Vol.12-1(2025) 67–84

International Journal of Computers and Information

(IJCI)

Available online at https://ijci.journals.ekb.eg/

Unravelling Schizophrenia: An Attention-Based Deep Learning Approach for Detection Using EEG Signals

Mohamed A. Elgendy [a,*], Sherif Eletriby [a], Arabi E. Keshk [a], Mohamed Sakr [a]

[a] Computer Science Department, Computers and Information Faculty, Menoufia University, Menoufia 32511, Egypt
[*] Muhammad.Elgendi@ci.menofia.edu.eg

Abstract

Schizophrenia (SZ) affects over 20 million people globally, with many patients diagnosed too late to receive appropriate treatment. Current diagnostic methods are time-consuming, requiring skilled psychiatrists, underscoring the need for more efficient approaches. This work explores using attention-based deep learning models to classify EEG signals, a non-invasive and cost-effective method, into healthy individuals and SZ patients. The proposed attention-GRU model incorporates convolutional neural networks (CNNs) for spatial feature extraction, gated recurrent units (GRUs) for sequence interpretation, and attention layers to highlight the most relevant inputs. Unlike previous works that require time-consuming manual feature extraction, our end-to-end model learns directly from EEG data, reducing preprocessing steps and enhancing the potential for real-time clinical application. Experimental results show a significant improvement in SZ detection, reaching a competitive 98.52% accuracy on an open-source EEG dataset, overcoming the accuracy reported in previous studies. This work highlights the potential of advanced deep learning models in improving the accuracy and efficiency of SZ diagnosis, addressing standardization challenges, and paving the way for more reliable diagnostic tools in psychiatric care. Our results indicate that, with further validation, AI-driven assessments can support early intervention and broader access to treatment for mental disorders.

Keywords: schizophrenia (SZ) detection; electroencephalogram (EEG) signals; attention; long short-term memory (LSTM); gated recurrent unit (GRU); convolutional neural network (CNN).

1. Introduction

Schizophrenia (SZ) is an acute neurological disorder affecting over 20 million people worldwide, approximately 1% of the global population [1, 2]. According to the World Health Organization (WHO), it can be cured through a long process of treatment [1], which exhibits a long-term economic burden on families [3]. Approximately 69% of SZ patients do not receive appropriate therapy and care [4] due to incorrect identification and restricted access to treatment. Environmental factors and genetics play a part in the onset of SZ. Patients with SZ die at three times the rate of the general population due to metabolic, circulatory, and infectious illnesses [5]. Skilled psychiatrists diagnose SZ patients through behavioral observation and standardized assessments such as those outlined in DSM-5, ICD-11, and CCMD-3. In [6], At least one of the delusions, hallucinations, or disorganized speech symptoms must be present for psychiatrists to diagnose patients with SZ. Manual observation and interview methods could be more reliable, but they are time-consuming, prone to human error, and have the possibility of false positive detection. As such, developing a robust automatic method for identifying SZ is crucial.

Researchers have recently addressed the automated detection of SZ using a wide range of techniques and different types of feature-extraction methods. However, because of the lack of standardization, Early detection of SZ has been hindered.

Profound disruptions in thinking, perception, emotions, language, sense of self, and behavior characterize SZ disorder. Individuals with SZ often experience hallucinations, which are false perceptions such as hearing voices that are not present, and delusions, which are strong beliefs in things that are not real. Additionally, they may exhibit disorganized speech, making it challenging to communicate effectively and significantly impairing their ability to function socially and occupationally [6].

SZ typically emerges in late adolescence or early adulthood, with men often exhibiting symptoms earlier than women. The exact causes of SZ are complex and multifaceted, involving a combination of genetic, neurobiological, and environmental factors. Genetic predisposition plays a significant role, as individuals with a family history of SZ have a higher risk of developing the disorder. Neurobiological research has identified structural and functional abnormalities in the brains of those with SZ, particularly in cognition and emotion regulation regions. Dysregulation of neurotransmitter systems, especially dopamine and glutamate, is also implicated in the disorder's pathophysiology [7].

Environmental factors contributing to SZ include prenatal exposures to infections, malnutrition, and stress, as well as complications during birth. Psychosocial stressors such as trauma, urban upbringing, and cannabis use during adolescence have been associated with an increased risk of developing the disorder [8].

One of the Biomarkers for early detection of SZ is electroencephalogram signals [9, 10]. Electroencephalogram (EEG) signals offer intensive data over alternative methods for identifying SZ [11]. EEG signal acquisition is non-invasive, cheap, and radioactive-free. Therefore, it has recently been extensively used to identify brain abnormalities.

While EEG readings are continuous and reflect the brain's electrical activity in real-time, event-related potentials (ERPs) are discrete chunks of EEG data tied to specific events and averaged across several trials. During vocalization, the auditory cortex in the temporal lobe shows reduced activity compared to when humans are passively listening. This reduction is due to the corollary discharge process, which helps the brain differentiate between sounds humans produce and external sounds. By dampening the response to self-generated sounds, the brain avoids being 'surprised' by expected noises and can concentrate better on unexpected external sounds [12].

The N100, a negative-going potential that peaks 100 ms post-stimulus, is reduced in first-degree relatives of SZ patients and both not-medicated and treated SZ patients. Another positive potential, the P200, reaches its maximum 200 ms following stimulation and follows the N100. An empirical way exists to tell N100 and P200 apart, even if they covary. Decreased N100 and P200 levels have been associated with SZ [13].

In this work, EEG signals are captured from 64 electrodes. Two deep learning models, attention-LSTM and attention-LSTM stacked models, are trained to recognize whether a subject has SZ. Since EEG signals have high temporal resolution, using an automatic feature extraction method will help simplify the preprocessing phase of such signals. Our models have reached state-of-the-art results in SZ identification using EEG signals, simplifying the identification of SZ procedures and helping as second advice to medical professionals.

The contributions of this work are:

- Provide an end-to-end framework to extract features from EEG signals automatically.
- Address the problem of understanding the most prominent features of potentially long sequences of EEG data by using long short-term memory (LSTM) units and gated recurrent units (GRU) with attention.
- Present cost-effective and accurate models to help detect SZ in EEG signals, which gives practitioners a second opinion.

The structure of this work is as follows: after being introduced to the problem, related works are presented in section 2, and the proposed method in section 3. Section 4 presents the experimental results, the dissection in section 5, and the conclusion in section 6.

2. Related Work

Researchers extensively examined the use of EEG signals for identifying schizophrenia (SZ) since these signals are non-invasive, non-radioactive, and portable. Wavelet analysis, autoregressive models, spectral

analysis of alpha, beta, theta, and delta waves; self-organizing competitive maps; entropy; Lempel-Ziv complexity (LZC); fast Fourier transform (FFT); independent component analysis (ICA); non-linear methods based on HFD; and band power are among the techniques that have been employed to extract features in EEG signals for SZ detection. The following methods have been used to classify EEG signals: multilayer perceptron (MLP) classifier, adaptive weighted distance nearest neighbor algorithm, leave-one-subject-out cross-validations, support vector machine (SVM) classifiers, random forest (RF) classifiers, convolutional neural network (CNN), and adaptive AdaBoost classifiers. Due to the absence of standardized techniques and procedures, previous attempts to diagnose SZ using EEG data have produced unsatisfactory outcomes.

In [14], it uses time-frequency representation (TFR) techniques to obtain the temporal details of nonstationary EEG signals, and the images obtained by different TFR methods are classified using convolutional neural networks (CNNs). This work used the open-sourced dataset from [15] and extracted spectrograms, scalograms, and smoothed pseudo-Wigner-Ville distribution (SPWVD) in combination with CNN. This work's best-obtained accuracy is 93.36% with the SPWVD-based TFR with a CNN classifier. However, this work transforms EEG signals into images without dealing with the intensive nature of the temporal resolution of EEG.

Also, [16] transformed the dataset from [15] to scalograms using the CWT technique with Morlet wavelet and used CNN with spatial and channel-wise attention, achieving an accuracy of 95% on [15] and 99% on [17]. Random forest classifier is used in [18] and [19] for SZ identification and achieved 81% accuracy on the classification of EEG signals from [15] and ERP analysis and found that the accuracy could be improved by feeding the model with certain features extracted from EEG signals. But only nine channels of the ERP averaged data were used, namely electrodes CP4, CP3, C4, C3, FC4, FC3, Cz, FCz, and Fz, in [19], used the same electrodes and reported an accuracy of 96.4% and sensitivity of 92.8%.

Artificial neural network is trained in [20] on Sample Entropy and Kolmogorov Complexity features extracted from electrodes CP4, CP3, C4, C3, FC4, FC3, Cz, FCz, and Fz in [15] dataset, achieved 88.75% accuracy, precision of 81.2%, specificity of 93% and sensitivity of 88%. Also, this work found that training the same model on four electrodes, Cz, FCz, C3, and FC3, can improve the results and achieve an accuracy of 91.25%, precision of 96%, specificity of 93.2%, and sensitivity of 90.8%. In the study of [21], the HC group and SZ patients were classified using Artificial Neural Networks (ANN), K-nearest neighbor (KNN), and Support Vector Machines (SVM). The researchers in [15] relied on ERP analysis features, and the ANN classifier obtained an accuracy rate of 93.9%. An ensemble bagged tree classifier was used in [22], producing 93.21% accuracy for classifying SZ; intrinsic mode functions (IMFs) achieved an overall accuracy of 89.59%.

Time series data from [15] were converted into images in [23], using Gramian Angular Field (GAF) and Recurrence Plot (RP) algorithms and a CNN based on VGGNet. The classification accuracy of 94.5% was achieved when classifying Recurrence Plot (RP) images, and 96.3% was achieved on GAF images. A flexible least-squares support vector machine (F-LSSVM) classifier achieved an accuracy of 91.39% and an F-1 score of 93.06%; the Fisher score, in conjunction with a flexible tunable Q wavelet transform (F-TQWT) to decompose EEG signals, was used in [24]. A spatial-temporal residual graph convolutional neural network was used in [25] detecting SZ from a publicly available dataset [26, 27]. EEG signals were treated as a graph to get connections between different channels, and the network achieved an accuracy of 96.32%.

Two different EEG datasets [17, 28] with 19 and 16 channels, respectively, were used in [29]; both datasets were used to validate a 3D convolutional information fusion network with node topology regression network, evaluated at accuracy of 96.04% and 92.67%. Graph attention networks (GAT) are used in [30] with a bilinear convolution neural network on resting-state functional MRI data. GATs obtain the node-level features, ignoring the spatial information that proved to be essential for SZ detection. Accuracy of 90.48%, sensitivity of 92%, and specificity of 89.29% were achieved. The comparison in Table 1 shows key differences among the related works in terms of methodology and highlights the strengths and weaknesses of each work.

In summary, the existing literature on the diagnosis of SZ using EEG data addressed the problem in one of the following three ways: first, using EEG sequential data alongside feature extraction algorithms and feeding that to a deep learning model or a traditional machine learning algorithm; second, transforming the EEG data to images; or third, using discrete event-related potential (ERP) data.

A gap exists for more advanced, end-to-end comprehensive deep learning architectures; traditional machine learning algorithms frequently need help dealing with the inherent noise, complex temporal dynamics, and long

sequences in EEG data. These characteristics made them unsuitable for real-time or clinical applications wherein accuracy and efficiency are paramount. Advancements in deep learning architectures, such as recurrent neural networks (RNNs), long short-term memory networks (LSTMs), gated recurrent unit (GRU), and more recently, transformers, have facilitated the development of models capable of autonomously recognizing patterns in EEG data, eliminating the necessity for manual feature extraction.

In this work, hybrid models that integrate the advantages of deep learning without conventional feature extraction methods are trained on EEG data, enabling early-stage automated detection of SZ.

Table 1: Comparison of Related Works on Schizophrenia Diagnosis

Ref	Pros	Cons
[14]	- High accuracy (93.36%) using SPWVD and CNN. Efficient combination of time-frequency analysis and deep learning.	- Requires empirical parameter selection. Increased memory and computational complexity. Limited generalizability.
[16]	- High accuracy uses depth-wise separable convolutions for computational efficiency. Incorporates attention mechanisms to focus on relevant features.	- Model complexity due to the use of attention mechanisms and their placement. Significant computational resource requirements.
[18]	- Simple, interpretable Random Forest model. Highlights EEG-derived features. Low error rate (18.11%).	- Small dataset size. High false positive rate. Lack of comparison with deep learning models. Uses ERP features.
[19]	- High accuracy (96.4%) with Random Forest. Uses additional sensors on both hemispheres for comprehensive EEG analysis.	- Limited generalizability of the model. Focuses on a specific ERP N100 component.
[20]	- High accuracy (91.25%) with a small number of electrodes. Utilizes Kolmogorov Complexity and Sample Entropy effectively. Robust classification using Artificial Neural Network.	- Limited in real-time application. Focus on a subset of electrode analysis. Lack of comparison with deep learning models.
[21]	- High accuracy (93.9%) using ERP features with ANN. Uses SMOTE for dataset balancing.	- Relies solely on ERP signals. Uses only a subset of EEG channels. Lack of comparison with deep learning models.
[22]	- Accuracy of (89.59%) using Ensemble Bagged Tree. Comprehensive evaluation of multiple classifiers.	- Limited dataset size may impact generalizability. The empirical mode decomposition (EMD) technique requires significant computational requirements.
[23]	- High accuracy (93.2%) using GAF and VGGNet. Uses time series image conversion techniques. Practical deep learning application for complex EEG patterns.	- High computational complexity due to image conversion. Lack of comparison with other deep learning models.
[24]	- High accuracy (91.39%) with F-TQWT and F-LSSVM. Fisher score for channel selection reduces computational complexity. Uses Grey Wolf Optimization for parameter tuning.	- Small sample size. High computational complexity. It is not validated in real-time settings.
[25]	- High accuracy (96.32%) with STRGCN. Combines spatial and temporal features effectively. Demonstrates strong generalization capabilities on two datasets.	- High computational complexity. Focus on offline analysis without clinical validation. Two small datasets may affect model generalizability.
[29]	- High accuracy (99.46% on Dataset 1 and 98.06% on Dataset 2). Novel 3DCNN-based model enhances interpretability. Uses knowledge distillation to improve feature diversity.	- High computational complexity due to 3DCNN and knowledge distillation techniques. Lack of comparison with other deep learning models.
[30]	- High accuracy (90.48%). Integrates multiple correlation measures and graph topologies. Captures complex brain network features with MGAT and BC.	- High computational complexity due to multi-graph learning and bilinear convolution. Lack of comparison with deep learning models for recognizing temporal features.

3. Proposed methodology

This section describes the EEG data used and presents our pre-processing steps. Afterward, two model architectures are proposed, and a brief introduction to each block is given. Figure 1 presents the proposed methodology for automatic SZ detection.

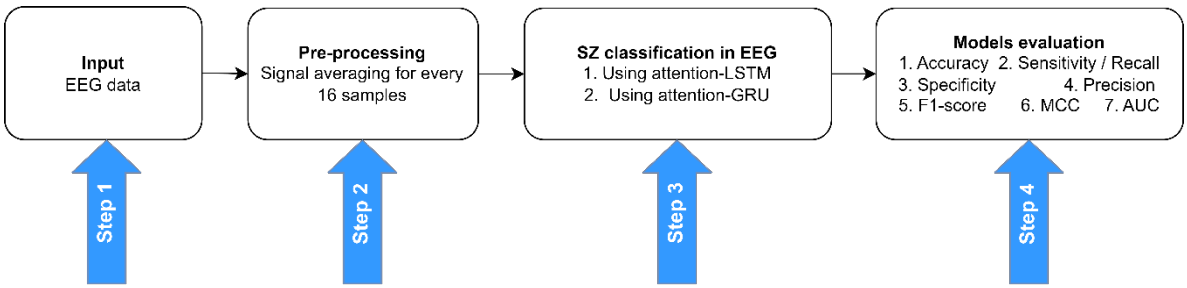


Figure 1: The proposed methodology for SZ detection.

3.1. *Dataset*

EEG data were initially collected in this experiment [31]. EEG data from 10 healthy subjects and 13 patients with SZ was collected. Later, 22 healthy controls and 36 patients were added to the data. Data was collected from participants from three different conditions. Namely, the first condition was when the subject pressed a button to generate a tone immediately, the second condition was when the subject passively listened to the tone, and the last was when the subject pressed the button and didn't get a tone as feedback. The following phases cleaned the data that were acquired. A re-referencing of the electrodes of the averaged ear lobes was conducted. The baseline was corrected into the range -100ms to 0ms. Muscle and high-frequency noise artifacts were removed. Outlier components from a spatial independent components' analysis were removed as defined in [32]. Outlier channels, as described in [32], were interpolated within single trials and in the continuous EEG data. A high-pass filter was used to pass frequencies greater than 0.1 Hz. Data was chopped into epochs of 1.5 seconds before and after task events (each epoch was 3 seconds in total). The dataset is publicly available on Kaggle [15]. Figure 2 shows a sample trial for the three different conditions of both healthy control subjects and SZ patients.

3.2. *Preprocessing*

The dataset used in this work has undergone a thorough preprocessing pipeline to ensure the integrity and dependability of the input features for further analysis. Initially, signal averaging was employed, in which the data obtained from each electrode was averaged across every 16 consecutive samples of EEG recordings. Minimizing noise and resolving the inherent unpredictability in EEG recordings is a crucial step in the preprocessing pipeline since it improves the quality of the input features [33, 34].

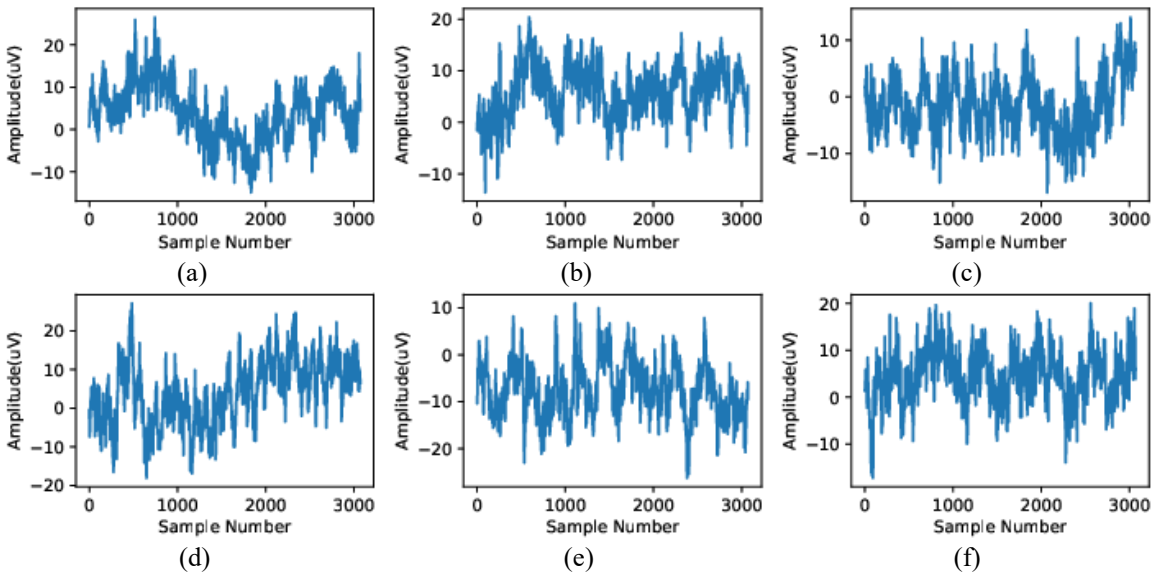


Figure 2: comparison of SZ patient and healthy control trial
a) SZ condition 1, b) SZ condition 2, c) SZ condition 3, d) HC condition 1, e) HC condition 2, f) HC condition 3

Furthermore, one-hot was implemented encoding on the condition feature, transforming categorical data into a format appropriate for training the models. The dataset was shuffled to prepare the data for model evaluation, and a stratified splitting was conducted, assigning 80% of the data for training and 20% for testing. This maintains class distribution consistency throughout the training and test sets.

Applying feature scaling is crucial before training deep learning models as it guarantees that all features contribute equally to the models' accurate predictions. Standardizing features to a uniform range enhances the efficiency of several algorithms, especially those that are responsive to the size of the features, such as gradient descent-based approaches and distance-based algorithms. The maximum of the absolute values rescales input features, guaranteeing uniform and equivalent input data for the models.

After the preprocessing phase, the output data shape is a feature matrix consisting of 576 time steps and 67 features, including 64 electrodes and the three values for the condition feature stored as one-hot. This preprocessing pipeline is specifically developed to clean the data for input into the proposed models, enabling precise and resilient performance throughout the training and evaluation processes.

3.3. *Proposed model architectures*

Two slightly different architectures are proposed, namely, attention-LSTM and attention-GRU models. Both are stacked of convolutional neural network (CNN), either a long short-term memory (LSTM) layer or a gated recurrent units layer, then an attention layer, and finally, a fully connected layer, as presented in Figures 3 and 4.

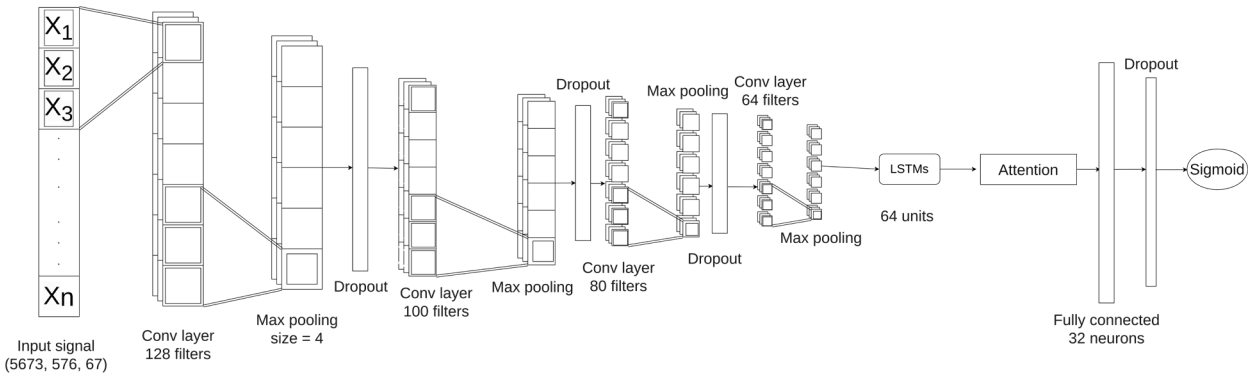


Figure 3: Attention-LSTM proposed model architecture

The attention-LSTM and attention-GRU comprise several stacked layers, each with a specific function in processing and transforming the incoming data. These layers are intended to effectively extract features, capture complex patterns, and provide the required output. Table 2 details each layer's summary, configuration, and level in architectures and its precise hypermeter, including the number of units, filter sizes, and activation functions.

EEG signals are fed to the input layer. Then, convolutional layers play an important role in feature extraction since they apply filters to input data to capture spatial hierarchy and patterns. Each convolutional layer is followed by an activation function, ReLU, which adds nonlinearity. Max pooling layers minimize the dimensionality of feature maps produced by convolutional layers, lowering computational costs and the likelihood of overfitting.

Attention units improve models' performance in sequence-based problems by adding trainable parameters focusing on the input's most essential features. The fully connected layers are used as the final layers in architectures, transforming flattened feature maps into a one-dimensional vector that may be translated to output classes. The final prediction is generated by the output layer in the last fully connected layers of the models. This layer often contains a sigmoid activation function used for binary classification tasks.

3.3.1. Convolutional neural network (CNN)

Convolutional neural networks (CNN) were created explicitly for pixel processing and are employed in image recognition. In CNN architecture, there are six distinct components. The convolutional layer is a linear operation that performs multiplication between a two-dimensional matrix of weights and the two-dimensional image input. In image processing, the weights matrix is called a kernel or filter.

Our models use the CNN block to extract spatial information from data. A convolutional layer is a linear operation that multiplies a two-dimensional matrix of weights with the two-dimensional image input. In image processing, the weights matrix is called a kernel or filter.

Four one-dimensional convolution layers are used: the first has 128 kernels, the second has 100 kernels, the third has 80 kernels, and the fourth has 64 kernels. A kernel size of 10 was used alongside a stride of one. ReLU is used as an activation function after the four layers. Batch normalization and a max-pooling layer are used with a pool size of four elements and a stride of one. A dropout layer was used after each layer with a dropping ratio of 0.2.

Table 2: Comparison of Models Attention-LSTM (1) and Attention-GRU (2) with Hyperparameters.

Layer (type)	Model 1 Output Shape	Model 2 Output Shape	Hyperparameters	Param #
Conv1D	(567, 128)	(567, 128)	kernel size=10, strides=1	85,888
Activation (ReLU)	(567, 128)	(567, 128)	-	0
Batch Normalization	(567, 128)	(567, 128)	-	512
Max Pooling1D	(564, 128)	(564, 128)	pool size=4, strides=1	0
Dropout	(564, 128)	(564, 128)	rate=0.2	0
Conv1D	(555, 100)	(555, 100)	kernel size=10, strides=1	128,100
Activation (ReLU)	(555, 100)	(555, 100)	-	0
Batch Normalization	(555, 100)	(555, 100)	-	400
Max Pooling1D	(552, 100)	(552, 100)	pool size=4, strides=1	0
Dropout	(552, 100)	(552, 100)	rate=0.2	0
Conv1D	(543, 80)	(543, 80)	kernel size=10, strides=1	80,080
Activation (ReLU)	(543, 80)	(543, 80)	-	0
Batch Normalization	(543, 80)	(543, 80)	-	320
Max Pooling1D	(540, 80)	(540, 80)	pool size=4, strides=1	0
Dropout	(540, 80)	(540, 80)	rate=0.2	0
Conv1D	(531, 64)	(531, 64)	kernel size=10, strides=1	51,264
Activation (ReLU)	(531, 64)	(531, 64)	-	0
Batch Normalization	(531, 64)	(531, 64)	-	256
Max Pooling1D	(528, 64)	(528, 64)	pool size=4, strides=1	0
Dropout	(528, 64)	(528, 64)	rate=0.2	0
LSTM (Model 1)	(528, 64)	-	kernel regularizer=l2(0.000001), recurrent regularizer=l2(0.000001)	33,024
GRU (Model 2)	-	(528, 64)	kernel regularizer=l2(0.000001), recurrent regularizer=l2(0.000001)	24,960
Attention	(64)	(64)	-	4,224
Dropout	(64)	(64)	rate=0.2	0
Flatten	(64)	(64)	-	0
Dense	(32)	(32)	activation='relu'	2,080
Dropout	(32)	(32)	rate=0.2	0
Dense	(1)	(1)	activation='sigmoid'	33
Total Params	386,181 (Model 1)		378,117 (Model 2)	-
Trainable Params	385,437 (Model 1)		377,373 (Model 2)	-
Non-trainable Params	744 (Model 1)		744 (Model 2)	-

3.3.2. Long short-term memory (LSTM)

Long short-term memory (LSTM) networks are a specific category of recurrent neural network (RNN) architectures specifically developed to recognize and memorize long-term relationships within sequential input efficiently. Unlike conventional recurrent neural networks (RNNs), long short-term memory (LSTM) models overcome the issue of inflating gradients during training. This characteristic renders LSTMs highly efficient for tasks that include lengthy sequences. A Long short-term memory (LSTM) cell consists of three primary gates: the forget gate, the input gate, and the output gate. The gates above regulate the transmission of information inside the cell, enabling the network to keep or eliminate information as required selectively.

The attention-LSTM model uses a single layer of 64 long-short-term memory units. Weights are regularized with L2 penalties, which multiply weights by 10^{-5} in the cost function to prevent high values from being assigned as a weight to an input vector. Out of 386,181 parameters, 385,437 parameters are tuned in the training process.

3.3.3. *Gated recurrent units (GRU)*

Gated recurrent units (GRU) are commonly used in sequence learning tasks and help resolve the vanishing gradient problem, which is troublesome in deep-layered recurrent neural networks (RNN). They were introduced as the supreme successor of long short-term memory (LSTM) units; they had a simpler construction with only reset and update gates. The sigmoid function was used in both gates, followed by a tanh function. Both LSTMs and GRUs are used for processing sequences. GRU uses less memory, fewer gates and operations, and trains faster than LSTM.

The attention-GRU model uses a single layer of 64 gated recurrent units. Weights are regularized with L2 penalties, which multiply weights by 10^{-5} in the cost function to prevent high values from being assigned as a weight to an input vector. Out of 378,117 parameters, 377,373 parameters are tuned in the training process.

3.3.4. *Attention operation*

Attention operation mimics cognitive attention in our brains. It's known for its outstanding performance in machine translation systems. It was first presented for machine translation problems in [35].

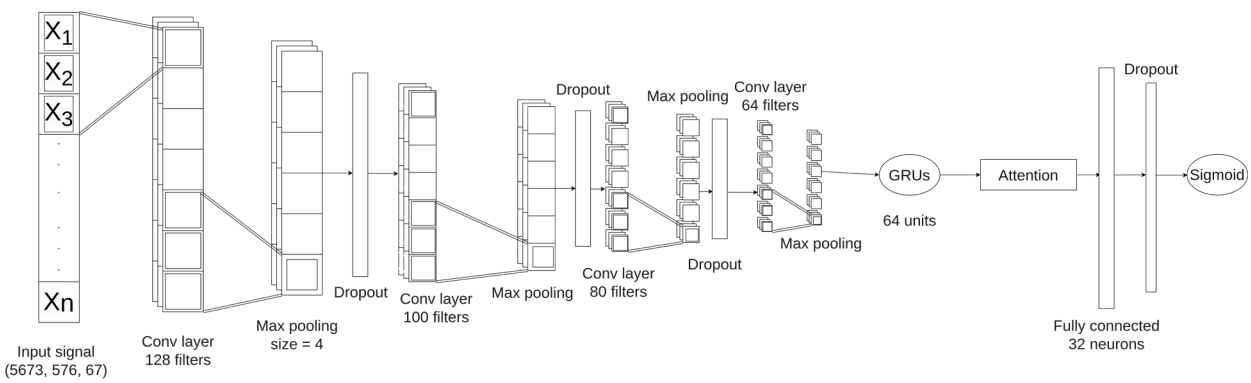


Figure 4: Attention-GRU proposed model architecture

Humans tend to focus on some words more than others during reading, and some words are more informative in determining the category/context of a text than others. Including attention in the proposed models outperforms other models concerned with the sequential nature of the data without giving the importance of a feature any weight. Attention is used in [36] for document classification problems and in [37] for machine translation problems.

Attention is implemented in a model by introducing a randomly initialized new vector called context vector(u) with weight matrix (W) and bias vector (b). The model is then tuned using the Backpropagation algorithm and the propagation of the errors.

The attention score is computed the same as [36] by equations 2, 3, and 4, input matrix X from previous layers is fed to the attention block, A MLP layer is used to get u_{it} which is considered as the hidden representation of x_{it}, then the importance of x_{it} is measured as the similarity of u_{it} with a channel level context vector u_c, and the importance weight α_{it} is normalized through SoftMax function 1, after that the output is computed by multiplying the attention weight by the input x_{it}.

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \quad for\ i = 1,2,\dots,K \tag{1}$$

$$u_{it} = \tanh(W_c x_{it} + b_i) \tag{2}$$

$$\alpha_{it} = \frac{\exp(u_{it}^{\top} u_c)}{\sum_t \exp(u_{it}^{\top} u_c)} \tag{3}$$

$$s_i = \sum_t \alpha_{it}\, x_{it} \tag{4}$$

With LSTMs and GRUs, sequential patterns are recognized, but to improve our ability to assign a higher weight to more critical channels. Stacked with attention scores, our proposed models can grasp the temporary nature of data and assign more weight to discriminative channels.

4. Experimental results

4.1. *Evaluation metrics*

Deep learning models for detecting schizophrenia are evaluated using various performance metrics. Accuracy quantifies the overall correctness of the model's predictions, reflecting the rate at which the model correctly classifies data. The recall (sensitivity) metric evaluates the model's ability to detect individuals with SZ accurately. Specificity evaluates the model's capacity to correctly identify individuals without SZ. Precision refers to the proportion of accurately detected SZ cases among all instances recognized as SZ by the model, which is crucial for minimizing false positives in clinical settings. The Matthews Correlation Coefficient (MCC) is a balanced metric that considers all potential outcomes in the classification process, making it particularly valuable when working with imbalanced datasets. Finally, the F1 Score combines precision and recall into a single measure, providing a comprehensive assessment of the model's effectiveness, especially in scenarios where accurate identification and reduction of false positives are critical. Together, these metrics [5,6,7,8,9,10] ensure the accuracy and reliability of the deep learning model for detecting SZ in a clinical setting.

$$\text{Accuracy} = \frac{T_p + T_n}{T_p + F_p + T_n + F_n} \tag{5}$$

$$\text{Recall} = \text{Sensitivity} = \frac{T_p}{T_p + F_n} \tag{6}$$

$$\text{Specificity} = \frac{T_n}{T_n + F_p} \tag{7}$$

$$\text{Matthews correlation coefficient} = \frac{T_p * T_n - F_p * F_n}{\sqrt{(T_p + F_p)(T_n + F_n)(T_n + F_p)(T_p + F_n)}} \tag{8}$$

$$\text{Precision} = \frac{T_p}{T_p + F_p} \tag{9}$$

$$F-1 \quad \text{Score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{10}$$

Where T_p (True Positive) refers to the number of instances correctly identified as positive, while T_n (True Negative) represents the instances correctly identified as negative, F_p (False Positive) indicates the instances incorrectly labeled as positive, and F_n (False Negative) refers to those that were missed, being wrongly classified as negative.

4.2. *Experimental setup*

In this experiment, multiple models are trained; the training process involves setting up the environment, loading and preprocessing the data, defining the model, and running the training and evaluation phases. Performance metrics such as accuracy, precision, recall, and f1-score are monitored and recorded throughout the experiment. All models are written with the TensorFlow framework version 2.16 in Python on a Nvidia T4

machine with 15 GB of GPU memory and an Intel Xeon CPU with two vCPUs and 13GB of RAM. A batch size of 32 and a maximum of 200 epochs with early stopping are configured through the experiment.

4.3. *Results*

EEG dataset from basic sensory task in Schizophrenia [15] available on Kaggle is used to evaluate the proposed models, EEG data from 32 healthy subjects and 47 SZ patients in trials of three conditions. Conventional methods typically require extensive signal processing and manual feature extraction. Many works have addressed this problem using signal analysis techniques.

In this work, signal averaging for the preprocessing phase is used to keep consistent features of trail samples and remove features that vary across the trail samples. End-to-end deep learning models are presented to identify SZ from EEG data without manual feature extraction. Different deep learning architectures were investigated to solve EEG data's intensive temporal resolution and noisy nature.

The best-performing model consists of four stacked layers of CNNs for automatic spatial feature extraction, GRUs for comprehending the sequential features of EEG signals, an attention layer to focus on the essential temporal features, and a fully connected layer in the end. A deep-learning model is trained to classify EEG signals and detect whether they belong to SZ patients. Challenges associated with EEG data have been overcome, and an automatic end-to-end solution is achieved for this classification.

4.4. *Long short-term memory (LSTM) vs. gated recurrent units (GRU)*

Long short-term memory (LSTM) networks and gated recurrent unit (GRU) networks are two distinct categories of recurrent neural networks (RNNs) that are commonly employed in the processing of sequential data. Although GRUs are generally more straightforward and exhibit comparable performance to LSTMs, some scenarios exist in which LSTMs may surpass GRUs in terms of performance. Both LSTMs and GRUs are specialized in mitigating the issue of vanishing gradients commonly encountered in conventional RNNs. However, they employ distinct internal methods and exhibit varying levels of complexity in doing so.

4.4.1. *Complexity of Architecture*

LSTMs have a higher level of architectural complexity than GRUs. The LSTM cell comprises three main distinct gates: the forget gate, the input gate, and the output gate. Equations [11,12,13,14,15,16] are computed. These gates play a crucial role in regulating the transmission of information and facilitating the network's ability to either keep or forget that knowledge as time progresses. In contrast, gated recurrent units (GRUs) streamline this architecture by consolidating the forget and input gates into a solitary update gate and deleting the output gate, enhancing the computational efficiency of GRUs. Equations of [17,18,19,20] are computed inside GRU.

- LSTM equations are listed as the following:

1. **Forget Gate** (f_t):

$$f_t = \sigma\big(W_f \cdot [h_{t-1}, x_t] + b_f\big) \tag{11}$$

Where σ is the sigmoid function, W_f is the weight matrix for the forget gate, h_{t-1} is the previous hidden state, x_t is input at the current time step, and b_f is the bias for the forget gate.

2. **Input Gate** (i_t):

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{12}$$

Where W_i is the weight matrix for the input gate, and b_i is the bias for the input gate

3. **Cell State Candidate** ($\widetilde{C_t}$):

$$\widetilde{C_t} = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (13)$$

Where tanh is the hyperbolic tangent function, W_C is the weight matrix for the cell state candidate, and b_C is the bias for the cell state candidate

4. **Cell State** (C_t):

$$C_t = f_t \odot C_{t-1} + i_t \odot \widetilde{C_t} \quad (14)$$

Where $\odot$ is element-wise multiplication, and C_{t-1} is the previous cell state

5. **Output Gate** (o_t):

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (15)$$

Where W_o is the weight matrix for the output gate, and b_o is the bias for the output gate

6. **Hidden State** (h_t):

$$h_t = o_t \odot \tanh(C_t) \quad (16)$$

Where tanh is the hyperbolic tangent function, and $\odot$ is element-wise multiplication

- GRU equations are the following:

1. **Update Gate** (z_t):

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \quad (17)$$

Where W_z is the weight matrix for the update gate, and b_z is the bias for the update gate

2. **Reset Gate** (r_t):

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \quad (18)$$

Where W_r is the weight matrix for the reset gate, and b_r is the bias for the reset gate

3. **Candidate Hidden State** ($\widetilde{h_t}$):

$$\widetilde{h_t} = \tanh(W_h \cdot [r_t \odot h_{t-1}, x_t] + b_h) \quad (19)$$

Where W_h is the weight matrix for the candidate hidden state, and b_h is the bias for the candidate hidden state

4. **Hidden State** (h_t):

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \widetilde{h_t} \quad (20)$$

Where $1 - z_t$ is the complement of the update gate

4.4.2. *Comparative Analysis of Performance*

Although GRUs are generally more straightforward in architecture and exhibit comparable performance to LSTMs, some scenarios exist in which LSTMs may surpass GRUs in terms of performance.

Studies [38, 39, 40, 41] have examined these disparities and suggest that LSTM models exhibit superior performance in accurate predictions compared to GRUs in scenarios where the data sequence is characterized by high complexity or necessitates capturing complicated relationships across extended time series. However,

GRUs may exhibit superior or comparable performance in conditions of more straightforward data or when computational efficiency is given greater importance. The superior performance of LSTMs compared to GRUs in tasks involving lengthy sequences can be attributed to their enhanced memory cell architecture, which enables the retention of information for extended durations. Because of their more straightforward architecture, GRUs exhibit superior computational efficiency and demonstrate quicker training speeds compared to LSTMs. Utilizing LSTMs with longer sequences may confer a competitive advantage, as the increased complexity of these models enables them to capture complicated patterns more efficiently. However, GRUs, despite their quicker training time, may exhibit underfitting when applied to lengthy or complex datasets.

LSTM units offer more tuning flexibility due to their larger parameter count. This feature enables higher levels of precision in tuning the model, which is advantageous for optimizing the network for problems with meticulous temporal pattern recognition, such as EEG data. On the other hand, GRUs, characterized by a more streamlined structure and using fewer parameter counts, frequently exhibit comparable performance to LSTMs and sometimes better performance across some tasks [42, 43] while demonstrating superior computational efficiency. The efficiency shown by GRUs renders them more desirable for real-time applications or scenarios involving short sequences [38].

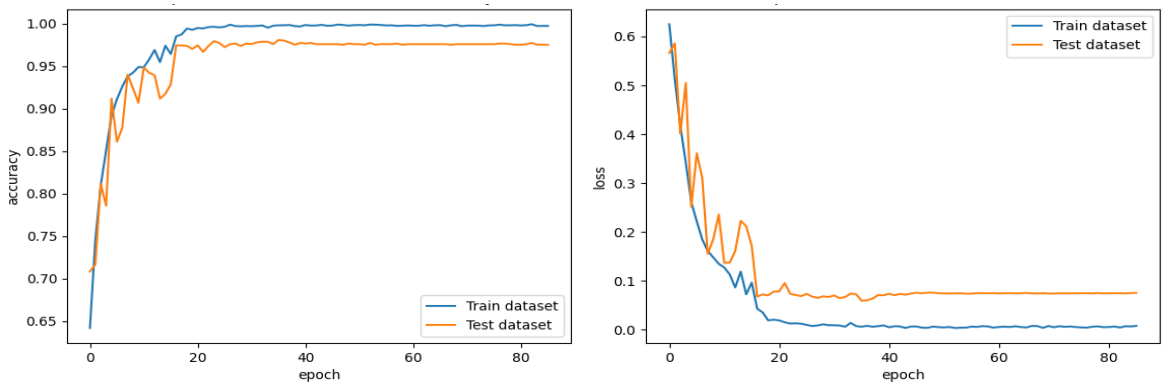
All proposed models have been evaluated with all metrics in Table 3 on a test (not-seen) dataset containing 1419 samples. As shown in Table 3, the superior attention-GRU proposed model achieved an accuracy of 98.52% as demonstrated in Figure 6a, 98.70% precision, a recall and sensitivity of 98.82%, a specificity of 98.08%, Matthews correlation coefficient of 96.92% and f-1 score of 98.76%. The confusion matrices in Figure 7 showed a thorough breakdown of the models' predictions, including the number of true positives, true negatives, false positives, and false negatives. This allowed us to examine the total accuracy and the balance of different misclassifications, which is critical for understanding how each model could behave in real-world circumstances.

The loss function during the training procedure was tracked, as in Figures 6b and 5b. The loss function supplied information on how effectively the model learned from the data, with lower loss values indicating better model performance. Assessing the loss function over time allowed us to ensure that the model efficiently minimized errors during training while not overfitting the data. These evaluation criteria indicated that the proposed models work effectively, demonstrating a great capacity to detect SZ.

As a baseline and in comparison to previous works, the proposed models have shown better performance, making them state-of-the-art for this problem. An overview of the comparison between the earlier works that addressed the same problem is illustrated in Table 4.

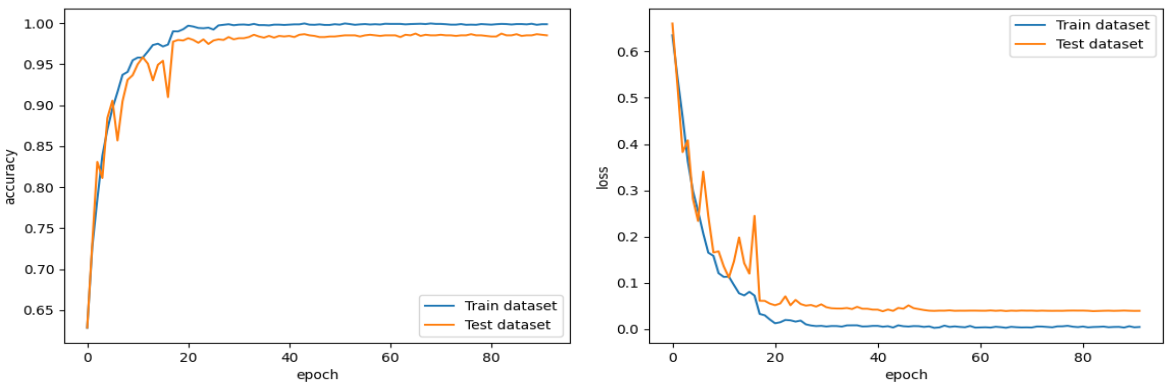
Table 3: Evaluation of the Proposed Models

Metric	Attention-LSTM	Attention-GRU
Accuracy	0.9753	0.9852
Precision	0.9788	0.9870
Recall	0.9799	0.9882
Specificity	0.9685	0.9808
Sensitivity	0.9799	0.9882
F1 Score	0.9794	0.9876
MCC	0.9487	0.9692
Training Time	828s	900s
Epochs	86	92



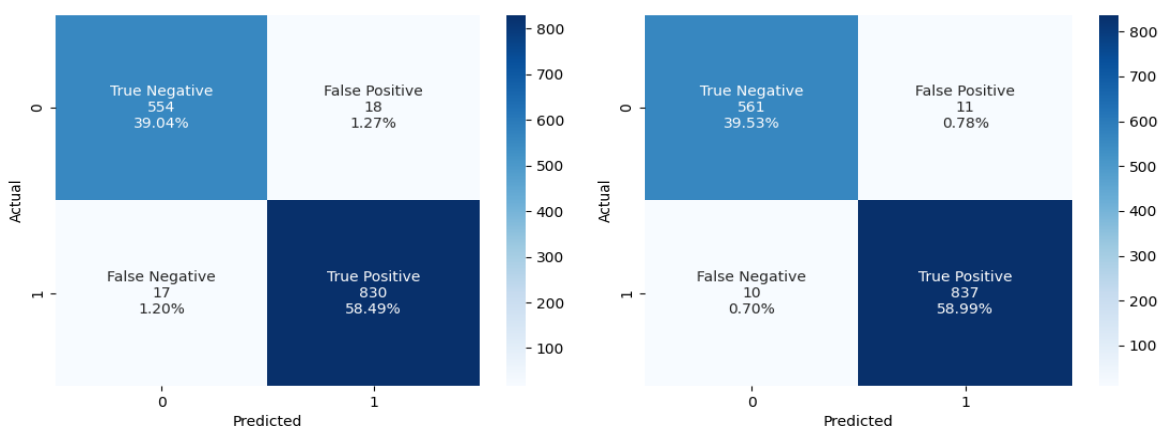
(a) (b)

Figure 5: Comparison of (a) attention-LSTM accuracy and (b) attention-LSTM loss function on test and train data.



(a) (b)

Figure 6: Comparison of (a) attention-GRU accuracy and (b) attention-GRU loss function on test and train data.



(a) (b)

Figure 7: Comparison of (a) attention-LSTM and (b) attention-GRU confusion matrices.

Table 4: Comparison of different models on the Kaggle EEG dataset

Author, Year	EEG dataset	Feature extraction	Classifier	Accuracy (%)
Ko et al. [23], 2022	32HC and 49SZ	Recurrence Plot (RP)	CNN	94.5
Ko et al. [23], 2022	32HC and 49SZ	Gramian Angular Field (GAF)	CNN	96.3
Siuly et al. [22], 2020	32HC and 49SZ	Empirical mode decomposition	Ensemble Bagged Tree	89.59
Aks¨oz et al. [21], 2022	32HC and 49SZ	Event-related potential	SVM	64.30
Aks¨oz et al. [21], 2022	32HC and 49SZ	Event-related potential	KNN	73.50
Aks¨oz et al. [21], 2022	32HC and 49SZ	Event-related potential	ANN	93.90
Prabhu et al. [20], 2020	32HC and 49SZ	Sample Entropy, Kolmogorov Complexity on four electrodes	ANN	91.25
Prabhu et al. [20], 2020	32HC and 49SZ	Sample Entropy, Kolmogorov Complexity on nine electrodes	ANN	88.75
Frick et al. [19], 2021	32HC and 49SZ	Event-related potential	Random Forest	96.4
Zhang [18], 2019	32HC and 49SZ	Event-related potential	Random Forest	81.10
Khare et al. [24], 2021	32HC and 49SZ	Flexible tunable Q wavelet transform (F-TQWT)	SVM	91.39
khare et al. [14], 2021	32HC and 49SZ	Smoothed pseudo-Wioner Ville distribution	CNN	93.36
Sahu et al. [16], 2023	32HC and 49SZ	Depth-wise separable convolution with SA and CWA	CNN	95
Proposed model 1	32HC **and** 49SZ	**Automatic Feature Extraction on EEG**	**Attention-LSTM**	**97.53**
Proposed model 2	32HC **and** 49SZ	**Automatic Feature Extraction on EEG**	**Attention-GRU**	**98.52**

5. Discussion

The results show that the proposed deep learning models using attention, GRUs, and LSTMs can accurately detect SZ from EEG brain signal data. With an accuracy of 98.52%, the model performs better than previous approaches that used manual feature extraction algorithms before applying the model. A vital advantage of this approach is that it learns to automatically extract relevant features from the EEG data end-to-end, avoiding the need for extensive manual feature engineering. The CNN layers learn spatial features, while the LSTM and GRU units recognize the temporal/sequential patterns. The attention mechanism helps focus on the most discriminative features of the signals. While promising, there are some limitations. The dataset size is relatively small, with only 81 subjects. Evaluating more extensive and more diverse datasets is needed. Additionally, the models only classify subjects as schizophrenic or healthy controls; they do not identify subtypes or severity.

Furthermore, with the growing complexity of deep learning models, it is imperative to prioritize their transparency and interpretability, particularly in clinical settings where comprehending the underlying reasoning behind a diagnosis has equal significance to the diagnosis itself. Overall, the shift towards end-to-end deep learning models in identifying SZ using EEG data shows excellent potential as it effectively tackles several constraints observed in prior methodologies. With further validation and refinement, such AI-based methods could provide an objective diagnosis tool to complement psychiatrists' current clinical assessments.

6. Conclusion

In this work, Two proposed cutting-edge deep learning models that combine convolutional neural networks, GRUs, LSTMs, and attention to detect SZ from EEG signals. The proposed models achieved 98.52% and 97.53% accuracy in classifying SZ patients versus healthy controls on a publicly available EEG dataset. The end-to-end deep learning method eliminates extensive manual feature engineering compared to previous approaches. It learns discriminative spatial and temporal representations directly from clean EEG signals through CNNs, LSTMs, and GRUs. The attention layer makes models focus on the most relevant features from previous layers. While evaluated on a dataset of only 81 subjects, the high accuracy demonstrates the promise of this approach. This work contributes to the literature by advancing automated EEG analysis for neurological disorders and offers a potential tool for an objective SZ diagnosis.

Future work should involve validating the models on larger datasets and exploring integration with other data modalities to enhance diagnostic capabilities. With larger datasets of more clinical diagnoses, such AI-powered EEG analysis could assist in efficient and objective diagnosis for SZ to enable early intervention. The automated analysis also has potential applications in monitoring the prognosis of SZ and treatment response. This novel deep learning approach offers a robust computational framework for aiding schizophrenia detection using EEG data. The proposed methodology overcomes limitations of prior detection techniques and could be an essential step towards improving access to diagnosis and care for this debilitating neurological disorder.

Acknowledgments

The authors are grateful to Brain Roach for opening access to the EEG dataset; the dataset was initially collected by the fund of the U.S. National Institute of Mental Health (NIMH).

References

[1] W. Pacific and S. A. W. Hasan, 'Magnitude and impact', 2022. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/schizophrenia. [Accessed: 24-Aug-2024].
[2] S. B. Guze, 'Diagnostic and statistical manual of mental disorders, (DSM-IV)', American Journal of Psychiatry, vol. 152, no. 8, pp. 1228–1228, 1995.
[3] T. H. McGlashan, 'Early detection and intervention of schizophrenia: rationale and research', The British Journal of Psychiatry, vol. 172, no. S33, pp. 3–6, 1998.
[4] M. Baygin, 'An accurate automated schizophrenia detection using TQWT and statistical moment based feature extraction', Biomedical Signal Processing and Control, vol. 68, p. 102777, 2021.
[5] T. M. Laursen, M. Nordentoft, and P. B. Mortensen, 'Excess early mortality in schizophrenia', Annual review of clinical psychology, vol. 10, pp. 425–448, 2014.
[6] M. B. First, DSM-5-TR® Handbook of Differential Diagnosis. American Psychiatric Pub, 2024.
[7] T. R. Insel, 'Rethinking schizophrenia', Nature, vol. 468, no. 7321, pp. 187–193, 2010.
[8] A. S. Brown and E. J. Derkits, 'Prenatal infection and schizophrenia: a review of epidemiologic and translational studies', American Journal of Psychiatry, vol. 167, no. 3, pp. 261–280, 2010.
[9] M. T. Blom et al., 'Brugada syndrome ECG is highly prevalent in schizophrenia', Circulation: Arrhythmia and Electrophysiology, vol. 7, no. 3, pp. 384–391, 2014.
[10] L. Flyckt et al., 'Muscle biopsy, macro EMG, and clinical characteristics in patients with schizophrenia', Biological psychiatry, vol. 47, no. 11, pp. 991–999, 2000.
[11] T. Miyauchi, K. Tanaka, H. Hagimoto, T. Miura, H. Kishimoto, and M. Matsushita, 'Computerized EEG in schizophrenic patients', Biological Psychiatry, vol. 28, no. 6, pp. 488–494, 1990.
[12] S. J. Eliades and X. Wang, 'Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations', Journal of neurophysiology, vol. 89, no. 4, pp. 2194–2207, 2003.
[13] H. K. Hamilton, A. K. Boos, and D. H. Mathalon, 'Electroencephalography and event-related potential biomarkers in individuals at clinical high risk for psychosis', Biological psychiatry, vol. 88, no. 4, pp. 294–303, 2020.
[14] S. K. Khare, V. Bajaj, and U. R. Acharya, 'SPWVD-CNN for automated detection of schizophrenia patients using EEG signals', IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1–9, 2021.
[15] B. Roach, 'EEG data from basic sensory task in Schizophrenia', 2021. [Online]. Available: https://www.kaggle.com/datasets/broach/button-tone-sz. [Accessed: 24-Aug-2024].
[16] G. Sahu, M. Karnati, A. Gupta, and A. Seal, 'SCZ-SCAN: An automated Schizophrenia detection system from electroencephalogram signals', Biomedical Signal Processing and Control, vol. 86, p. 105206, 2023.
[17] E. Olejarczyk and W. Jernajczyk, 'EEG in schizophrenia', 2017. [Online]. Available: https://repod.icm.edu.pl/dataset.xhtml?persistentId=doi:10.18150/repod.0107441. [Accessed: 24-Aug-2024].
[18] L. Zhang, 'EEG signals classification using machine learning for the identification and diagnosis of schizophrenia', in 2019 41st annual international conference of the ieee engineering in medicine and biology society (EMBC), 2019, pp. 4521–4524.
[19] J. Frick, T. Rieg, and R. Buettner, 'Detection of schizophrenia: A machine learning algorithm for potential early detection and prevention based on event-related potentials', in HICSS, 2021, pp. 1–10.
[20] S. Prabhu and R. J. Martis, 'Diagnosis of schizophrenia using Kolmogorov complexity and sample entropy', in 2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), 2020, pp. 1–4.
[21] A. Aksöz, D. Akyüz, F. Bayir, N. C. Yildiz, F. Orhanbulucu, and F. Latifoğlu, 'Analysis and Classification of Schizophrenia Using Event Related Potential Signals', Computer Science, pp. 32–36, 2022.
[22] S. Siuly, S. K. Khare, V. Bajaj, H. Wang, and Y. Zhang, 'A computerized method for automatic detection of schizophrenia using EEG signals', IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 28, no. 11, pp. 2390–2400, 2020.
[23] D.-W. Ko and J.-J. Yang, 'EEG-based schizophrenia diagnosis through time series image conversion and deep learning', Electronics, vol. 11, no. 14, p. 2265, 2022.
[24] S. K. Khare and V. Bajaj, 'A self-learned decomposition and classification model for schizophrenia diagnosis', Computer Methods and Programs in Biomedicine, vol. 211, p. 106450, 2021.
[25] X. Xu, G. Zhu, B. Li, P. Lin, X. Li, and Z. Wang, 'Automated diagnosis of schizophrenia based on spatial--temporal residual graph convolutional network', BioMedical Engineering OnLine, vol. 23, no. 1, p. 55, 2024.

[26] M. Albrecht, J. Waltz, J. Cavanagh, M. Frank, and J. Gold, 'GM EEG data zip 1 of 2', 2015. [Online]. Available: https://doi.org/10.5281/zenodo.29601. [Accessed: 24-Aug-2024].
[27] M. Albrecht, J. Waltz, J. Cavanagh, M. Frank, and J. Gold, 'GM EEG data zip 2 of 2', 2015. [Online]. Available: https://doi.org/10.5281/zenodo.29604. [Accessed: 24-Aug-2024].
[28] N. N. Gorbachevskaya and S. Borisov, 'EEG of healthy adolescents and adolescents with symptoms of schizophrenia', 2002. [Online]. Available: http://brain.bio.msu.ru/eeg_schizophrenia.htm. [Accessed: 24-Aug-2024].
[29] Z. Guo, J. Wang, T. Jing, and L. Fu, 'Investigating the interpretability of schizophrenia EEG mechanism through a 3DCNN-based hidden layer features aggregation framework', Computer Methods and Programs in Biomedicine, vol. 247, p. 108105, 2024.
[30] R. Yu, C. Pan, X. Fei, M. Chen, and D. Shen, 'Multi-graph attention networks with bilinear convolution for diagnosis of schizophrenia', IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 3, pp. 1443–1454, 2023.
[31] J. M. Ford, V. A. Palzes, B. J. Roach, and D. H. Mathalon, 'Did I do that? Abnormal predictive processes in schizophrenia when button pressing to deliver a tone', Schizophrenia bulletin, vol. 40, no. 4, pp. 804–812, 2014.
[32] H. Nolan, R. Whelan, and R. B. Reilly, 'FASTER: fully automated statistical thresholding for EEG artifact rejection', Journal of neuroscience methods, vol. 192, no. 1, pp. 152–162, 2010.
[33] S. J. Luck, An introduction to the event-related potential technique. MIT press, 2014.
[34] T. W. Picton et al., 'Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria', Psychophysiology, vol. 37, no. 2, pp. 127–152, 2000.
[35] D. Bahdanau, K. Cho, and Y. Bengio, 'Neural machine translation by jointly learning to align and translate', arXiv preprint arXiv:1409. 0473, 2014.
[36] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, 'Hierarchical attention networks for document classification', in Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies, 2016, pp. 1480–1489.
[37] A. Vaswani et al., 'Attention is all you need', Advances in neural information processing systems, vol. 30, 2017.
[38] R. Cahuantzi, X. Chen, and S. Güttel, 'A comparison of LSTM and GRU networks for learning symbolic sequences', in Science and Information Conference, 2023, pp. 771–785.
[39] C. J. Barberan, S. Alemmohammad, N. Liu, R. Balestriero, and R. Baraniuk, 'Neuroview-rnn: It's about time', in Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency, 2022, pp. 1683–1697.
[40] A. Pudikov and A. Brovko, 'Comparison of LSTM and GRU recurrent neural network architectures', in International Scientific and Practical Conference in Control Engineering and Decision Making, 2020, pp. 114–124.
[41] S. Yang, X. Yu, and Y. Zhou, 'Lstm and gru neural network performance comparison study: Taking yelp review dataset as an example', in 2020 International workshop on electronic communication and artificial intelligence (IWECAI), 2020, pp. 98–101.
[42] F. M. Shiri, T. Perumal, N. Mustapha, and R. Mohamed, 'A comprehensive overview and comparative analysis on deep learning models: CNN, RNN, LSTM, GRU', arXiv preprint arXiv:2305. 17473, 2023.
[43] S. Narayana, S. N. D. Sri, S. R. Kumar, T. Ajay, and S. S. Vasiq, 'Predicting the stock market index using GRU for the year 2020', in 2024 International Conference on Emerging Systems and Intelligent Computing (ESIC), 2024, pp. 399–404.

المجلد () العدد () ()

المجلة الدولية للحاسبات والمعلومات

متاح على الإنترنت على الرابط https://ijci.journals.ekb.eg

فك شفرة مرض الفصام: نهج تعلم عميق قائم على الانتباه للكشف باستخدام إشارات التخطيط الكهربائي للدماغ

محمد عبد الحميد الجندي*، شريف العتربي، عربي كشك، محمد صقر

قسم علوم الحاسب، كلية الحاسبات والمعلومات، جامعة المنوفية، المنوفية، مصر

Muhammad.Elgendi@ci.menofia.edu.eg

الملخص

يؤثر مرض الفصام على أكثر من ٢٠ مليون شخص على مستوى العالم، حيث يتم تشخيص العديد من المرضى في وقت متأخر، مما يمنعهم من تلقي العلاج المناسب. تتسم أساليب التشخيص الحالية باعتمادها على أطباء نفسيين ذوي خبرة، مما يبرز الحاجة إلى طرق أكثر كفاءة. يستكشف هذا العمل استخدام نماذج التعلم العميق القائمة على الانتباه لتصنيف إشارات التخطيط الكهربائي للدماغ EEG كطريقة غير جراحية وفعالة من حيث التكلفة، لتحديد الأفراد الأصحاء مقابل مرضى الفصام. يعتمد النموذج المقترح (attention-GRU)على شبكات الالتفاف العصبية (CNNs) لاستخراج الخصائص المكانية، ووحدات التكرار ذات البوابات (GRUs) لتحليل التسلسل، وطبقات الانتباه لتحديد المدخلات الأكثر أهمية. وعلى عكس الأعمال السابقة التي تتطلب استخراج الخصائص يدويًا بشكل شاق ومستهلك للوقت، فإن نموذجنا الشامل يتعلم مباشرة من بيانات EEG ، مما يقلل من خطوات المعالجة المسبقة للبيانات ويعزز إمكانية التطبيق الفوري في البيئات السريرية. تُظهر النتائج التجريبية تحسنًا ملحوظًا في اكتشاف مرض الفصام، حيث بلغت دقة التصنيف ٩٨,٥٢٪ باستخدام مجموعة بيانات مفتوحة المصدر لإشارات EEG، متفوقةً بذلك على الدقة التي تم الإبلاغ عنها في الدراسات السابقة. يسلط هذا العمل الضوء على إمكانيات نماذج التعلم العميق المتقدمة في تحسين دقة وكفاءة تشخيص مرض الفصام، ومعالجة تحديات تطبيع البيانات، وفتح المجال لتطوير أدوات تشخيص أكثر موثوقية في رعاية الطب النفسي. تشير نتائجنا إلى أنه مع المزيد من التحقق والتجارب، يمكن للتقييمات المعتمدة على الذكاء الاصطناعي دعم التدخل المبكر وتوفير وصول أوسع للعلاج لاضطرابات الصحة النفسية.