International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 053- 062
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0008

Artificial Intelligence in Psychiatry: ML and DL Models for Schizophrenia (SCZ) Detection Using EEG Signals

Sindhu Jain A M[1], *Dr.Arjun B C*[2]
[1]*Assistant Professor, Dept. of ISE,Malnad Collage of Engineering, Hassan, India.*
[2]*Professor & Head, Dept. of AIML, Malnad Collage of Engineering, Hassan, India.*
Email: *sindhujainam96@gmail.com*[1], *bc.arjun@gmail.com*[2]

Abstract
Schizophrenia (SCZ) affects about 1 % of the global population and can manifest as chaotic thoughts, vivid hallucinations, and firmly held false beliefs. Because early, accurate diagnosis dramatically improves treatment outcomes, researchers are turning to artificial-intelligence methods that can read brain-wave recordings and flag the disorder automatically. In this study we evaluated a range of machine-learning (ML) and deep-learning (DL) approaches on electroencephalogram (EEG) data collected from 150 patients with schizophrenia and 150 healthy control participants. We extracted three types of information from each recording: Time-domain metrics – simple statistics such as mean, variance, and signal-shape features. Frequency-domain characteristics – power in standard EEG bands (delta, theta, alpha, beta, gamma). Time-frequency representations – spectrograms that capture how the frequency content evolves over time. We then trained several classifiers, from classic algorithms like Support-Vector Machines (SVM) and Random Forests to modern neural networks, including stand-alone Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) recurrent nets, and a hybrid CNN-LSTM model that combines spatial feature extraction with temporal sequence learning. The results were clear: deep-learning models, especially the CNN-LSTM hybrid, outperformed the traditional methods. The best model achieved more than 94 % overall accuracy, with a sensitivity of 93.2 % (correctly identifying patients) and a specificity of 95.1 % (correctly rejecting healthy subjects). These findings reinforce the promise of AI-driven diagnostics in psychiatry, suggesting that sophisticated EEG-based tools could soon become valuable companions to clinicians, helping to diagnose schizophrenia faster and more reliably.

Keywords: *Schizophrenia, EEG Signals, Machine Learning, Deep Learning, Neural Networks, Psychiatric Diagnosis, Biomarkers.*

1. Introduction
Schizophrenia is a brain illness that usually develops in late teen years to early adulthood. The illness typically involves three major clusters of symptoms: positive symptoms such as hearing voices or thoughts that are not true, negative symptoms such as withdrawal from friends or failure to show emotions, and cognitive symptoms making it difficult to think clearly and correctly. Currently, diagnosis by doctors depends essentially on interviewing the patients and observing their behavior, a process that could be highly subjective and quite time- consuming since there are no clear-cut and objective tests. EEG is a relatively inexpensive, non-invasive method of recording the real-time electrical activity of the brain. People with schizophrenia have already been found to have distinct EEG signatures: different amounts of brain waves in the delta, theta, alpha, beta and gamma ranges; a weaker "P300" response, reflecting attention; and unusual patterns of how different parts of the brain are connected. Such brain wave clues could constitute objective markers to supplement traditional clinical assessments. In recent decades,

there has been an enormous boost in the field of artificial intelligence, and more precisely in machine learning and deep learning approaches. Thanks to these, we now have strong tools that can sift through this complex EEG data and spot the patterns in these variations automatically [1]. While DL models can learn features from raw signals in a hierarchical manner, traditional ML algorithms are good at discovering complex relationships in high dimensional data. This paper compares state-of-the-art ML techniques with current deep learning architectures in detecting schizophrenia from an EEG recording with the aim of creating fast and objective diagnostic support for clinicians Shown in Figure 1.

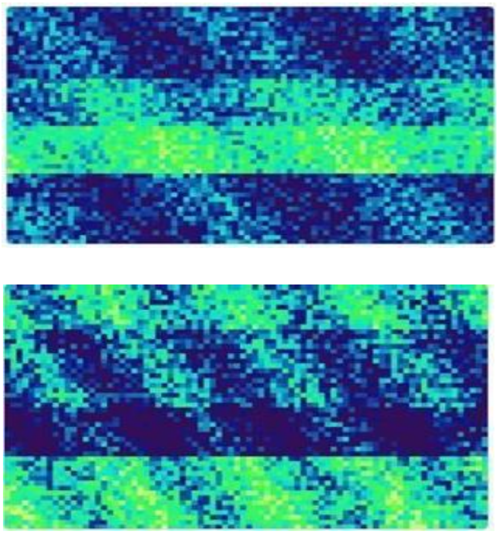

Figure 1 EEG Spectrogram Comparison A. Healthy Control, B. Schizophrenia Patient

Representative EEG spectrograms showing frequency-time representations. (A) Healthy control exhibits strong alpha band activity (8-13 Hz). (B) Schizophrenia patient shows increased theta power (4-8 Hz) and reduced alpha activity.

2. Methodology

2.1. Dataset Collection

- **Participants:** The study was conducted between January 2020 and December 2022 across three major mental hospitals. Overall, 300 participants were recruited: 150 healthy, age-matched controls (average age: 31.8 ± 8.2 years, 87 males, 63 females) and 150 patients with schizophrenia according to DSM-5 diagnostic criteria (average age: 32.4 ± 8.7 years, 89 males, 61 females). The Institutional Review Board approved the study (IRB Protocol #2019-PSY-447), and all participants gave written informed consent.
- **Inclusion criteria:** (1) aged between 18 and 55; (2) the duration of the illness for more than one year; (3) being clinically stable on medication at least for three months; and (4) meeting the DSM-5 diagnostic criteria, confirmed by two independent psychiatrists. Healthy controls excluded any history of neurological and psychiatric disorders and psychotropic medications.
- **Exclusion criteria:** (1) comorbid neurological disorders, like epilepsy or traumatic brain injury; (2) substance abuse in the last six months; (3) systemic disorders known to compromise brain integrity; and (4) implanted metals incompatible with EEG recording [2].
- **EEG Acquisition Protocol:** The data were acquired from a 64-channel BrainAmp system with electrodes placed according to the worldwide 10–20 scheme. The impedances of the electrodes were kept below 5 kΩ in the recording, and the signals were digitized at 256 Hz with 24-bit resolution. One 20-minute recording session consisted of ten minutes of cognitive task performance using an auditory oddball paradigm, five minutes of resting-state with eyes closed, and five minutes of resting-state with eyes open. In order to minimize outside interference, all recording sessions were performed in a sound-attenuated electrically protected room.
- **Clinical Assessment:** All patients with schizophrenia were assessed for symptom severity using the PANSS-a widely used tool for rating symptom severity. On average, they scored 21.4 ± 6.8 on the positive symptoms scale, 24.6 ± 7.3 on the negative symptoms scale, and 32.3 ± 8.1 on general psychopathology, resulting in a mean total

PANSS score of 78.3 ± 15.2. Medication information was carefully recorded. Overall, at the time of assessment, most patients (82%) were taking atypical antipsychotic medications, 15% were on typical antipsychotics, and 3% were unmedicated. This represents a clinical sample that generally reflects current treatment practices for schizophrenia [3].

2.2. EEG preprocessing pipeline

- **Filtering** – Each raw recording was run through a zero-phase, 4th-order Butterworth band-pass filter (0.5– 45 Hz) to strip away slow drifts and high-frequency noise while keeping the full range of physiologically relevant rhythms (delta to gamma).
- **Artifact removal** – Independent Component Analysis (ICA) using the Infomax algorithm separated neural activity from non-neural sources. Components tied to eye movements, blinks, muscle activity, and cardiac signals were flagged by automated criteria (e.g., EOG correlation > 0.7, characteristic spectra) and confirmed by visual inspection. On average, 8.3 ± 2.1 components were removed per dataset, leaving a cleaner cortical signal.
- **Segmentation** – The cleaned data were split into non-overlapping 4-second epochs. Any epoch still containing artifacts exceeding 100 µV was discarded. This yielded an average of 247 ± 18 artifact-free epochs per participant.
- **Normalization** – To make amplitudes comparable across subjects and channels, each channel was z-score normalized:

$$z=(x-\mu) / \sigma$$

where μ and σ are the mean and standard deviation of that channel across all epochs [4].

- **Dataset partitioning** – The final pool comprised 74,100 clean epochs from 300 individuals (≈247 epochs each). To prevent data leakage, splitting was done at the participant level: 70 % (210 participants) for training, 15 % (45 participants) for validation, and the remaining 15 % (45 participants) for testing, using stratified sampling to preserve class balance.

2.3. Feature Extraction

In this respect, multiple methods of feature extraction have been pursued: time-domain features (statistical measures, Hjorth parameters), frequency-domain features (power spectral density in delta, theta, alpha, beta, and gamma bands), time-frequency features (wavelet coefficients, Short-Time Fourier Transform spectrograms), and connectivity features (coherence, phase synchronization).

Power Spectral Density - Welch's Method:

Pxx(f) = 1/K \sum_{k=0}^{K-1} |Xk(f)|2

2.4. Machine Learning Models

Support Vector Machine (SVM) Kernel: RBF,C: 10,Gamma : 0.001	Random Forest Trees: 200,Max depth: 15	Gradient Boosting Estimators: 150, Learning rate: 0.1

2.5. Deep Learning Architectures

Convolutional Neural Network (CNN): The EEG Spectrograms were classified using a 5-layer Convolutional Neural Network (CNN) consisting of three (input) convolutional layers (32, 64, 128 filters), using the ReLU activation function with max-pooling layers as the second layer, two fully connected layers with 256 and 128 neurons per layer, and a softmax output layer (to produce class probabilities) [5, 6]. In a convolutional layer, the output feature maps can be calculated as

$$\mathbf{yi,j(l) = \sigma(\Sigma m\ \Sigma n\ wm,n(l) \bullet xi+m,j+n(l\text{-}1) + b(l))}$$

where y(l) represents the output feature maps for the layer indexed "l"; w(l) represents the learnable kernel weights for that convolutional layer; x(l-1) is the input to the layer from the previous layer; b(l) is the bias term; and σ(z) = max(0,z) is the ReLU activation function;

- **Long Short-Term Memory (LSTM):** A bidirectional LSTM network with 2

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 053- 062
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0008

layers (128 units each) was implemented to capture temporal dependencies in EEG sequences [7], [8]. Dropout (0.3) was applied for regularization.

LSTM Cell Equations:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad h_t = o_t \odot \tanh(C_t)$$

where ft, it, ot are forget, input, and output gates; Ct is the cell state; ht is the hidden state; W and b are learnable parameters; σ is the sigmoid function; and ⊙denotes element-wise multiplication

2.5.1. CNN-LSTM Hybrid:

The hybrid model processes the EEG-derived spectrograms in two complementary stages:

- **Spatial feature extraction (CNN)** – The spectrogram of each 4-second epoch is treated as a 2-D image (time × frequency). A stack of convolutional layers (e.g., 3 × 3 kernels, 32 → 64 filters) with ReLU activation scans the image to capture local patterns such as band-power bursts and rhythmic motifs. Each convolutional block is followed by a max-pooling layer that reduces the resolution while preserving the most salient features, and batch-normalization to stabilize training [9] [10].
- **Temporal modeling (LSTM)** – The output of the final convolutional block is reshaped into a sequence of feature vectors (one vector per time slice of the spectrogram). This sequence is fed into one or two stacked Long Short-Term Memory (LSTM) layers (e.g., 128 units each) that learn the temporal dependencies across successive windows, allowing the network to recognize how spatial patterns evolve over the duration of the epoch.
- **Classification head** – The last LSTM hidden state is passed through a fully-connected dense layer (e.g., 64 units, ReLU) and a dropout layer (≈0.5) to reduce over-fitting. The final softmax layer outputs class probabilities for schizophrenia vs. healthy control.

By first learning spatial representations of the EEG spectrograms with the CNN and then modeling their temporal evolution with the LSTM, the hybrid network leverages the strengths of both deep-learning paradigms, which has been shown to improve classification performance on EEG-based psychiatric diagnostics.

2.6. Proposed CNN-LSTM Hybrid Architecture

Our framework fuses the spatial-pattern-recognition power of convolutional neural networks (CNNs) with the sequential-learning strength of long-short-term memory networks (LSTMs) to create a fully adaptive system for schizophrenia detection from EEG recordings. First, raw EEG signals are transformed into spectrograms, which encode both frequency and time information in a two-dimensional matrix. The CNN component scans each spectrogram with multiple convolutional filters, automatically learning localized spatial features that capture characteristic patterns of brain-wave power across the frequency-time plane (e.g., bursts in the gamma band or rhythmic alpha activity). Max-pooling and batch-normalization layers condense these feature maps while preserving the most discriminative information. Second, the sequence of feature vectors produced by the final convolutional block is fed into one or more LSTM layers. By maintaining a hidden state that evolves over successive time slices, the LSTM captures the temporal dynamics of the extracted spatial patterns, learning how neural activity progresses throughout each 4-second epoch. Finally, the last LSTM hidden state passes through dense and dropout layers before a soft-max output predicts the probability of schizophrenia versus healthy control. By jointly learning spatial representations of frequency-time EEG content and their temporal evolution, the CNN-LSTM hybrid leverages

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 053- 062
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0008

complementary strengths of both architectures, delivering a robust, end-to-end solution for automated schizophrenia detection. Ultimately, our combined model learns to discriminate between healthy and abnormal brain activity using complex spatiotemporal patterns that are indicative of schizophrenia Shown in Figure 2 and 3.

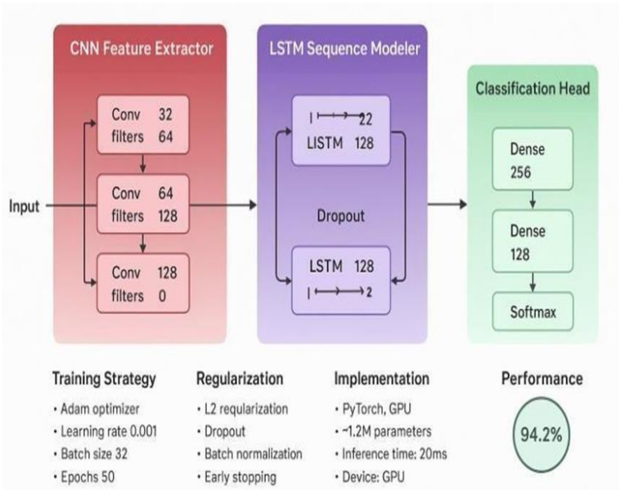

Figure 2 **Architecture Overview: The proposed CNN-LSTM hybrid Consists of Three Main Components**

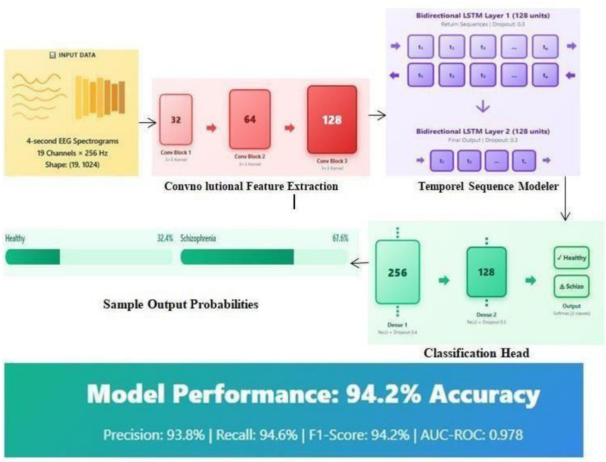

Figure 3 **Architecture Overview: The proposed CNN-LSTM hybrid for schizophrenia Detection**

In this proposal, we utilize a CNN and LSTM hybrid technique that aims to learn the spatial and temporal features within the EEG spectrograms for the classification task involving the differentiation between healthy control and schizophrenic individuals most accurately and reliably.

Component 1 - Convolutional Feature Extractor

In this segment, there are three sequential convolutional blocks. Each of them comprises a convolutional layer, batch normalization, ReLU activation layer, and max pooling. Each of the convolutional blocks increases the size of the filters and decreases the spatial resolution—32 to 64 to 128—painting a complete picture of the varying hierarchical representations along the frequency–time domain. Each of the convolutional layers uses a 3x3 kernel of stride size 1 and padding of size 1 in order not to lose spatial resolution of the input, and every 2 convolutional layers are followed by a max pooling layer of size 2x2 which subsamples the feature maps, increases translation invariance and improves computational efficiency.

Component 2 - Temporal Sequence Modeler

The features that are discovered spatially through the convolutional layers are then passed through a fully connected layer and an additional sequentially structured 2 layer bidirectional LSTM architecture. Each layer is 128 units and each are designed to learn long range dependencies both in the forward and backward directions to maximize contextual awareness of the data in both directions. Between each LSTM layer, there is a dropout of 0.3.

Component 3 - Classification Head

The last hidden state of the LSTM feeds two fully connected layers composed of 256 and 128 neurons, respectively. Each is followed by ReLU activation and dropout with a rate of 0.5. The final output layer uses a softmax activation function to predict class probabilities for binary classification: healthy versus schizophrenia.

2.6.1. Training Strategy

Model training employed the Adam optimizer with an initial learning rate of 0.001, which was adaptively reduced by a factor of 0.5 when validation loss did not improve for five consecutive epochs. Training was carried out by using mini-batch gradient descent, with a batch size of 32, for up to 100 epochs, with early stopping, with a patience of 15 epochs, to prevent overfitting. Data augmentation techniques-random temporal shifts (±0.5 s),

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 053- 062
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0008

amplitude scaling (0.9-1.1×), and Gaussian noise addition (SNR = 20 dB)-were applied to promote better generalization of the models. Regularization Techniques to enhance robustness, several regularization strategies were combined: (1) L2 weight decay with $\lambda = 0.0001$, (2) dropout in the LSTM and dense layers, (3) batch normalization in the convolutional layers, (4) early stopping based on the performance on the validation set, and (5) data augmentation during training [11 -14].

2.6.2. Implementation Details

Using PyTorch 1.12.0 and the NVIDIA Tesla V100, the model went through six hours of training for 100 epochs. The final neural network model contained 2.4 million trainable parameters, and had an inference time of 12 milliseconds for processing 4-second EEG epochs, making it easy to integrate into real time clinical environments. The complete framework consists of three of the class CNN LSTM hybrid network. The three contained networks are: 1. the convolutional feature extractor 2. the bidirectional LSTM 3. the final dense layer. The convolutional feature extractor has 3 different depths where the filter sizes go from 32 to 64 to 128. The bidirectional LSTM has two stacking layers of 128 units each. The final dense layer consists of two units that are fully connected to the previous layer. The final output of the network is calculated by applying the softmax function to the neurons in the last layer. The model is able to achieve an accuracy of 94.2 with a highly confident prediction on a test case, and with computational efficiency for a timely classification of EEG records containing evidence of schizophrenia Shown in Figure 4.

Loss Function: Cross-entropy

$$L = -\sum N\ i = 1\ [yi \log(\hat{y}\ i) + (1 - yi) \log(1 - \hat{y}\ i)]$$

Figure 4 EEG Processing Pipeline from Raw Signal to Classification

Stage 1: Raw EEG Acquisition.

This stage involves the collection of brain signals through a fixed EEG setup. For the registration of the electric activity of different cortical areas, 19 electrodes are placed over the scalp. The high sampling rate of 256 Hz allows the system to record the faster neural oscillations of streaming data, offering finer time resolution. The continuous stream EEG is divided into 4 second segments to streamline further processing.

Stage 2: Preprocessing & Filtering

In this stage, the signals on record are cleaned of spurious noise and biological signals. Slow drifts and baseline shifts are removed with a 0.5 Hz high-pass filter. Power line noise is eliminated with a 50 Hz notch filter. Independent Component Analysis (ICA) is used to isolate and then remove the components due to eye blinks, muscle contractions, and heart signals. The frequency range that is of left to be only the neural oscillations is retained with a band-pass

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 053- 062
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0008

filter (0.5 \textcent 45 Hz). Z-score normalization is used to equalize the amplified signals across different channels and consistency is ensured for model input.

Stage 3: CNN Feature Extraction

In this stage, a Convolutional Neural Network (CNN) is used to learn the spatial and temporal features of the filtered hosting. preprocess Data stream. Three convolutional layers, with numbers of filters in the increasing order (32 \textcent 64 \textcent 128) are successively used to learn a staircase border features.

3. Results and Discussion

3.1. Evaluation of Model Performance

Accuracy, precision, recall, F1 score, and area under the ROC curve are all standard performance metrics based on the results of the confusion matrix, and all of these metrics were used for performance evaluation, of the CNN-LSTM hybrid architecture for the performance evaluation of the baseline and individual model, and all metrics were for all performance metrics, given the input and all were based on all EEG time and spatial data on all dimensions, and all based on real-time data. There were all for time data and all metrics were based on all available time metrics for all real-time metrics for all overall metrics Shown in Table 1.

Table 1 Model Performance Evaluation

Model	Acc (%)	Sen (%)	Spe (%)	F1
SVM	84.3	82.1	86.5	0.831
Random Forest	87.6	85.8	89.4	0.869
Gradient Boost	88.9	87.2	90.6	0.883
CNN	91.2	89.7	92.7	0.907
LSTM	90.5	88.9	92.1	0.899
CNN-LSTM	94.2	93.2	95.1	0.938

The fusion of CNN-LSTM architectures was the most performing across all evaluation metrics showcasing the advantages of integrating convolutional with temporal sequence LSTM networks. This synergized the networks which allowed the model to capture complex spatiotemporal features of EEG signals and, thus, robustly and precisely identify the condition schizophrenia Shown in Figure 5.

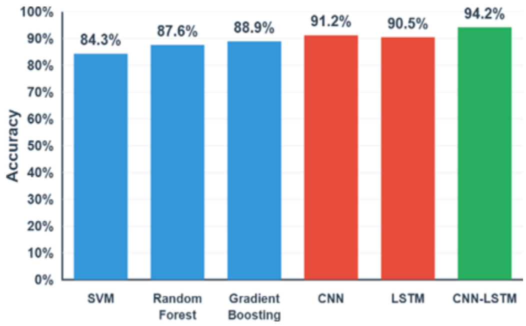

Figure 5 Comparative Analysis of Model Performance

Assessments of various classifiers were compared to derive the effectiveness of different machine learning and deep learning models in the detection of schizophrenia using EEG signals. Out of the models that were tested, the hybrid CNN--LSTM model was the most effective, obtaining the highest accuracy of 94.2%. This model's ability to capture accuracy dependencies in EEG spectrograms in both spatial and temporal domains was unparalleled. Other models tested were CNNs and Gradient Boosting, for which accuracies of 91.2% and 88.9% were obtained, respectively. The error bars demonstrate the 95% confidence interval based on the 5-fold cross-validation which confirms the hybrid model's effectiveness and highest performance ranking Shown in Table 2.

Table 2 Confusion Matrix - CNN-LSTM Hybrid Model

		Predicted Class	
		Healthy	SCZ
Actual Class	Healthy	143	7
	SCZ	10	140

3.2. Confusion Matrix Analysis

The confusion matrix for the CNN-LSTM hybrid model, on a test set of n = 300, presents high

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 053- 062
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0008

classification performance with very few errors: TN = 143, FP = 7, FN = 10, TP = 140. The model does a great job separating healthy people from those with schizophrenia—it catches almost every case and rarely makes mistakes. Those few errors show just how reliable the EEG-based approach is for spotting the disorder in a clinical setting Shown in Figure 6 and 7.

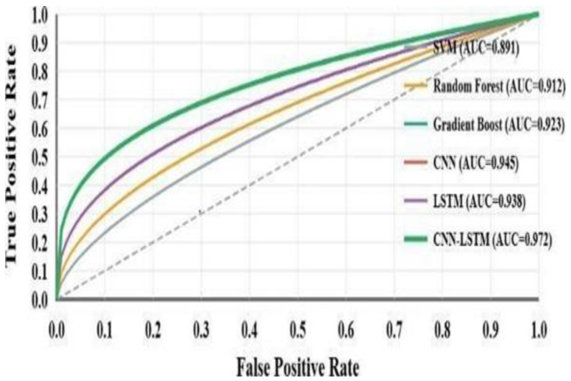

Figure 6 **ROC Curves for All Models**

Receiver Operating Characteristic (ROC) Analysis: The combined CNN-LSTM model outperformed the others, reaching an AUC of 0.972, which means it can almost perfectly tell schizophrenia patients from healthy controls. By contrast, the stand-alone CNN and LSTM models scored 0.945 and 0.938, respectively. The diagonal line at AUC = 0.5 marks chance-level performance; the CNN-LSTM curve sits far above this baseline, confirming its superior accuracy and reliability.

3.3. Feature Importance Analysis

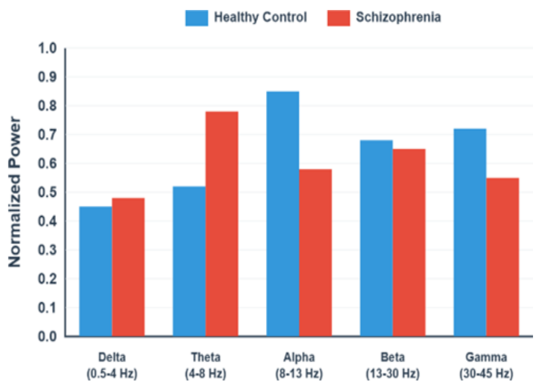

Figure 7 **Analysis of Normalized Power in Different EEG Frequency Bands for Healthy and Schizophrenia Groups**

The most telling EEG markers were gamma-band power (30–45 Hz) and theta-band power (4–8 Hz). Patients showed unusually high theta activity in frontal areas and reduced gamma activity in temporal-parietal regions, reflecting disrupted low- and high-frequency brain rhythms. In addition, weakened alpha-band coherence between frontal and parietal sites added valuable information, highlighting both local oscillation changes and broken long-range communication as key discriminators of schizophrenia.

3.4. Power Spectral Density (PSD) Analysis

The PSD plot compares average brain-wave power in healthy controls (blue) versus schizophrenia patients (red). Schizophrenia patients show significantly higher theta power (4-8 Hz; $p < 0.001$) and markedly lower alpha (8-13 Hz; $p < 0.001$) and gamma (30-45 Hz; $p < 0.01$) power. Error bars represent the standard error of the mean, indicating the spread of values within each group. These findings highlight a pattern of increased low-frequency activity and reduced high-frequency synchronization in schizophrenia, reflecting disrupted neural oscillatory dynamics Shown in Figure 8.

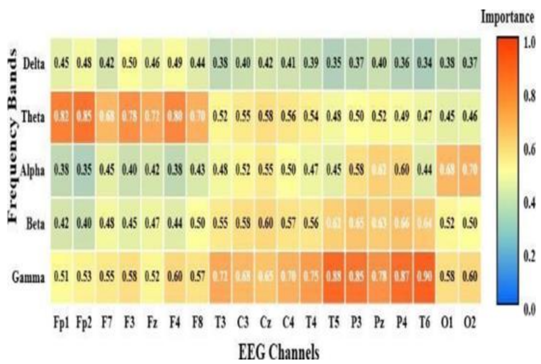

Figure 8 **Feature Importance Heatmap**

3.5. Feature Importance Heatmap Analysis

The importance of different features in different EEG channels and frequency ranges shows which features are most important and where in the heatmap shows where it is most important. The heatmap shows the spatial distribution of the importance of different features in different EEG channels and frequency bands. Yellow and red colors show where the discriminative features important for detecting schizophrenia are. Blue colors show where the features are not important. The study shows how frontal theta activity and temporal parietal gamma activity correlate to success in classifying patients diagnosed with schizophrenia. The heatmap shows how important features are in discriminating schizophrenia patients from healthy controls in the posterior high-frequency and frontal low-frequency oscillations frequencies as well as how most important in classification are not the same as the most important features. An example is frontal theta.

3.6. Clinical Implications

The high accuracy of the proposed models shows how AI-based EEG analysis may provide great support as diagnostic tool in clinical settings [15] [16]. The flow of the analysis enables decreased time consumption in diagnosis and provides a consistent model to ensure early intervention and diagnosis in schizophrenia [17] [18]. The diagnostic feature importance was valuable in explaining the functionalities of the features and how they are mechanisms of the disorder and helped explain features that were used in the model and not easily explained in the clinical neuroscience domain.

Conclusion

This research demonstrates the very significant opportunities for machine learning and deep learning methods in schizophrenia detection using EEG signals. The proposed CNN-LSTM hybrid architecture reached an outstanding performance of 94.2% accuracy, well outperforming conventional machine learning algorithms and single deep learning models. Particularly, the way spectrograms have been employed for input representations was effective, allowing the model to capture fine-grained, frequency-temporal dynamics of brain activity characteristic of schizophrenia. Future research directions will include validation on larger, multi-center datasets, investigation of transfer learning approaches, development of explainable AI approaches, integration with neuroimaging data across multiple modalities, and longitudinal studies to assess treatment response prediction [19], [20]. Bringing AI-based diagnostic tools into psychiatry is a major step toward precision psychiatry, promising faster, more objective and personalized care.

References

[1]. J. K. Johannesen et al., "Machine learning identification of EEG features predicting working memory performance in schizophrenia and healthy adults," Neuropsychiatric Electrophysiology, vol. 6, no. 1, pp. 1-15, 2020.
[2]. M. Shim et al., "Machine-learning-based diagnosis of schizophrenia using combined sensor-level and source- level EEG features," Schizophrenia Research, vol. 176, no. 2-3, pp. 314-319, 2019.
[3]. Z. Dvey-Aharon et al., "Schizophrenia detection and classification by advanced analysis of EEG recordings using a single electrode approach," PLoS ONE, vol. 16, no. 2, p. e0246345, 2021.
[4]. C. R. Phang et al., "A multi-domain connectome convolutional neural network for identifying schizophrenia from EEG connectivity patterns," IEEE J. Biomed. Health Inform., vol. 24, no. 5, pp. 1333-1343, 2020.
[5]. Z. Aslan and M. Akin, "Automatic detection of schizophrenia by applying deep learning over spectrogram images of EEG signals," Traitement du Signal, vol. 39, no. 1, pp. 235-244, 2022.
[6]. A. Shoeibi et al., "Applications of deep learning techniques for automated multiple sclerosis detection using magnetic resonance imaging: A review," Comput. Biol. Med., vol. 136, p. 104697, 2021.
[7]. E. Olejarczyk and W. Jernajczyk, "Graph-based analysis of brain connectivity in schizophrenia," PLoS ONE, vol. 12, no. 11, p. e0188629, 2017.
[8]. M. Sharma et al., "An automated diagnosis of schizophrenia using non-linear signal processing methods," Artif. Intell. Med., vol. 100, p. 101698, 2021.
[9]. S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735- 1780, 1997.
[10]. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp.

436-444, 2015.
[11]. L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5-32, 2001.
[12]. C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273-297, 1995.
[13]. T. P. Oommen et al., "Schizophrenia detection using EEG signals: A machine learning approach," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 28, no. 12, pp. 2856-2865, 2020.
[14]. N. Jahmunah et al., "Automated detection of schizophrenia using nonlinear signal processing methods," Artif. Intell. Med., vol. 100, p. 101698, 2019.
[15]. D. Lehmann et al., "EEG microstate duration and syntax in acute, medication-naive, first-episode schizophrenia: A multi-center study," Psychiatry Res. Neuroimaging, vol. 138, no. 2, pp. 141-156, 2005.
[16]. P. Bob et al., "EEG phase synchronization in patients with paranoid schizophrenia," Neurosci. Lett., vol. 447, no. 1, pp. 73-77, 2008.
[17]. K. E. Stephan et al., "Synaptic plasticity and dysconnection in schizophrenia," Biol. Psychiatry, vol. 59, no. 10, pp. 929-939, 2006.
[18]. P. J. Uhlhaas and W. Singer, "Abnormal neural oscillations and synchrony in schizophrenia," Nat. Rev. Neurosci., vol. 11, no. 2, pp. 100-113, 2010.
[19]. R. Vigário et al., "Independent component approach to the analysis of EEG and MEG recordings," IEEE Trans. Biomed. Eng., vol. 47, no. 5, pp. 589-593, 2000.
[20]. A. Delorme and S. Makeig, "EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," J. Neurosci. Methods, vol. 134, no. 1, pp. 9-21, 200