. This complex, multifactorial behavior is often the tragic endpoint of a progression that begins with Suicidal Ideation (SI) and may advance to Suicidal Attempts (SA) [2]. The traditional suicidal risk assessment commonly employed in clinical psychiatry primarily relies on self-reported questionnaires and structured interviews based on diagnostic manuals like the Diagnostic and Statistical Manual of Mental Disorders 5th edition (DSM-5) and the International Classification of Diseases 11th revision (ICD-11) [3]. However, these methods are inherently subjective, susceptible to patient denial, clinician bias, and lack the objective biomarkers needed for precise, individualized risk stratification [4]. These critical limitations underscore the need for the development of novel, neurophysiology-based tools that can provide quantifiable and reliable biomarkers for identifying individuals at imminent risk.

Electroencephalography (EEG) has emerged as a premier modality in this pursuit, offering a non-invasive, cost-effective, and high-temporal-resolution window into the brain's electrophysiological dynamics [5]. The core hypothesis driving this research is that the neurophysiological correlates of emotional dysregulation, cognitive deficits, and impulsivity manifest in aberrant oscillatory patterns and functional connectivity within the brain [6]. Specific EEG biomarkers have been extensively investigated, with Frontal Alpha Asymmetry (FAA) being one of the most replicated findings, where a relative increase in alpha power in the right frontal hemisphere is often associated with withdrawal-related motivation and negative effects, potentially serving as a trait marker for depression and suicide risk [7, 8]. Furthermore, alterations in Power Spectral Density (PSD) across other frequency bands, such as increased theta and beta power in frontal regions, have been linked to the severity of depressive symptoms and acute suicidal states [9, 10].

The application of Machine Learning (ML) and Deep Learning (DL) to EEG analytics has revolutionized the field, enabling the extraction of complex, non-linear patterns that are often imperceptible to traditional analysis [11]. Supervised learning models have been trained on features derived from time-domain, frequency-domain, and connectivity metrics to classify individuals with SI/SA from healthy controls or those with non-suicidal depression [12, 13]. More recently, deep Convolutional Neural Networks (CNNs) and recurrent architectures like Long Short-Term Memory (LSTM) networks have demonstrated remarkable efficacy in automatically learning discriminative spatio-temporal features directly from raw or preprocessed EEG signals, often surpassing the performance of hand-crafted features [14, 15]. Despite these advancements, significant challenges persist. For instance, EEG signals are notoriously susceptible to artifacts from muscular, ocular, and environmental sources, necessitating robust preprocessing and denoising pipelines to ensure signal fidelity [16]. The high-dimensional nature of multi-channel EEG data also introduces substantial dimensionality, requiring effective feature selection and dimensionality reduction techniques to improve model generalizability and avoid overfitting [17]. Moreover, many existing DL models operate as "black boxes", lacking mechanisms to highlight the most salient features contributing to a decision, which is crucial for clinical interpretability and trust [18].

Despite these limitations, several ML and DL models have been successfully developed to identify high-risk individuals using clinical registry data from psychiatric care visits [19], electronic health records during hospitalizations [20], and to distinguish suicide attempters from those with ideation alone [21]. Furthermore, predictive accuracy is being enhanced through the analysis of novel data types, including social media language patterns [22], structural brain imaging [23], and resting-state EEG signals in depressed patients [24], collectively contributing to the development of powerful early-warning systems [25]. Related work on advanced Artificial Intelligence (AI) for EEG analysis includes [26], which presents an IoT-enabled DL approach for epilepsy detection.

To overcome limitations regarding noise artifacts, high-dimensionality, and limited model interpretability, this study proposes an integrated framework for suicide risk prediction from EEG data. The principal contributions are as follows:

- To propose a noise-robust preprocessing pipeline combining a modified Finite Impulse Response (FIR) filter and Fast Independent Component Analysis (Fast-ICA) for enhanced artifact removal.
- To extract a comprehensive suite of features spanning time-domain, frequency-domain, time-frequency, connectivity, and DL-based domains to capture the multifaceted nature of neural signatures associated with suicidality.
- To introduce a modified t-Distributed Stochastic Neighbor Embedding (t-SNE) technique for effective dimensionality reduction and visualization while preserving critical local structures in the data.
- To design AttentionBiSqueezeNet, which synergistically combines a CNN with an attention mechanism, an optimized Bidirectional LSTM (Bi-LSTM), and a SqueezeNet backbone. The activation functions of the Bi-LSTM are further optimized using a Self-Adaptive Chaotic Cuckoo Search (SA-CCS) algorithm to enhance the model's capacity to learn complex temporal dependencies inherent in EEG signals.

II. PROPOSED METHODOLOGY

The proposed methodology for predicting suicidal risk using EEG data is presented in Figure 1.

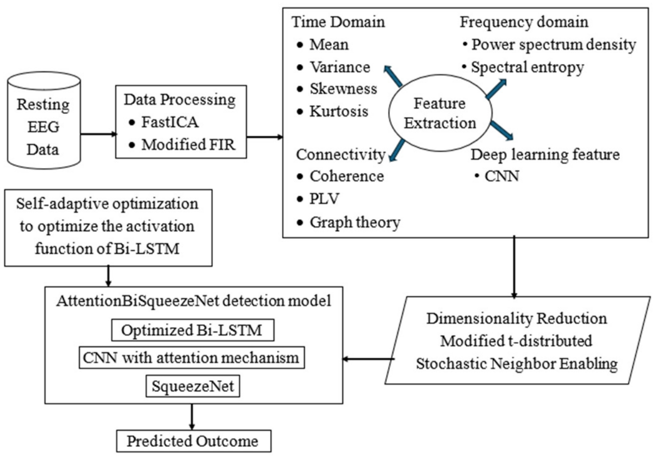

Fig. 1. Workflow of the proposed methodology.

A. *Data Collection and Validation Protocol*

The study utilized a publicly available EEG dataset licensed under CC-BY 4.0 International and specifically designed for suicide ideation detection [27]. The dataset comprises EEG recordings from 2,758 participants, including 1,500 healthy controls and 1,258 individuals clinically diagnosed with Major

Depressive Disorder (MDD) exhibiting SI. Data acquisition parameters included:

- EEG System: 128-channel HydroCel Geodesic Sensor Net.
- Sampling Rate: 500 Hz.
- Recording Condition: Resting state (eyes closed).
- Duration: 5-minute recordings per participant.
- Filtering: Band-pass filtered between 0.5-50 Hz.

1) *Validation Protocol*

To ensure robust and generalizable performance estimates, a stratified 10-fold cross-validation protocol (90% training and 10% testing) was employed for all experiments, maintaining the original class distribution in each fold.

B. *Data Preprocessing*

The first preprocessing step employed was the Fast-ICA algorithm [16] for artifact removal, denoising of EEG signals, and enhancing signal fidelity for subsequent analysis. This blind source separation technique effectively isolates neural signals from muscular and ocular artifacts by maximizing the non-Gaussianity of independent components. An example of the algorithm's output compared to the original signal is presented in Figure 2. Additionally, a customized FIR filter with a Kaiser window was designed to eliminate high-frequency noise and artifacts while maintaining linear phase response. The filter order M was optimized to achieve a flat frequency response in the passband with minimal phase distortion. A representative example of its performance is presented in Figure 3.

Both Figures 2 and 3 use diverse multicolored lines (cyan, orange, green, red, etc.) to distinguish individual EEG channels and components, illustrating the transformation from chaotic, noisy raw signals to separated components, and finally to smooth, oscillation-rich filtered waveforms.

C. *Feature Extraction*

Feature extraction was employed to capture various aspects of brain activity, identifying patterns relevant to subsequent analysis.

1) *Time-Domain Features*

Statistical measures, including mean, variance, skewness, and kurtosis of the EEG signal, were computed, serving as features that describe the central tendency, variability, and asymmetry of EEG signals and provide an overview of the underlying neural dynamics and their temporal variations.

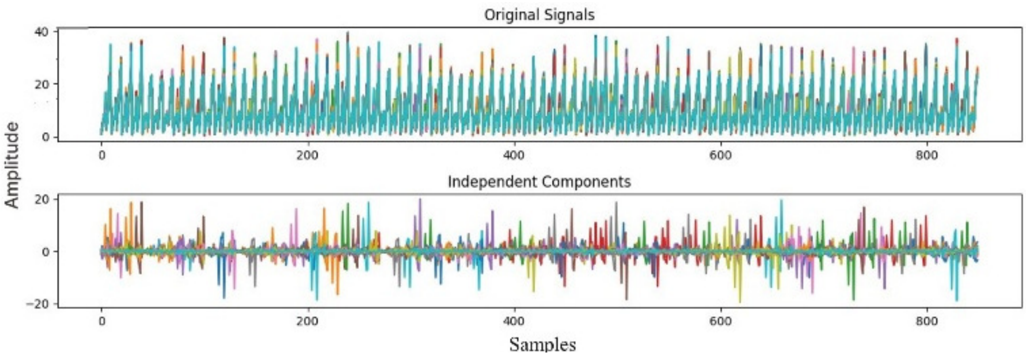

Fig. 2. Original signal vs Fast-ICA signal.

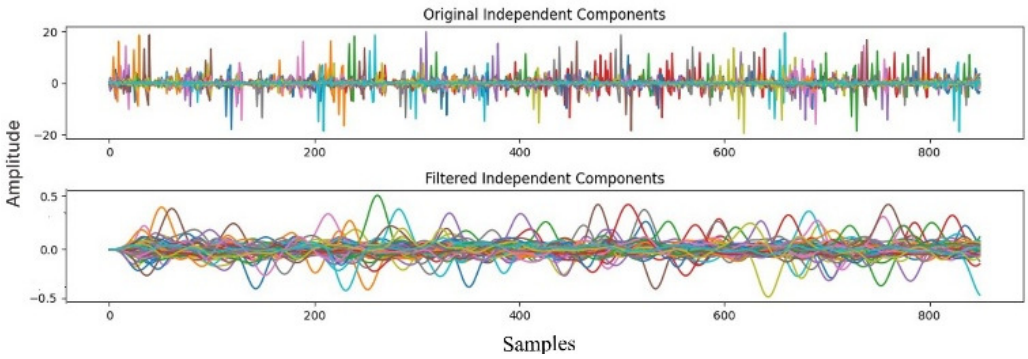

Fig. 3. Fast-ICA signal vs FIR filtered signal.

2) *Frequency Domain Features*

Frequency-domain analysis included exploring the distribution of frequency components in EEG signals using PSD and spectral entropy. PSD provides information on dominant frequencies and their intensities throughout various frequency bands, while spectral entropy represents a normalized form of Shannon entropy computed from the power spectrum amplitude aspects of the time series.

3) Time Frequency Features

For time-frequency analysis, we employed the Continuous Wavelet Transform (CWT) feature extraction technique to capture the evolution of signal characteristics in both time and frequency domains.

4) Connectivity Measures

The interactions and communication between various brain regions are explained by connectivity features such as coherence, Phase-Locking Value (PLV), and graph-theoretical measures. By demonstrating the synchronization and functional relationships among neural oscillations in different brain regions, these metrics offer information about the network dynamics that underlie cognitive functions and behaviors.

a) Coherence

Coherence quantifies the correlation or synchronization between two EEG signals at specific frequencies by examining the consistency of their phase and magnitude relationships.

b) Phase-Locking Value (PLV)

PLV measures the consistency of phase differences between two signals over time at a given frequency.

c) Graph-Theoretical Measures

The clustering coefficient, used as a graph-theoretical measure, measures the ratio of a node's connected neighbors to the total number of possible connections, and helps characterize network organization. To achieve this, brain regions are treated as nodes and the connections between them as edges.

5) Deep Learning Features

DL-based feature extraction was performed using CNNs, which enabled the automatic identification of complex patterns that may not be captured by traditional feature extraction techniques. CNNs are particularly effective in learning hierarchical representations from EEG data, allowing the extraction of intricate spatial features relevant to mental health conditions.

A typical CNN architecture consists of an input layer, three convolutional layers, two max-pooling layers, and one fully connected layer. The input layer receives the EEG-derived representation, which is processed through kernel filters to extract salient features. Convolutional layers apply successive filtering operations followed by activation functions and pooling layers, enabling the network to learn discriminative spatial patterns from the EEG signals. These learned features support the subsequent prediction or classification of conditions related to suicidal tendencies.

D. Dimensionality Reduction

In this study, complex, high-dimensional EEG features are mapped into a lower-dimensional space using a modified t-SNE approach. In the first phase, pairwise similarities between data points in the high-dimensional space are modeled as conditional probabilities based on their distances. In the second phase, the algorithm minimizes the mismatch between these high-dimensional similarities and the corresponding similarities in the low-dimensional embedding, thereby preserving local neighborhood structures.

E. Deep Learning Based Detection

The reduced-dimensional features obtained from the previous stage are provided as input to the detection model, AttentionBiSqueezeNet, which is a DL-based detection model that integrates multiple neural network architectures and optimization techniques to enhance EEG-based suicidal risk detection. This architecture combines SqueezeNet, a CNN with an attention mechanism, and an optimized Bi-LSTM network, as shown in Figure 4.

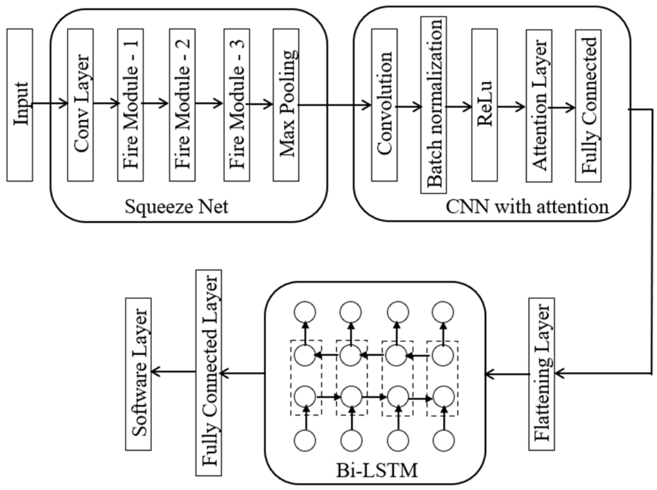

Fig. 4. Architecture of the proposed AttentionBiSqueezeNet model.

1) Bidirectional Long Short-Term Memory (Bi-LSTM)

The Bi-LSTM processes EEG sequences bidirectionally, enabling the model to simultaneously understand the context and dependencies of past and future information. The standard LSTM gate mechanisms (forget, input, and output gates) regulate information flow through the network, with activation functions optimized using the SA-CCS algorithm [16]. This self-adaptive optimization strategy dynamically adjusts parameters during training, enabling improved learning of complex temporal relationships present in EEG signals.

2) SqueezeNet

The SqueezeNet backbone provides efficient feature extraction with reduced parameter complexity through its fire modules and global average pooling layers.

3) Convolutional Neural Network (CNN)

The CNN component [14, 15] extracts spatial features from EEG data through convolutional and pooling layers, forming the foundational feature learning mechanism of our hybrid architecture.

III. RESULTS AND DISCUSSION

A. Performance Metrics and Analysis

The proposed AttentionBiSqueezeNet model was implemented in MATLAB R2023a and evaluated against several state-of-the-art DL models, including Bi-LSTM, SqueezeNet, CNN, Gated Recurrent Unit (GRU), and Multilayer Perceptron (MLP), using hyperparameter tuning via grid search for each model and reporting the average performance results across all 10 folds of the 10-fold cross-validation protocol.

The performance was evaluated using a comprehensive set of metrics, including accuracy, precision, sensitivity (Recall), specificity, F1-score, False Positive Rate (FPR), False Negative Rate (FNR), and Matthews Correlation Coefficient (MCC). The comparative results are presented in Table I.

TABLE I. COMPARISON OF DIFFERENT PERFORMANCE METRICS AND ANALYSIS

Model	Proposed	Bi-LSTM	Squeeze Net	CNN	GRU	MLP
Accuracy	98.40%	92.10%	94.30%	93.50%	91.80%	89.50%
Precision	98.87%	90.50%	92.80%	91.20%	89.70%	87.10%
Sensitivity	97.61%	89.80%	91.50%	92.10%	90.30%	86.90%
Specificity	99.07%	94.20%	96.80%	94.80%	93.20%	91.80%
F1-Score	98.24%	90.10%	92.10%	91.60%	90.00%	87.00%
FPR	0.93%	5.80%	3.20%	5.20%	6.80%	8.20%
FNR	2.39%	10.20%	8.50%	7.90%	9.70%	13.10%
MCC	0.967	0.841	0.885	0.87	0.835	0.787

The proposed AttentionBiSqueezeNet model outperformed all baseline models across all evaluation metrics. Specifically, the accuracy of 98.40%, combined with the high precision (98.87%) and sensitivity (97.61%), demonstrates the model's strong capability to correctly classify both high-risk and low-risk subjects. Additionally, the high specificity of 99.07% and the very low FPR of 0.93% are particularly important for clinical screening tools, as they minimize false alarms that could otherwise lead to unnecessary psychological distress and inefficient allocation of limited mental health resources. Equally important, the low FNR of 2.39% indicates that only a small proportion of high-risk individuals were misclassified as low risk, highlighting the model's effectiveness in minimizing missed detections. Lastly, the high MCC value of 0.967 further confirms the overall reliability and balanced predictive performance of the proposed approach, even under potential class imbalance conditions. These results compare very favorably to the current state-of-the-art in EEG-based suicide risk prediction, which typically reports accuracies in the range of 85-95% for complex DL models.

To further validate the performance of the proposed model, Table II depicts the confusion matrix, showcasing that the model correctly identified 1,228 subjects as high-risk and 1,486 subjects as low-risk, while only 30 false negatives and 14 false positives were observed.

B. *Feature Importance Analysis*

Feature importance was analyzed using the attention weights generated by the AttentionBiSqueezeNet model. Table III presents the average attention weights assigned to key feature categories, which are derived from the attention layer applied to the fused feature vector immediately before classification. Higher weights indicate a greater contribution of the corresponding features to the final prediction.

The analysis revealed that connectivity-related features, particularly PLV in frontal brain regions, along with PSD features in the alpha (8-13 Hz) and theta (4-8 Hz) frequency bands, were the most discriminative for suicide risk prediction. These findings are consistent with established neurophysiological biomarkers, such as FAA and altered theta-band connectivity, which have been associated with SI in prior studies.

TABLE II. CONFUSION MATRIX FOR ATTENTIONBISQUEEZENET

TYPE OF RISK	Predicted: High Risk	Predicted: Low Risk
Actual: High Risk	1228	30
Actual: Low Risk	14	1486

TABLE III. AVERAGE ATTENTION WEIGHTS FOR KEY FEATURE CATEGORIES

Feature Category	Specific Feature	Average Attention Weight
Connectivity Features	PLV (Frontal Regions)	0.217
PSD	Alpha Band (8-13 Hz)	0.198
PSD	Theta Band (4-8 Hz)	0.185
Time-Frequency Features	CWT Coefficients (Theta Range)	0.124
Graph-Theoretical Measures	Clustering Coefficient (Frontal Network)	0.112
Time-Domain Features	Kurtosis	0.088
DL Features	CNN-extracted Spatio-temporal Patterns	0.076

IV. CONCLUSION

This study introduced an integrated Electroencephalogram (EEG)-based suicide risk prediction framework centered on the proposed AttentionBiSqueezeNet architecture. The framework was designed to address critical limitations observed in existing approaches, including sensitivity to artifacts, incomplete characterization of EEG dynamics, and insufficient modeling of temporal dependencies.

Artifact contamination was mitigated through the joint application of a modified Finite Impulse Response (FIR) filter and Fast Independent Component Analysis (Fast-ICA), ensuring that relevant neural activity was preserved. Feature extraction was performed across time, frequency, time-frequency, and functional connectivity domains, enabling the model to capture complementary aspects of EEG signals that are often overlooked when relying on a single representation. To manage feature redundancy and high dimensionality, a modified t-Distributed Stochastic Neighbor Embedding (t-SNE)-based dimensionality reduction strategy was employed.

At the classification stage, the proposed AttentionBiSqueezeNet model integrates convolutional layers for spatial feature learning with a Bidirectional Long Short-Term Memory (Bi-LSTM) to model temporal dependencies in EEG sequences. Additionally, the activation functions within the Bi-LSTM were optimized using the Self-Adaptive Chaotic Cuckoo Search (SA-CCS) algorithm.

Quantitative results demonstrate that the proposed model consistently outperforms all baseline methods across standard evaluation metrics, including accuracy (98.40%), precision (98.87%), sensitivity (97.61%), and specificity (99.07%), indicating strong discriminative capability for both high-risk and low-risk subjects

Nevertheless, a significant limitation of this study is the limited size of the dataset used of 2,758 subjects. Future work will involve validating this framework on larger, multi-center, and more diverse datasets to enhance generalizability. Furthermore, although the attention mechanism provides a degree of interpretability, further analysis of specific EEG biomarkers, such as distinct frequency bands and connectivity patterns, will be pursued. This deeper investigation is essential for strengthening clinical trust and supporting the translation of the proposed framework from a predictive model into a clinically actionable diagnostic aid.

ACKNOWLEDGMENT

This research involved the analysis of human EEG data for suicide risk prediction. The study protocol was approved by the Institutional Review Board (IRB) / Ethics Committee of Bangalore Institute of Technology (Approval No: BIT/EC/2023/476). The research was conducted in accordance with the ethical principles of the Declaration of Helsinki.

Informed consent was obtained from all participants prior to their inclusion in the study. All data were anonymized and de-identified to protect participant confidentiality and privacy. Given the sensitive nature of suicide-related research, additional precautions were taken to ensure data security and ethical handling throughout the study.

REFERENCES

[1] World Health Organization, *Live Life: an Implementation Guide for Suicide Prevention in Countries*, 1st ed. Geneva: World Health Organization, 2021.

[2] M. K. Nock, G. Borges, E. J. Bromet, C. B. Cha, R. C. Kessler, and S. Lee, "Suicide and Suicidal Behavior," *Epidemiologic Reviews*, vol. 30, no. 1, pp. 133–154, May 2008, https://doi.org/10.1093/epirev/mxn002.

[3] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition. American Psychiatric Association, 2013.

[4] J. C. Franklin *et al.*, "Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research.," *Psychological Bulletin*, vol. 143, no. 2, pp. 187–232, 2017, https://doi.org/10.1037/bul0000084.

[5] B. H. Brinkmann, M. R. Bower, K. A. Stengel, G. A. Worrell, and M. Stead, "Large-scale electrophysiology: Acquisition, compression, encryption, and storage of big data," *Journal of Neuroscience Methods*, vol. 180, no. 1, pp. 185–192, May 2009, https://doi.org/10.1016/j.jneumeth.2009.03.022.

[6] S. Tian *et al.*, "Prediction of suicidality in bipolar disorder using variability of intrinsic brain activity and machine learning," *Human Brain Mapping*, vol. 44, no. 7, pp. 2767–2777, May 2023, https://doi.org/10.1002/hbm.26243.

[7] R. J. Davidson, "Anterior cerebral asymmetry and the nature of emotion," *Brain and Cognition*, vol. 20, no. 1, pp. 125–151, Sept. 1992, https://doi.org/10.1016/0278-2626(92)90065-T.

[8] J. A. Coan and J. J. B. Allen, "Frontal EEG asymmetry as a moderator and mediator of emotion," *Biological Psychology*, vol. 67, no. 1–2, pp. 7–50, Oct. 2004, https://doi.org/10.1016/j.biopsycho.2004.03.002.

[9] A. C. Nugent, E. D. Ballard, J. R. Gilbert, P. K. Tewarie, M. J. Brookes, and C. A. Zarate, "Multilayer MEG functional connectivity as a potential marker for suicidal thoughts in major depressive disorder," *NeuroImage: Clinical*, vol. 28, 2020, Art. no. 102378, https://doi.org/10.1016/j.nicl.2020.102378.

[10] N. Krepel, L. Benschop, C. Baeken, A. T. Sack, and M. Arns, "An EEG signature of suicidal behavior in female patients with major depressive disorder? A non-replication," *Biological Psychology*, vol. 161, Apr. 2021, Art. no. 108058, https://doi.org/10.1016/j.biopsycho.2021.108058.

[11] B. Abibullaev, A. Keutayeva, and A. Zollanvari, "Deep Learning in EEG-Based BCIs: A Comprehensive Review of Transformer Models, Advantages, Challenges, and Applications," *IEEE Access*, vol. 11, pp. 127271–127301, 2023, https://doi.org/10.1109/ACCESS.2023.3329678.

[12] D. Arunachalam and N. Kumar, "Benefit-based consumer segmentation and performance evaluation of clustering approaches: An evidence of data-driven decision-making," *Expert Systems with Applications*, vol. 111, pp. 11–34, Nov. 2018, https://doi.org/10.1016/j.eswa.2018.03.007.

[13] M. Duan, L. Wang, X. Liu, F. Su, L. An, and S. Liu, "Abnormal brain activity in fronto-central regions in mental disorders with suicide: An EEG Study," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Mexico, Nov. 2021, pp. 1035–1038, https://doi.org/10.1109/EMBC46164.2021.9629876.

[14] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *Journal of Neural Engineering*, vol. 14, no. 1, Feb. 2017, Art. no. 016003, https://doi.org/10.1088/1741-2560/14/1/016003.

[15] S. U. Amin, M. Alsulaiman, G. Muhammad, M. A. Mekhtiche, and M. Shamim Hossain, "Deep Learning for EEG motor imagery classification based on multi-layer CNNs feature fusion," *Future Generation Computer Systems*, vol. 101, pp. 542–554, Dec. 2019, https://doi.org/10.1016/j.future.2019.06.027.

[16] I. Winkler, S. Brandl, F. Horn, E. Waldburger, C. Allefeld, and M. Tangermann, "Robust artifactual independent component classification for BCI practitioners," *Journal of Neural Engineering*, vol. 11, no. 3, June 2014, Art. no. 035013, https://doi.org/10.1088/1741-2560/11/3/035013.

[17] L. Van der Maaten and G. Hinton, "Visualizing Data using t-SNE," *Journal of Machine Learning Researc*h, vol. 9, pp. 2579–2605, Nov. 2008.

[18] A. Vaswani *et al.*, "Attention Is All You Need." arXiv, 2017, https://doi.org/10.48550/ARXIV.1706.03762.

[19] Q. Chen *et al.*, "Predicting suicide attempt or suicide death following a visit to psychiatric specialty care: A machine learning study using Swedish national registry data," *PLOS Medicine*, vol. 17, no. 11, Nov. 2020, Art. no. e1003416, https://doi.org/10.1371/journal.pmed.1003416.

[20] J. B. Edgcomb, T. Shaddox, G. Hellemann, and J. O. Brooks, "Predicting suicidal behavior and self-harm after general hospitalization of adults with serious mental illness," *Journal of Psychiatric Research*, vol. 136, pp. 515–521, Apr. 2021, https://doi.org/10.1016/j.jpsychires.2020.10.024.

[21] S. Ryu, H. Lee, D.-K. Lee, S.-W. Kim, and C.-E. Kim, "Detection of Suicide Attempters among Suicide Ideators Using Machine Learning," *Psychiatry Investigation*, vol. 16, no. 8, pp. 588–593, Aug. 2019, https://doi.org/10.30773/pi.2019.06.19.

[22] A. Roy, K. Nikolitch, R. McGinn, S. Jinah, W. Klement, and Z. A. Kaminsky, "A machine learning approach predicts future risk to suicidal ideation from social media data," npj Digital Medicine, vol. 3, no. 1, May 2020, Art. no. 78, https://doi.org/10.1038/s41746-020-0287-6.

[23] J.-C. Weng, T.-Y. Lin, Y.-H. Tsai, M. Cheok, Y.-P. Chang, and V. Chen, "An Autoencoder and Machine Learning Model to Predict Suicidal Ideation with Brain Structural Imaging," *Journal of Clinical Medicine*, vol. 9, no. 3, Feb. 2020, Art. no. 658, https://doi.org/10.3390/jcm9030658.

[24] F. Amico *et al.*, "A resting state EEG study on depressed persons with suicidal ideation," *IBRO Neuroscience Reports*, vol. 14, pp. 346–352, June 2023, https://doi.org/10.1016/j.ibneur.2023.03.012.

[25] L. Zheng *et al.*, "Development of an early-warning system for high-risk patients for suicide attempt using deep learning and electronic health

records," *Translational Psychiatry*, vol. 10, no. 1, Feb. 2020, Art. no. 72, https://doi.org/10.1038/s41398-020-0684-2.

[26] A. Jaffar, "IoT-enabled EEG-based Epilepsy Detection using Multilayer Deep Learning and the Evolutionary Algorithm Approach," *Engineering, Technology & Applied Science Research*, vol. 14, no. 5, pp. 16595–16603, Oct. 2024, https://doi.org/10.48084/etasr.8270.

[27] P. Wongupparaj, "Resting-state EEG of adolescents with minimal, mild, and moderate depression." OSF, 2021, https://doi.org/10.17605/OSF.IO/4HQ3Y.