IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 14, No. 2, April 2025, pp. 1311~1325
ISSN: 2252-8938, DOI: 10.11591/ijai.v14.i2.pp1311-1325

Accuracy of neural networks in brain wave diagnosis of schizophrenia

Sukemi[1], Gabriel Ekoputra Hartono Cahyadi[2], Samsuryadi[3], Muhammad Agung Akbar[4]
[1]Department of Computer System, Faculty of Computers Science, Sriwijaya University, Palembang, Indonesia
[2]Department of Information Systems, Faculty of Computer Science, Sriwijaya University, Palembang, Indonesia
[3]Department of Informatic Engineering, Faculty of Computer Science, Sriwijaya University, Palembang, Indonesia
[4]Nursing Study Program, STIKes Al-Ma'arif, Baturaja, Indonesia

Article Info

Article history:

Received Nov 22, 2023
Revised Oct 18, 2024
Accepted Nov 15, 2024

Keywords:

Convolutional
Deep learning
Diagnosis
Electroencephalography
Neural network
Signal classification

ABSTRACT

This research explores the application of a modified deep learning model for electroencephalography (EEG) signal classification in the context of schizophrenia diagnosis. This study aims to utilize the temporal and spatial characteristics of EEG data to improve classification accuracy. Four popular convolutional neural network (CNN) architectures, namely LeNet-5, AlexNet, VGG16, and ResNet-18, are adapted to handle 1D EEG signals. In addition, a hybrid architecture of CNN-gated recurrent unit (GRU) and CNN-long short-term memory (LSTM) is proposed to capture spatial and temporal dynamics. The model was evaluated on a dataset consisting of EEG recordings from 14 patients with paranoid schizophrenia and 14 healthy controls. The results show high accuracy and F1 scores for all modified models, with CNN-LSTM and CNN-GRU achieving the highest performance with scores of 0.96 and 0.97, respectively. Receiver operating characteristic (ROC) curves demonstrate the model's ability to distinguish between healthy controls and schizophrenia patients. The proposed model offers a promising approach for automated schizophrenia diagnosis based on EEG signals, potentially assisting clinicians in early detection and intervention. Future work will focus on larger data sets and explore transfer learning techniques to improve the generalization ability of the model.

Corresponding Author:

Sukemi
Department of Computer System, Faculty of Computer Science, Sriwijaya University
Masjid Al Gazali St., Bukit Lama, Ilir Bar. I, Palembang City, South Sumatra 30128, Indonesia
Email: sukemi@unsri.ac.id

1. INTRODUCTION

Schizophrenia is a severe and often chronic psychiatric disorder that causes substantial personal and social burden of long-term disability [1]. Schizophrenia is a severe mental disorder characterized by positive symptoms such as delusions and hallucinations, negative symptoms including amotivation and social withdrawal, and cognitive symptoms such as working memory deficits and cognitive flexibility [2]. Worldwide, schizophrenia is one of the top 20 causes of disability and the lifetime prevalence of schizophrenia is estimated to be around 0.7% [3]. Early detection and diagnosis of schizophrenia is challenging due to the many comorbidities associated with schizophrenia, suboptimal patient management and potentially limiting positive outcomes [4]. The main problem in detecting schizophrenia is that there are no diagnostic tests or biomarkers available as they are based on the patient's history and their overall mental state [5]. Because schizophrenia shares many clinical features with other mental disorders, traditional clinical methods are unreliable and less accurate [6]. To overcome such limitations, an automated, reliable and reproducible approach of brain imaging modalities is required [7]. Although schizophrenia is usually detected by specialists, in recent times, tools such as electroencephalography (EEG) have been used to automatically detect schizophrenia. As EEG is cheaper and more practical, the use of EEG in schizophrenia detection is widely preferred [8].

EEG/magnetoencephalography (MEG) can be used for non-invasive study of brain electrical activity. Scalp potential differences from electric fields driven by neural currents are measured using EEG [9]. EEG is a physiological technique that records spontaneous electrical brain activity originating from neurons with high resolution through electrodes connected to the scalp [10]. However, EEG is a signal with a very small amplitude; therefore, recognizing emotions from EEG is a very challenging task. Nevertheless, many researchers have attempted to overcome this problem by adopting advanced techniques, including deep learning [11]. Unlike other classification methods, deep learning does not require a separate algorithm to manually extract features from data. The use of deep learning has increased due to its automatic extraction of desired features and its good performance in classification [12].

Recently, convolutional neural network (CNN) has received a lot of attention and achieved great success in the visual field due to its ability to automatically extract strong features [13]. The output of EEG signal is time series data, therefore, CNN can automatically discover and extract the internal structure of the input time series to generate deep features for classification [14]. Various types of CNN architectures have also been implemented in various EEG signal classification systems. According Yıldırım *et al.* [15], a new one-dimensional (1D) CNN model is proposed for automatic recognition of normal and abnormal EEG signals. The developed model resulted in only 20.66% misclassification rate in classifying normal and abnormal EEG signals. The model proposed by Oh *et al.* [16] resulted in classification accuracies of 98.07% and 81.26% for non-subject-based testing and subject-based testing, respectively.

In the method proposed by Sheykhivand *et al.* [17], raw EEG signals are applied to CNN-long short-term memory (LSTM) networks, without involving feature extraction/selection. Simulation results of the proposed algorithm for two-stage classification (negative and positive) and three-stage classification (negative, neutral, and positive) of emotions for 12 active channels showed accuracies of 97.42% and 96.78% and Kappa coefficients of 0.94 and 0.93, respectively. Shoeibi *et al* [18] utilized various intelligent deep learning based methods for automatic schizophrenia diagnosis through EEG signals. Classification of EEG signals was first performed by conventional machine learning methods, e.g., support vector machine, k-nearest neighbor, decision tree, naïve Bayes, random forest, highly randomized trees, and bagging. Various deep learning models were proposed, namely, LSTM, CNN, and CNN-LSTM. The results show that CNN-LSTM has the best performance. The proposed CNN-LSTM model has achieved an accuracy percentage of 99.25%, better than previous machine learning models.

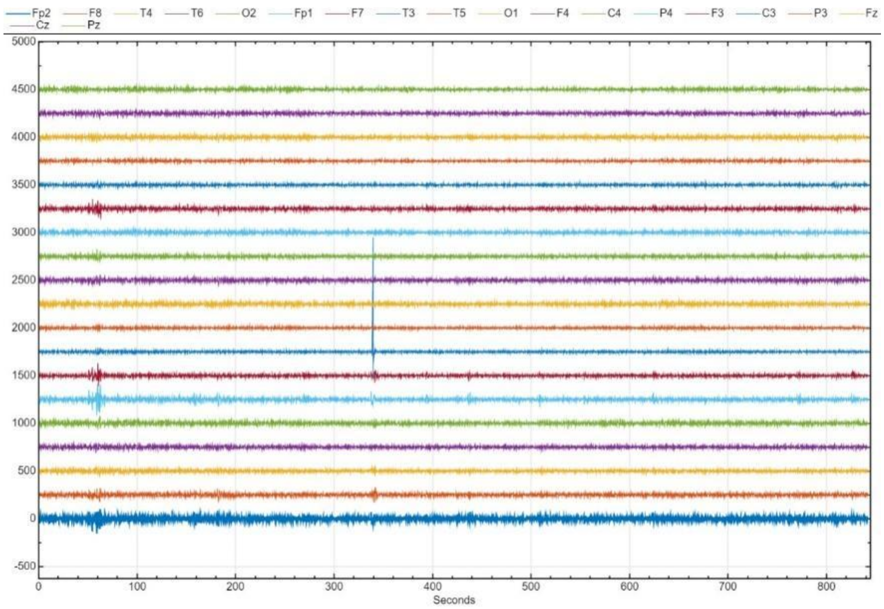
According to Yıldırım *et al.* [15], a new 1D-CNN model was proposed for automatic recognition of normal and abnormal EEG signals with 97% accuracy. Research by Sheykhivand *et al.* [17] recorded EEG signals from 14 subjects with musical stimulation for the process. Simulation results of the proposed algorithm for two-stage classification (negative and positive) and three-stage classification (negative, neutral, and positive) of emotions for 12 active channels showed accuracies of 97.42% and 96.78%, and Kappa coefficients of 0.94 and 0.93, respectively. Research by Nagabushanam *et al.* [19] analyzed deep learning for EEG signal classification. The improved LSTM and neural network were proposed for better performance with 71.3% and 78.9% accuracy in EEG classification, respectively. with comparative experiments, using OpenBCI to collect EEG action ideas during static action and dynamic action and using Conv1D-GRU EEG recognition model to train and recognize actions, respectively. The experimental results show that the brain wave action idea is easier to recognize in the static state. The recognition accuracy of the brain wave action idea in the dynamic state is only 72.27%, and the recognition accuracy of the brain wave action idea in the static state is 99.98%. An eleven-layer CNN model is proposed to detect schizophrenia using EEG signals. High classification accuracies of 98.07% and 81.26% were obtained for non-subject-based testing and subject-based testing, respectively, although the data set was small [16].

In this paper, we will test the diagnosis of EEG signals to predict schizophrenia disease using CNN methods with existing architectures (LeNet-5, AlexNet, VGG-16, and ResNet-18) modified to be adapted to 1D EEG signal datasets. The next CNN models are CNN-LSTM and CNN-gated recurrent unit (GRU). Then compare the accuracy of each architecture used and analyze the increase and decrease in accuracy in classifying EEG signals to diagnose people with schizophrenia.

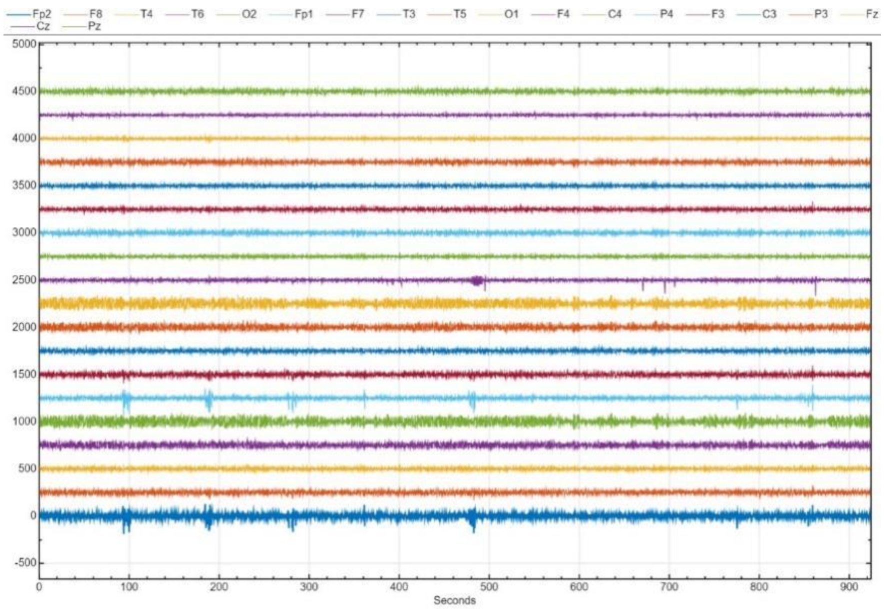
2. METHOD

2.1. EEG signal

EEG is a non-invasive measurement of the brain's electric fields. Electrodes placed on the scalp record voltage potentials resulting from current flow in and around neurons [20]. EEG is a measure of the electric fields generated by an active brain, is a brain mapping and neuro imaging technique widely used in and outside the clinical domain [21]. EEG is widely used in research involving neural engineering, neuroscience, and biomedical engineering. The output of an EEG signal is time series data. Therefore, comparing the output data with the standard EEG signal can determine the disease or problem [22]. The EEG signal visualization in this study represents data recorded over a duration of 800 to 900 seconds, as indicated by the range on the X-axis of the graph. The EEG signals consist of 19 channels recorded from standard electrode locations on the scalp, with the Y-axis representing each EEG channel. Each signal reflects the brain's electrical activity, with varying patterns and amplitudes depending on the condition of the subject. The signals are displayed with a separation offset to distinguish individual channels clearly. The y-axis values, ranging from -500 to 4,500, are likely numerical representations of the EEG signal amplitudes. This visualization style is typical in MATLAB, where the scale adapts automatically based on the signal values to facilitate clear observation of temporal patterns and variations within the EEG data. Figure 1 shows EEG Signals from normal subject (Figure 1(a)) and patients with schizophrenia (Figure 1(b)).



(a)



(b)

Figure 1. EEG signals from (a) normal subject and (b) patients with schizophrenia

2.2. Convolutional neural network

The term deep learning or deep neural network refers to an artificial neural network (ANN) [23] with multiple layers. One of the most popular deep neural networks is the CNN. One of the most popular deep neural networks is the CNN. The most beneficial aspect of CNN is that it reduces the number of parameters in the ANN. The most important assumption about the problem solved by CNN should not have spatially dependent features [24]. CNN architecture is inspired by visual perception [25]. One of the main differences is that the neuron layers in a CNN consist of neurons organized into three dimensions, the input spatial dimensions (height and width) and depth [26].

CNNs are generally designed to operate exclusively on 2D data such as images and videos. This is why they are often referred to as, "2D CNNs". Alternatively, a modified version of 2D CNNs called 1D CNNs has recently been developed. In the recent studies mentioned above, compact 1D CNNs have shown superior performance on applications that have limited labeled data and high signal variation obtained from different sources (i.e., patient EEG, civil, mechanical or aerospace structures, high power circuits, and engine or motor power) as shown in Figure 2 [27].

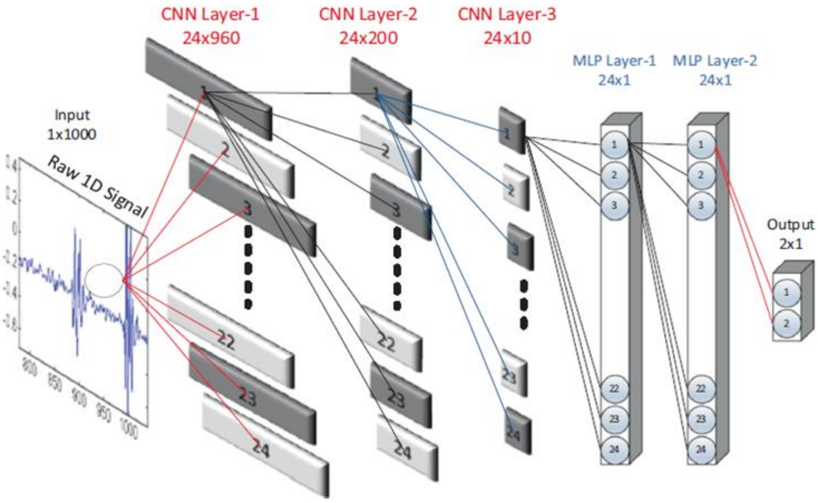


Figure 2. Example of 1D CNN architecture with 3 Conv1D and 2 dense layers [27]

2.3. Long short-term memory

Recurrent neural networks (RNNs) are a group of deep learning models used in speech recognition, natural language processing, and biomedical signal processing. CNN models are of the feed-forward type. However, RNNs have a feedback layer, where the network output returns to the network along with the next input. Since it has an internal memory, the RNN remembers previous inputs and uses them to process the input sequence. Simple RNNs, LSTMs, and GRU networks are three important groups of RNNs [18]. The specialized memory cell architecture in LSTMs makes it easy to store information for long periods of time. The cell structure has been modified by many people since then, but the standard formulation of a single LSTM cell can be given by (1)-(6) [28].

$$f_t = \sigma\left(W_f \cdot [h_t - 1, x_t] + b_f\right) \tag{1}$$

$$i_t = \sigma\left(W_i \cdot [h_t - 1, x_t] + b_i\right) \tag{2}$$

$$C_t^{\sim} = tanh\,(WC \cdot [h_t - 1, x_t] + bC) \tag{3}$$

$$C_t = f_t * (C_t - 1) + i_t * C_t^{\sim} \tag{4}$$

$$O_t = \sigma\left(W_o \cdot [h_t - 1, x_t] + b_o\right) \tag{5}$$

$$h_t = O_t * tanh\,(C_t) \tag{6}$$

RNNs are found to be an effective tool for approximating dynamic systems that deal with time and sequence-dependent data such as video, audio, and others. LSTM is part of an RNN with state memory and multilayer cell structure. Hardware acceleration of LSTMs using memristor circuits is an emerging topic of study. Figure 3 shows the example of LSTM, Figure 3(a) shows the example of original LSTM unit architecture: memory cell and two gates, Figure 3(b) shows LSTM cell with forget gate, and Figure 3(c) shows modern representation of LSTM with forget gate [29].

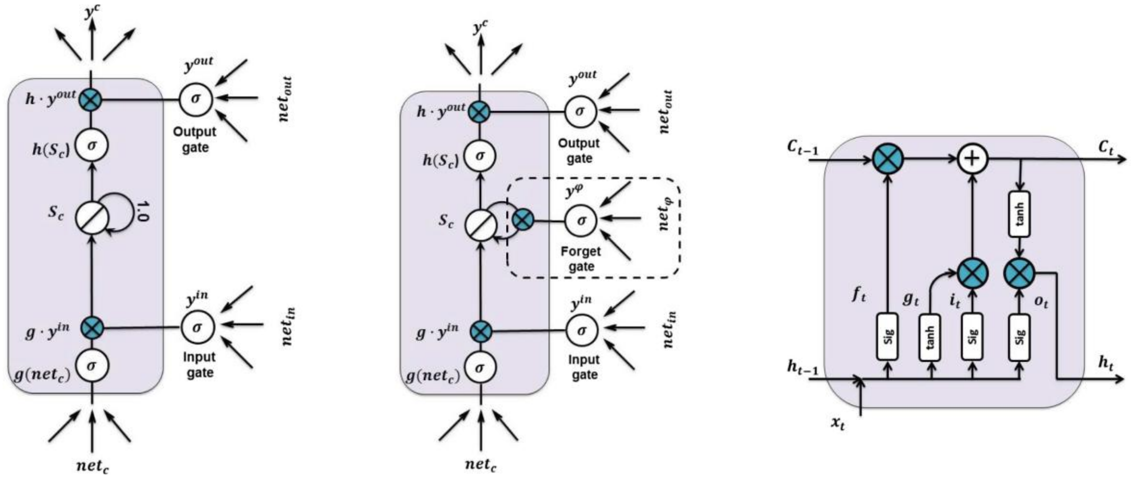


Figure 3. Type of LSTM architecture: (a) example of original LSTM unit architecture: memory cell and two gates, (b) LSTM cell with forget gate, and (c) modern representation of LSTM with forget gate [29]

2.4. Gated recurrent unit

Introduced in 2014, GRUs are similar to LSTMs but have fewer parameters. They also have gated units like LSTMs that control the flow of information within the unit but without having separate memory cells. Unlike LSTMs, GRUs do not have output gates, thus displaying their full content. The formulation of GRU can be given by (7)-(9) [28].

$$r_t = sigm(W_{xr}x_t + W_{hr}h_{t-1} + b_r) \quad (7)$$

$$z_t = sigm(W_{xz}x_t + W_{hz}h_{t-1} + b_z) \quad (8)$$

$$h_t = tanh(W_{xh}x_t + W_{hh}(r_t \odot h_{t-1}) + b_h) \quad (9)$$

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot h_t \quad (10)$$

GRU is a popular neural network nowadays, especially feed-forward networks, memory networks, and graph networks. In a feed-forward network, each neuron is on a different layer. Each layer of neurons can get signals from the previous layer of neurons and output signals to the next layer. The output is not only related to the current input and the network weights but also related to the previous inputs. It is mainly used in the fields of time series data and text data [30]. GRU has a less complicated structure compared to LSTM. It has no output gates but has an update z and a reset gate r. These gates are vectors that decide what information should be passed to the output. The reset gate defines how to combine the new input with the previous memory. The definition of how much of the last memory stored is done by the update [31]. Figure 4 shows example diagram of how the GRU works.

2.5. Proposed architecture

The CNN-1D structure used in this research is an implementation of the existing CNN-2D architecture with the addition of parameters so that it can work well in processing 1D data in the form of schizophrenia EEG signals. The CNN architectures include VGG-16, ResNet-18, LeNet-5, and AlexNet. The structure arrangement is by replacing the 2D convolutional layer with a 1D convolutional layer. Secondly, the pooling layer must also be replaced with a 1D pooling layer. Third, the fully connected layer should be removed and replaced with a 1D global average pooling layer. The output layer also needs to be modified to match the number of classes used for classification. The number of filters in each convolutional layer is reduced and each layer is increased with dropout to prevent overfitting. These modifications, by maintaining the basic structure of the CNN-2D architecture while adapting it to work with 1D data such as the schizophrenia EEG dataset.

To optimize the LeNet-5 architecture to work with 1D data, such as EEG signal data, the original 2D convolutional layer (Conv2D) is replaced with a 1D convolutional layer (Conv1D). The input_shape parameter is set to (n_timesteps, n_features), where n_timesteps is the length of the time series (in this case, the number of EEG samples) and n_features is the number of channels (in this case, the number of electrodes). The pooling layer (MaxPooling1D) is also modified to work with 1D data, and the fully connected layer (dense) remains the same as in the original LeNet-5 architecture as the number of filters in the convolutional layer and the kernel size need to be adjusted to capture relevant features in the EEG signal as shown in Figure 5. Once the architecture and hyperparameters are optimized (Table 1), the model can be trained using standard optimization techniques, such as stochastic gradient descent or Adam. Model performance can then be evaluated using appropriate metrics, such as accuracy or F1 score, on the validation set or through cross-validation.

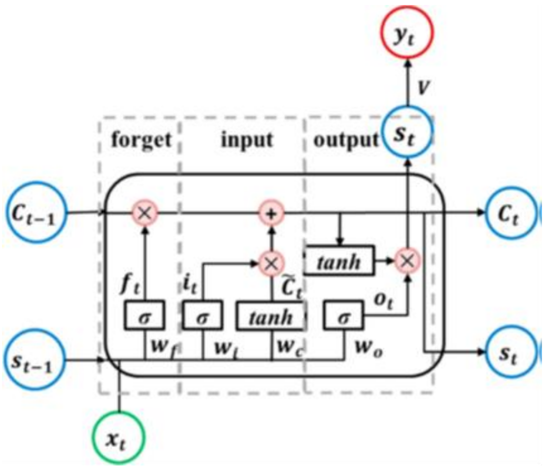


Figure 4. Example diagram of how the GRU works [32]

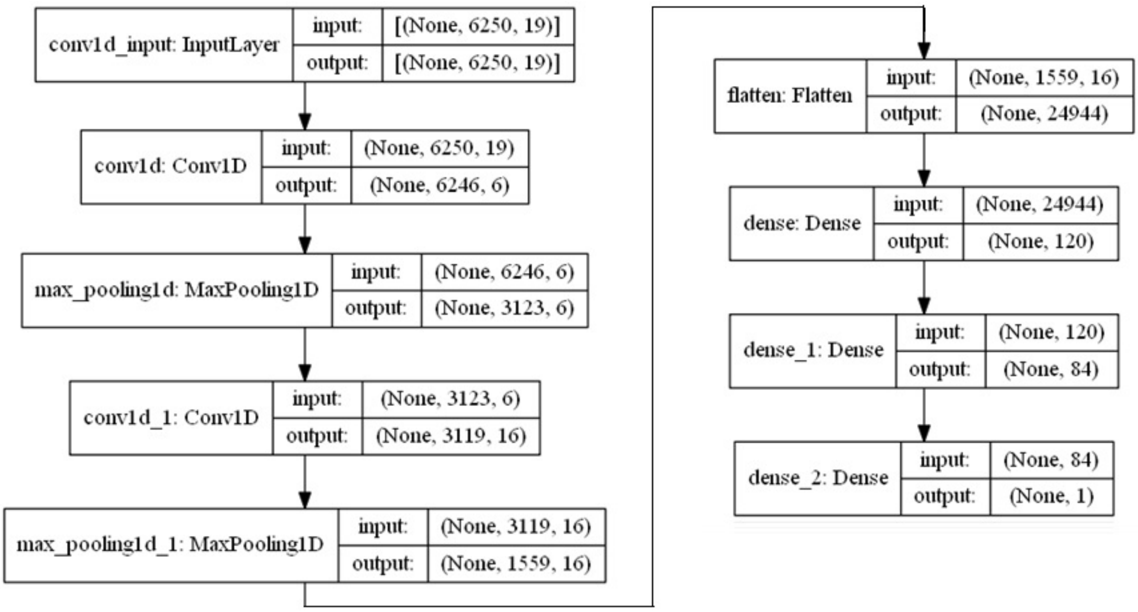


Figure 5. The LeNet-5 architecture

Table 1. The hyperparameter of 1D LeNet-5 architecture

Layers	Layers Name	Filters	Kernel Size	Strides	Activation
1	Conv1D_1	6	5	1	ReLU
2	MaxPooling1D_1	-	2	2	-
3	Conv1D_2	16	5	1	ReLU
4	MaxPooling1D_2	-	2	2	-
5	Conv1D_3	120	5	1	ReLU
6	GlobalAveragePooling1D	-	-	-	-
7	Dense_1	84	-	-	ReLU
8	Dense_2	1	-	-	Sigmoid

To adapt AlexNet to work on 1D data, it is necessary to modify the input layer to accept 1D signals rather than 3-channel images. It needs to adjust the number of filters and kernel size in the convolutional layer to better capture relevant features in the EEG signal. One approach to optimize AlexNet for 1D EEG data is to use Conv1D instead of Conv2D. It also replaces the Flatten layer with global average pooling 1D to reduce overfitting. The number of filters in the convolutional layer can be increased to capture complex patterns in the EEG signal. The last layer of AlexNet uses softmax layer for classification (Table 2).

Table 2. The hyperparameter of 1D AlexNet architecture

Layers	Layers Name	Filters	Kernel Size	Strides	Activation
1	Conv1D_1	96	11	4	ReLU
2	MaxPooling1D_1	-	3	2	-
3	Conv1D_2	256	5	1	ReLU
4	MaxPooling1D_2	-	3	2	-
5	Conv1D_3	384	3	1	ReLU
6	Conv1D_4	384	3	1	ReLU
7	Conv1D_5	256	3	1	ReLU
8	MaxPooling1D_3	-	3	2	-
9	GlobalAveragePooling1D	-	-	-	-
10	Dense_1	4,096	-	-	ReLU
11	Dropout_1	-	-	-	-
12	Dense_2	4,096	-	-	ReLU
13	Dropout_2	-	-	-	-
14	Dense_3	1	-	-	Sigmoid

For EEG data, we use a different output layer, such as a sigmoid layer for binary classification. In the original VGG-16 architecture, the input shape is (224, 224, 3), which corresponds to a 2D image with 224×224 pixels and 3 color channels. To make it work with 1D data, we can change the input shape to (n, 1, 1), where n is the length of the 1D signal. This will allow the architecture to accept 1D input signals. The VGG-16 architecture has several convolutional layers with large filter sizes, which may not be suitable for 1D data (Table 3). Some of these layers are removed and replaced with smaller filter sizes. The 3×3 filters in the two convolutional layers are replaced with 1×3 filters to capture temporal features along the signal.

More convolutional layers are added to capture more complex features along the signal. These additional layers should have smaller filter sizes to avoid overfitting and improve the model's ability to generalize. The VGG-16 architecture uses max layer pooling with a 2×2 pool size to shrink the feature map. For 1D data, we can use average pooling instead of max pooling to capture the average value of the feature map throughout the signal. The VGG-16 architecture has three fully connected layers at the end of the architecture, which may not be suitable for 1D data and can add more fully connected layers to the architecture to capture more complex relationships between features along the signal. The number of filters in each convolution layer was reduced to half to prevent overfitting and to allow the process to run on Microsoft Visual Studio Code (32-bit). Since the architecture is modified for 1D data, the optimal learning rate may be smaller than the original VGG-16 architecture. We can try different learning rates and choose the one that gives the best performance on the validation set.

In Table 4, the modified ResNet-18 architecture for 1D data, such as EEG signals, retains the basic principles of the original ResNet-18 while adapting it to handle 1D inputs. The architecture consists of several key components. First, the input layer takes the form of a 1D EEG signal. The subsequent stages follow the ResNet-18 structure, with each stage consisting of a series of residue blocks. However, in this modified version, the original 2D convolutional layer is replaced with a 1D convolutional layer. This adjustment allows the model to capture temporal dependencies in EEG data along one dimension. Residue blocks, which are the building blocks of ResNet, help to overcome the vanishing gradient problem and enable deep network training. Each block consists of convolutional layers with skip connections, which allow gradients to flow directly through the network. The number of filters in the convolutional layer varies across stages to capture different levels of features. To adapt the output of the ResNet-18 model to the classification task, a global average pooling layer is applied to reduce the spatial dimension of the feature map to one value per channel. This pooling operation averages the values of each channel, effectively summarizing the learned features. Dropout regularization is then applied to prevent overfitting by randomly disabling a small portion of neurons. Finally, the output layer consists of a dense layer with the desired number of classes, which yields the predicted class probabilities. Overall, this modified ResNet-18 architecture for 1D data provides a robust framework for capturing temporal patterns and features in EEG signals, making it suitable for tasks such as EEG-based classification and analysis in the context of neurological disorders and brain-computer interfaces.

Table 3. The hyperparameter of 1D VGG-16 architecture

Layers	Layers Name	Filters	Kernel Size	Strides	Activation
1	Conv1D_1	32	3	1	ReLU
2	Conv1D_2	32	3	1	ReLU
3	MaxPooling1D_1	-	2	2	-
4	Conv1D_3	64	3	1	ReLU
5	Conv1D_4	64	3	1	ReLU
6	MaxPooling1D_2	-	2	2	-
7	Conv1D_5	128	3	1	ReLU
8	Conv1D_6	128	3	1	ReLU
9	Conv1D_7	128	3	1	ReLU
10	MaxPooling1D_3	-	2	2	-
11	Conv1D_8	256	3	1	ReLU
12	Conv1D_9	256	3	1	ReLU
13	Conv1D_10	256	3	1	ReLU
14	MaxPooling1D_4	-	2	2	-
15	Conv1D_11	256	3	1	ReLU
16	Conv1D_12	256	3	1	ReLU
17	Conv1D_13	256	3	1	ReLU
18	MaxPooling1D_5	-	2	2	-
19	GlobalAveragePooling1D	-	-	-	-
20	Dense_1	2,048	-	-	ReLU
21	Dropout_1	0.5	-	-	-
22	Dense_2	2,048	-	-	ReLU
23	Dropout_2	0.5	-	-	-
24	Dense_3	1	-	-	Sigmoid

Table 4. The hyperparameter of 1D ResNet-18 architecture

Layers	Layers Name	Filters	Kernel Size	Strides	Activation
1	Conv1D_1	64	7	2	ReLU
2	MaxPooling1D_1	-	3	2	-
3	ResidualBlock_1	-	-	-	-
4	ResidualBlock_2	-	-	-	-
5	ResidualBlock_3	-	-	-	-
6	ResidualBlock_4	-	-	-	-
7	Conv1D_2	128	3	2	ReLU
8	ResidualBlock_5	-	-	-	-
9	ResidualBlock_6	-	-	-	-
10	ResidualBlock_7	-	-	-	-
11	ResidualBlock_8	-	-	-	-
12	Conv1D_3	256	3	2	ReLU
13	ResidualBlock_9	-	-	-	-
14	ResidualBlock_10	-	-	-	-
15	ResidualBlock_11	-	-	-	-
16	ResidualBlock_12	-	-	-	-
17	Conv1D_4	512	3	2	ReLU
18	ResidualBlock_13	-	-	-	-
19	ResidualBlock_14	-	-	-	-
20	ResidualBlock_15	-	-	-	-
21	ResidualBlock_16	-	-	-	-
22	GlobalAveragePooling1D	-	-	-	-
23	Dense_1	1	-	-	Sigmoid

In the proposed CNN-LSTM architecture, it consists of a total of 13 layers, combining CNN and LSTM layers. Similar to the model in [18], this architecture uses the global average pooling 1D layer instead of flatten to reduce the spatial dimension of the feature map. The first 10 layers of this CNN-LSTM model remain the same as the previous architecture, with the CNN layer responsible for extracting local features and patterns from the input data. The global average pooling 1D layer follows, which performs a pooling operation to summarize the extracted features in each channel. By using global average pooling 1D, aim for a fixed-length representation regardless of the length of the input sequence.

In the next layer, a dense layer with 50 neurons is used and a rectified linear unit (ReLU) activation function is used to introduce non-linearity to the model. This layer makes it possible to learn more complex feature representations. To prevent overfitting, a dropout rate of 0.25 was added to the 12th layer. Finally, at the 13th layer, a dense layer with a sigmoid activation function is used to perform the classification task. By combining the CNN and LSTM components, along with the global average pooling 1D layer, the proposed CNN-LSTM model is able to capture both spatial and temporal information from EEG signals. This architecture allows the network to learn meaningful representations and make accurate predictions for EEG signal classification in this study (Table 5).

Table 5. The hyperparameter of 1D CNN-LSTM architecture

Layers	Layers Name	Filters	Kernel Size	Strides	Activation
1	Conv1D	64	3	1	ReLU
2	Conv1D	64	3	1	ReLU
3	Dropout	0.5	-	-	-
4	MaxPooling1D	-	2	1	-
6	GlobalAveragePooling1D	-	-	-	-
7	LSTM	100	-	-	-
8	Dropout	0.5	-	-	-
9	Dense	100	-	-	ReLU
10	Dropout	0.25	-	-	-
11	Dense	50	-	-	ReLU
12	Dropout	0.25	-	-	
13	Dense	1	-	-	Sigmoid

The proposed CNN-GRU model is used for comparison in this study. This architecture combines a CNN layer with a GRU layer, providing a different approach to capture the temporal dynamics of EEG signals. A fixed global average pooling 1D layer is used to replace flatten in summarizing the extracted features in each channel. Similar to the previous model, the first 10 layers of this CNN-GRU architecture remain unchanged, consisting of a CNN layer that extracts local features from the input EEG data. The subsequent global average pooling 1D layer is used to reduce the spatial dimension and obtain a fixed-length representation, regardless of the length of the input sequence Following the pooling layer, the GRU layer is used instead of the LSTM. GRU is a type of RNN that offers a more efficient architecture by utilizing a gating mechanism to selectively store important information and discard irrelevant information from the sequence. This helps the model effectively capture the temporal dependencies present in EEG signals. To further enhance the representation capacity of the model, a dense layer with 50 neurons and a ReLU activation function was also applied to this model. This layer allows the network to learn complex feature representations. To reduce overfitting, a dropout layer with a rate of 0.25 is included after the dense layer. In the last layer, a dense layer with a sigmoid activation function is used for classification, allowing the model to output a probability value indicating the likelihood of the input EEG signal belonging to a particular class. By combining the CNN layer, GRU layer, and global average pooling 1D layer, the proposed CNN-GRU model captures both spatial and temporal information from the EEG signal. This architecture enables the network to effectively learn relevant patterns and make accurate predictions for the classification task in this study.

3. RESULTS

The performance of the modified models, including AlexNet, VGG-16, LeNet-5, and ResNet-18, was evaluated using two main metrics: test accuracy and F1 score. The results demonstrate the effectiveness of these models in classifying EEG signals for schizophrenia diagnosis. Both AlexNet and VGG-16 achieved outstanding performance, with test accuracy and F1 scores of 0.99, indicating highly accurate predictions. These models demonstrated outstanding discriminative ability, accurately distinguishing between healthy individuals and those with paranoid schizophrenia. In addition, LeNet-5 and ResNet-18 achieved test accuracy and F1 scores of 0.98, demonstrating strong classification performance.

The performance of the modified model was evaluated using a comprehensive EEG signal dataset collected from patients diagnosed with schizophrenia and a group of healthy individuals. Four widely recognized CNN architectures, namely LeNet-5, AlexNet, VGG-16, and ResNet-18, were adapted and modified to effectively handle the unique characteristics of 1D EEG data. The model undergoes rigorous training and testing procedures using appropriate data splits to ensure reliable evaluation. Evaluation metrics used to assess model performance include test accuracy and F1 scores, which are common measures used in classification tasks. The Figures 6 to 10 are the output results from each model used.

Among the modified models, the modified AlexNet and VGG-16 models stood out with outstanding performance, achieving impressive testing accuracy and an F1 score of 0.99. These models showed remarkable ability in accurately classifying EEG signals from individuals with schizophrenia and healthy controls. The robustness and effectiveness of this model can be attributed to the incorporation of specialized convolutional layers, pooling operations, and non-linear activation, which enable effective feature extraction and discrimination of relevant patterns in EEG data.

In addition, the modified LeNet-5 and ResNet-18 models also showed strong performance, achieving commendable test accuracy and an F1 score of 0.98. These models, although slightly behind their AlexNet and VGG-16 counterparts, still demonstrated their suitability for EEG classification tasks. The modified LeNet-5 model, with its relatively simpler architecture, exhibits the efficacy of utilizing 1D convolutional layers in capturing relevant temporal dependencies and extracting discriminative features from EEG signals. Similarly, the modified ResNet-18 model, with its deep residual architecture, demonstrated its ability to effectively learn complex representations and handle potential challenges in modeling EEG data.

Overall, the results obtained from the evaluation of modified CNN architectures underline their significant potential for the classification and diagnosis of schizophrenia based on EEG data. These findings provide valuable insights for the field of psychiatry research, highlighting the efficacy of deep learning approaches in analyzing and interpreting EEG signals. The high test accuracy and F1 scores achieved by the modified model reflect its robustness and reliability, demonstrating its potential for practical application in clinical settings for early detection and monitoring of schizophrenia.

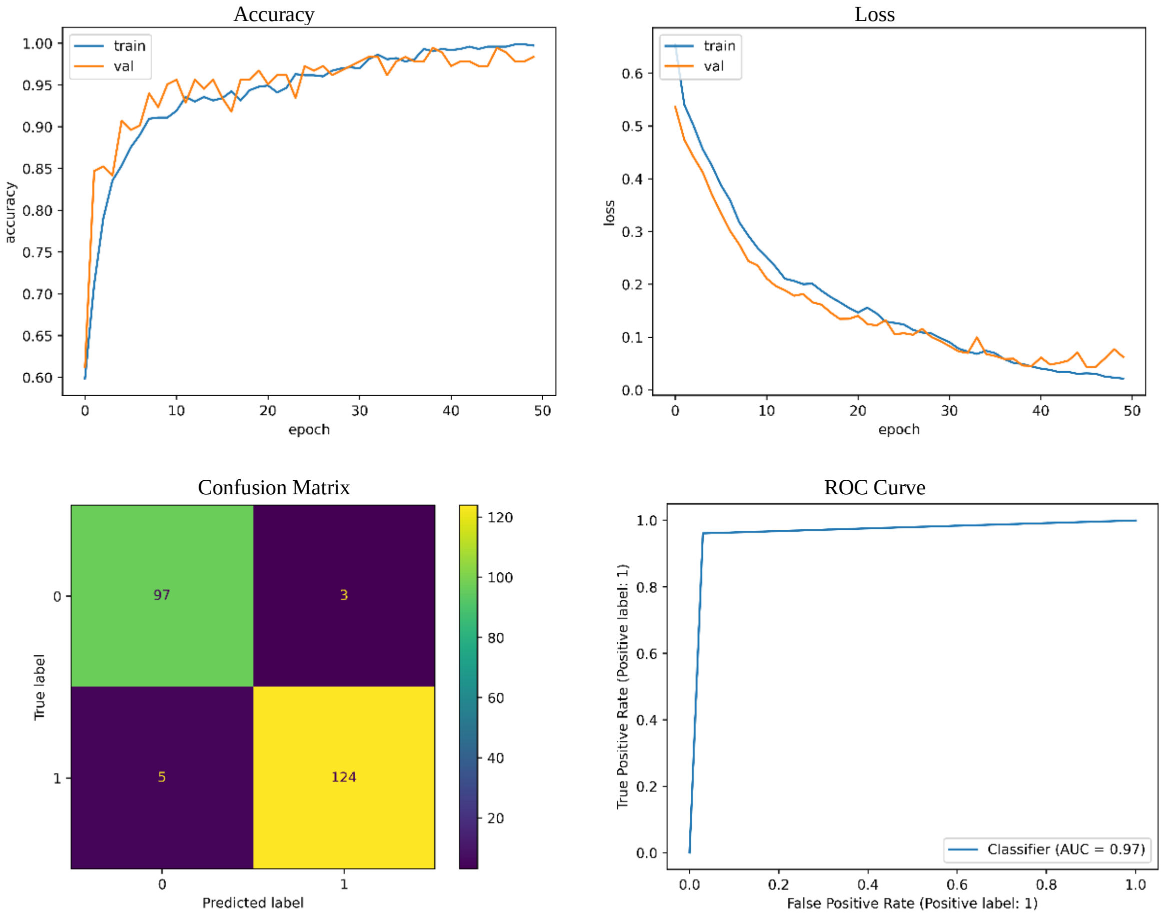


Figure 6. Accuracy curve and confusion matrix diagram of the 1D LeNet-5 model

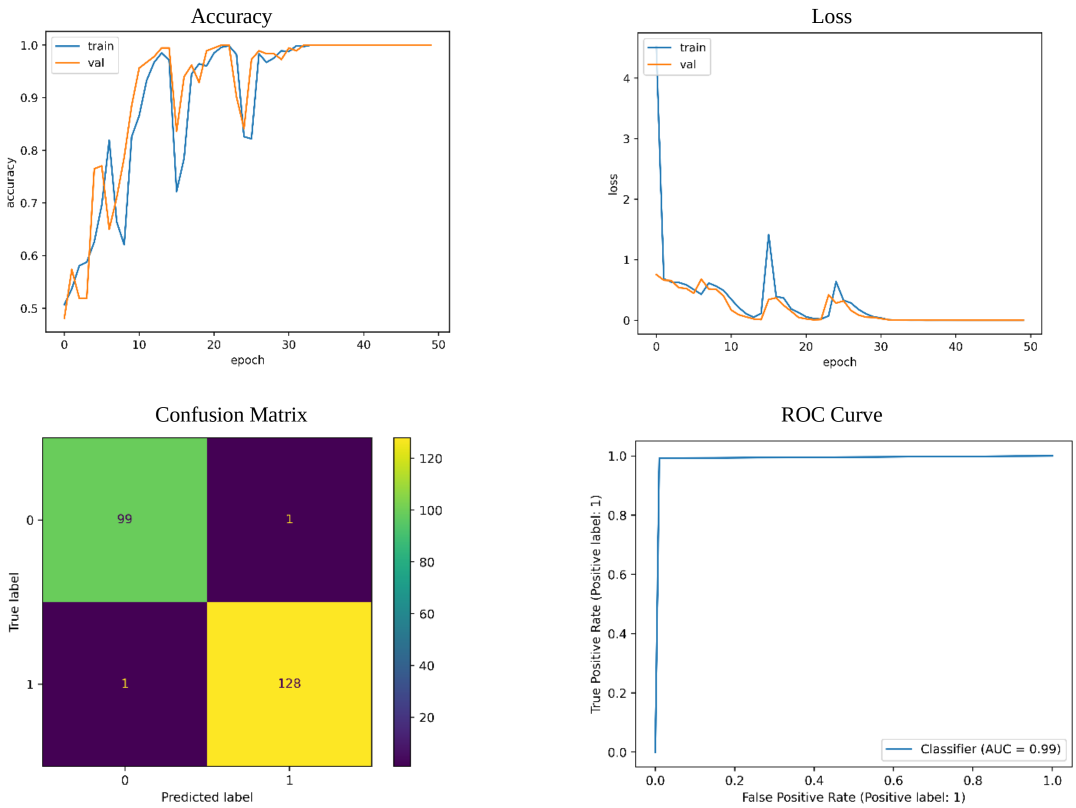


Figure 7. Accuracy curve and confusion matrix diagram of the 1D AlexNet model

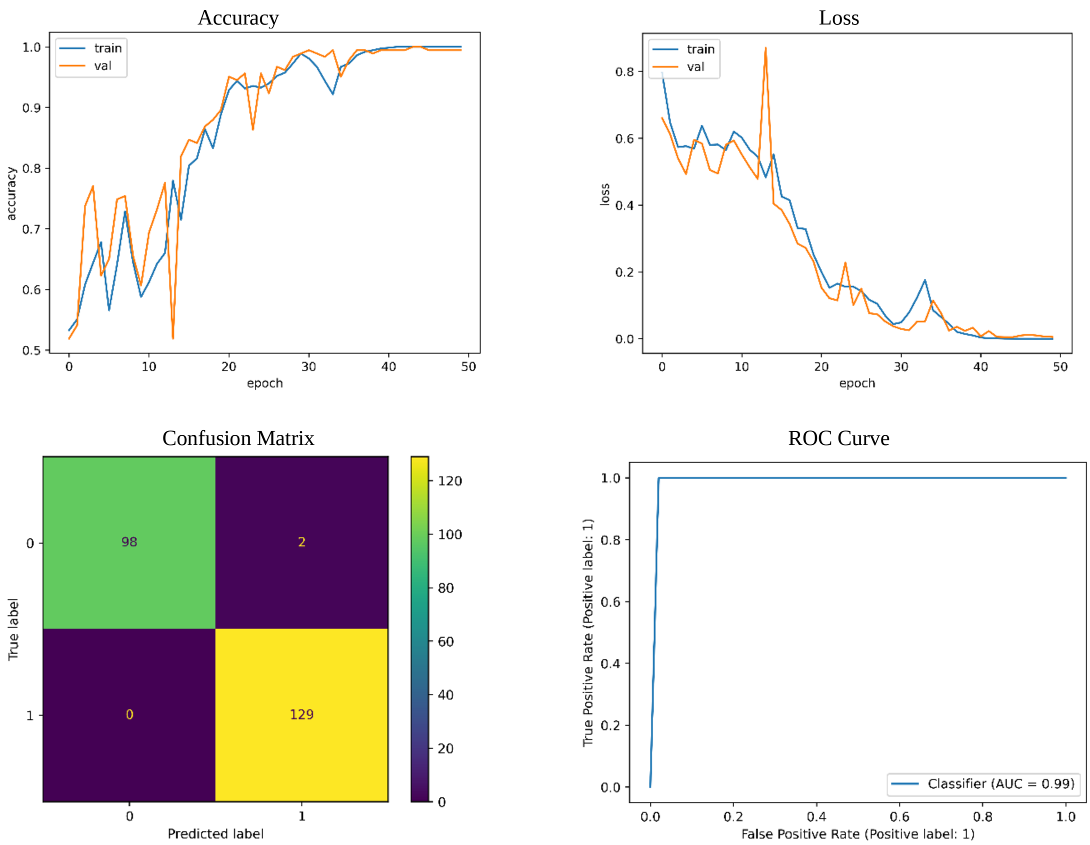


Figure 8. Accuracy curve and confusion matrix diagram of the 1D VGG-16 model

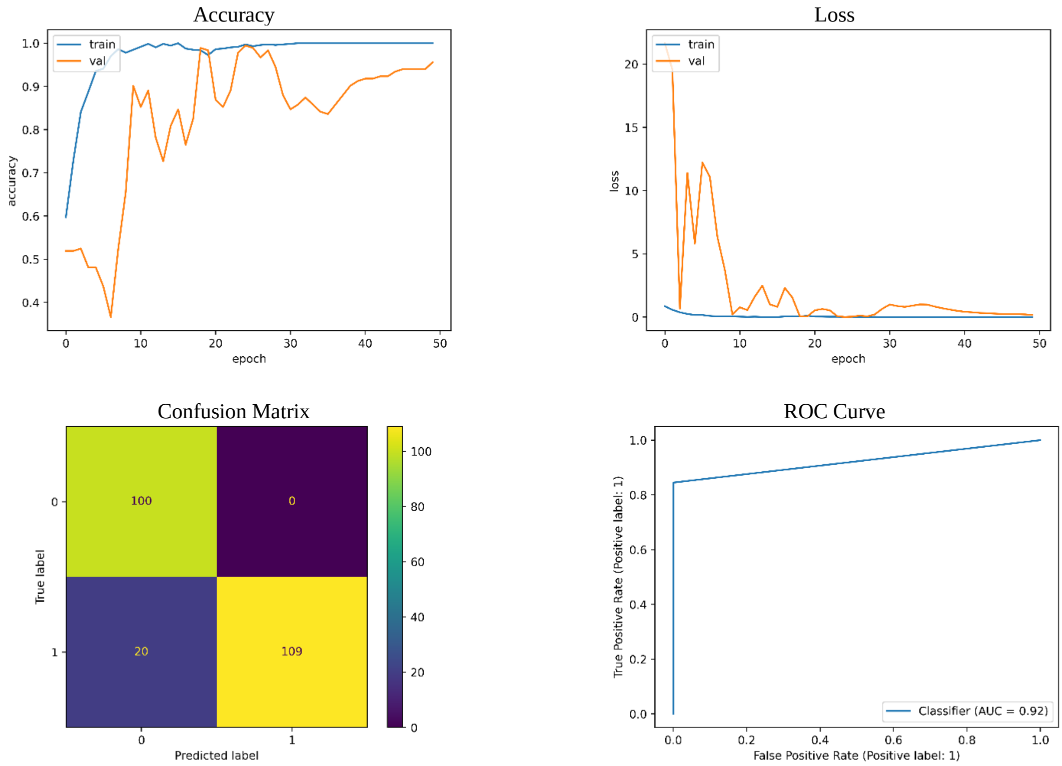


Figure 9. Accuracy curve and confusion matrix diagram of the 1D ResNet-18 model

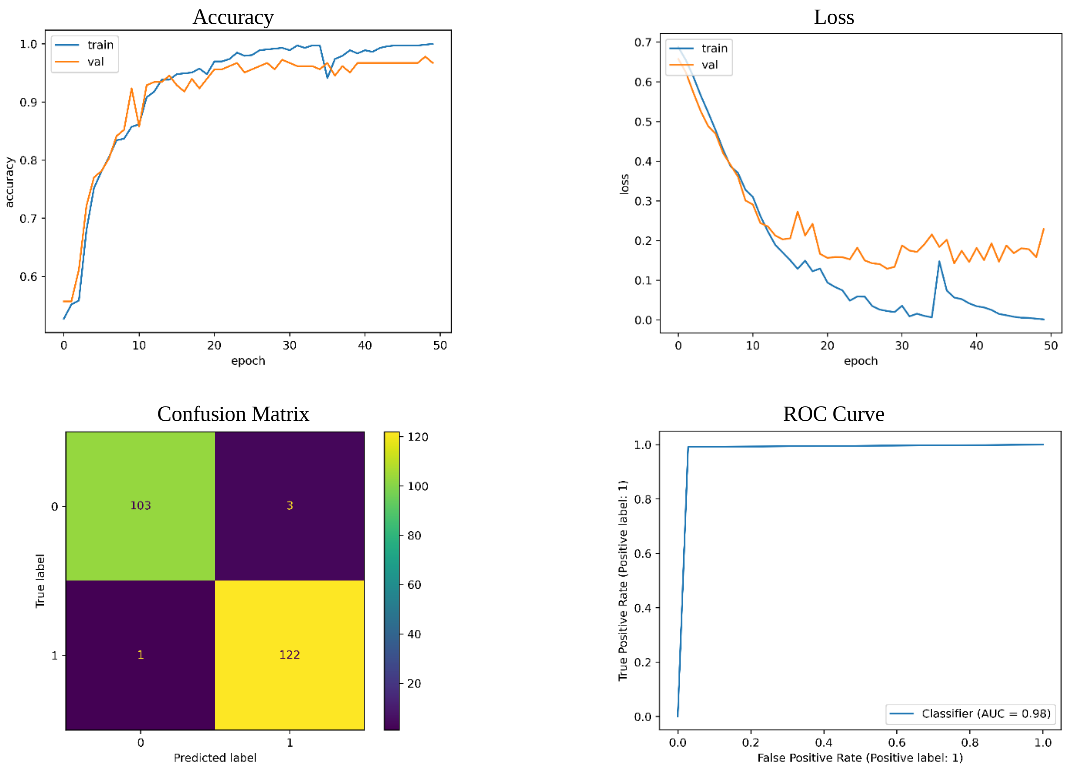


Figure 10. Accuracy curve and confusion matrix diagram of the proposed 1D CNN-LSTM model

The CNN-LSTM model with the architecture offered has an accuracy and F1 score of 0.99 which is greater than previous 1D CNN architecture models. This model has a confusion matrix comparison with a true positive of 103 and a true negative of 122 with an area under curve of 0.98. CNN-GRU on the other hand has an accuracy value of 0.986 and an F1 value of 0.987 and a confusion matrix value with true positive weights of 105 and true negative 122 is able to outperform the previous CNN architecture, but its accuracy and F1 values are not greater compared to the CNN-LSTM model. which has been tested. Figure 11 is a recap of the accumulated value of the CNN architectures that has been tested.

As the results, CNN-LSTM and CNN-GRU, achieve the highest performance with an accuracy of 0.996 and an F1 score of 0.997, indicating their superior ability to capture both spatial and temporal dependencies in EEG signals. Traditional CNN-based architectures like LeNet-5, AlexNet, and VGG-16 show competitive performance but fall slightly behind, while ResNet-18 records the lowest accuracy and F1 score. The Table 6 shows a comparative analysis of different deep learning architectures, including LeNet-5, AlexNet, VGG-16, ResNet-18, CNN-LSTM, and CNN-GRU, in terms of their accuracy and F1 score for EEG signal classification.

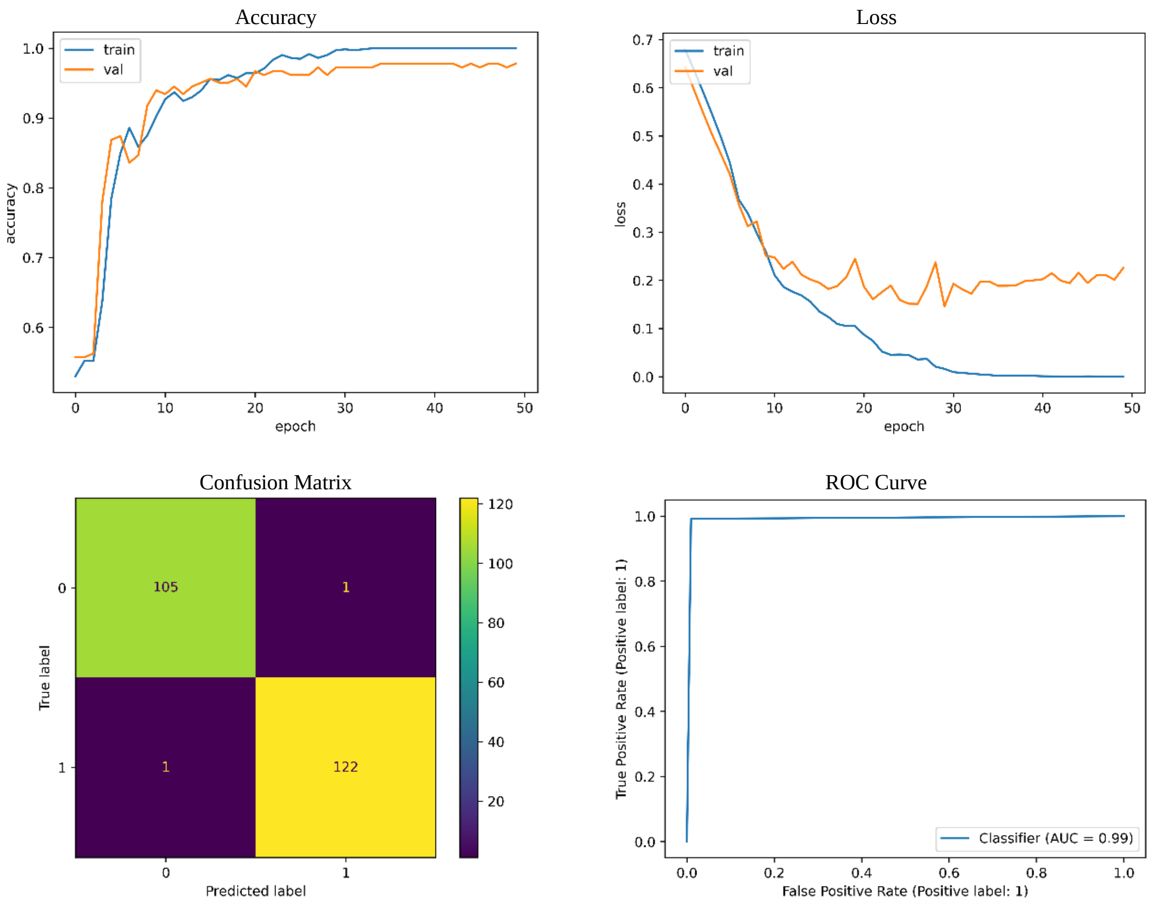


Figure 11. Accuracy curve and confusion matrix diagram of the proposed 1D CNN-GRU model

Table 6. Accuracy value and F1 score of each model

Metrics	Architecture					
	LeNet-5	AlexNet	VGG-16	ResNet-18	CNN-LSTM	CNN-GRU
Accuracy	0.983	0.975	0.976	0.962	0.996	0.996
F1 Score	0.985	0.975	0.977	0.963	0.997	0.997

4. CONCLUSION

This section presents the results obtained from training and evaluating the modified versions of the LeNet-5, AlexNet, VGG-16, and ResNet-18 architectures on EEG datasets and comparing the performance of the architectures with the proposed CNN-LSTM and CNN-GRU models. Modified versions of LeNet-5, AlexNet, VGG-16, and ResNet-18, were adapted to perform classification of 1D EEG data trained with an input shape of (6250, 19). The model was trained using cross validation to ensure robust evaluation. The 1D LeNet-5 architecture achieved an accuracy value of 98% and an F1 score of 0.98 on the test set. In the 1D AlexNet architecture, the accuracy achieved a value of 97% and an F1 score of 0.97. The 1D VGG-16 architecture has the same accuracy and F1 score as the 1D AlexNet architecture. The modified ResNet-18 architecture had the lowest score with an accuracy of 96% and an F1 score of 0.96. From these results, the modified CNN architecture is quite good at classifying schizophrenia based on 1D EEG signals. But the CNN-LSTM and CNN-GRU architecture shows slightly lower accuracy and F1 score compared to the proposed CNN models by showing significant potential for the classification of 1D data especially EEG signals. These models combine the CNN capability for feature extraction with recurrent neural networks (LSTM and GRU) to capture temporal dependencies in EEG data. The CNN-LSTM model achieves excellent performance by achieving 99% accuracy and an F1 score of 0.99 while CNN-GRU is in second place but still higher than previous CNN architectures on the test set.

ACKNOWLEDGEMENT

This publication was funded by the UNSRI's 2022 competitive research grant program from number of SP DIPA-023.17.2.677515/2022, 13 December 2022 according to Chancellor's Decree number: 0017/UN93.1/SK.LP2M.PT/2022

REFERENCES

[1] S. E. Legge, M. L. Santoro, S. Periyasamy, A. Okewole, A. Arsalan, and K. Kowalec, "Genetic architecture of schizophrenia: A review of major advancements," *Psychological Medicine*, vol. 51, no. 13, pp. 2168–2177, 2021, doi: 10.1017/S0033291720005334.

[2] R. A. McCutcheon, J. H. Krystal, and O. D. Howes, "Dopamine and glutamate in schizophrenia: biology, symptoms and treatment," *World Psychiatry*, vol. 19, no. 1, pp. 15–33, 2020, doi: 10.1002/wps.20693.

[3] G. A. Keepers *et al.*, "The American psychiatric association practice guideline for the treatment of patients with schizophrenia," *American Journal of Psychiatry*, vol. 177, no. 9, pp. 868–872, 2020, doi: 10.1176/appi.ajp.2020.177901.

[4] Y. J. Bae, M. Shim, and W. H. Lee, "Schizophrenia detection using machine learning approach from social media content," *Sensors*, vol. 21, no. 17, pp. 1–18, 2021, doi: 10.3390/s21175924.

[5] R. Buettner, D. Beil, S. Scholtz, and A. Djemai, "Development of a machine learning based algorithm to accurately detect schizophrenia based on one-minute EEG recordings," *Proceedings of the Annual Hawaii International Conference on System Sciences*, pp. 3216–3225, 2020, doi: 10.24251/hicss.2020.393.

[6] H. Akbari, S. Ghofrani, P. Zakalvand, and M. T. Sadiq, "Schizophrenia recognition based on the phase space dynamic of EEG signals and graphical features," *Biomedical Signal Processing and Control*, vol. 69, 2021, doi: 10.1016/j.bspc.2021.102917.

[7] A. Shalbaf, S. Bagherzadeh, and A. Maghsoudi, "Transfer learning with deep convolutional neural network for automated detection of schizophrenia from EEG signals," *Physical and Engineering Sciences in Medicine*, vol. 43, no. 4, pp. 1229–1239, 2020, doi: 10.1007/s13246-020-00925-9.

[8] Z. Aslan and M. Akin, "A deep learning approach in automated detection of schizophrenia using scalogram images of EEG signals," *Physical and Engineering Sciences in Medicine*, vol. 45, no. 1, pp. 83–96, 2022, doi: 10.1007/s13246-021-01083-2.

[9] S. Asadzadeh, T. Y. Rezaii, S. Beheshti, A. Delpak, and S. Meshgini, "A systematic review of EEG source localization techniques and their applications on diagnosis of brain abnormalities," *Journal of Neuroscience Methods*, vol. 339, 2020, doi: 10.1016/j.jneumeth.2020.108740.

[10] P. M. Rodrigues, B. C. Bispo, C. Garrett, D. Alves, J. P. Teixeira, and D. Freitas, "Lacsogram: A new EEG tool to diagnose Alzheimer's disease," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 9, pp. 3384–3395, 2021, doi: 10.1109/JBHI.2021.3069789.

[11] M. R. Islam *et al.*, "Emotion recognition from EEG signal focusing on deep learning and shallow learning techniques," *IEEE Access*, vol. 9, pp. 94601–94624, 2021, doi: 10.1109/ACCESS.2021.3091487.

[12] M. Moghaddari, M. Z. Lighvan, and S. Danishvar, "Diagnose ADHD disorder in children using convolutional neural network based on continuous mental task EEG," *Computer Methods and Programs in Biomedicine*, vol. 197, 2020, doi: 10.1016/j.cmpb.2020.105738.

[13] H. Cui, A. Liu, X. Zhang, X. Chen, K. Wang, and X. Chen, "EEG-based emotion recognition using an end-to-end regional-asymmetric convolutional neural network," *Knowledge-Based Systems*, vol. 205, 2020, doi: 10.1016/j.knosys.2020.106243.

[14] B. Zhao, H. Lu, S. Chen, J. Liu, and D. Wu, "Convolutional neural networks for time series classification," *Journal of Systems Engineering and Electronics*, vol. 28, no. 1, pp. 162–169, 2017, doi: 10.21629/JSEE.2017.01.18.

[15] Ö. Yıldırım, U. B. Baloglu, and U. R. Acharya, "A deep convolutional neural network model for automated identification of abnormal EEG signals," *Neural Computing and Applications*, vol. 32, no. 20, pp. 15857–15868, 2020, doi: 10.1007/s00521-018-3889-z.

[16] S. L. Oh, J. Vicnesh, E. J. Ciaccio, R. Yuvaraj, and U. R. Acharya, "Deep convolutional neural network model for automated diagnosis of Schizophrenia using EEG signals," *Applied Sciences*, vol. 9, no. 14, 2019, doi: 10.3390/app9142870.

[17] S. Sheykhivand, Z. Mousavi, T. Y. Rezaii, and A. Farzamnia, "Recognizing emotions evoked by music using CNN-LSTM networks on EEG signals," *IEEE Access*, vol. 8, pp. 139332–139345, 2020, doi: 10.1109/ACCESS.2020.3011882.

[18] A. Shoeibi *et al.*, "Automatic diagnosis of schizophrenia in EEG signals using CNN-LSTM models," *Frontiers in Neuroinformatics*, vol. 15, pp. 1–16, 2021, doi: 10.3389/fninf.2021.777977.

[19] P. Nagabushanam, S. T. George, and S. Radha, "EEG signal classification using LSTM and improved neural network algorithms," *Soft Computing*, vol. 24, no. 13, pp. 9981–10003, 2020, doi: 10.1007/s00500-019-04515-0.

[20] A. Biasiucci, B. Franceschiello, and M. M. Murray, "Electroencephalography," *Current Biology*, vol. 29, no. 3, pp. R80-R85, 2019, doi: 10.1016/j.cub.2018.11.052.

[21] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: A systematic review," *Journal of Neural Engineering*, vol. 16, no. 5, 2019, doi: 10.1088/1741-2552/ab260c.

[22] W. L. Mao, H. I. K. Fathurrahman, Y. Lee, and T. W. Chang, "EEG dataset classification using CNN method," *Journal of Physics: Conference Series*, vol. 1456, no. 1, 2020, doi: 10.1088/1742-6596/1456/1/012017.

[23] A. K. Jain, J. Mao, and K. M. Mohiuddin, "Artificial neural networks: A tutorial," *Computer*, vol. 29, no. 3, pp. 31–44, 1996, doi: 10.1109/2.485891.

[24] S. Albawi, T. A. M. Mohammed, and S. Alzawi, "Understanding of a convolutional neural network," *2017 International Conference on Engineering and Technology (ICET)*, pp. 1–6, 2017.

[25] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: analysis, applications, and prospects," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 12, pp. 6999–7019, 2022, doi: 10.1109/TNNLS.2021.3084827.
[26] K. O'Shea and R. Nash, "An introduction to convolutional neural networks," *arXiv-Computer Science*, pp. 1–11, 2015.
[27] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, "1D convolutional neural networks and applications: A survey," *Mechanical Systems and Signal Processing*, vol. 151, 2021, doi: 10.1016/j.ymssp.2020.107398.
[28] A. Shewalkar, D. Nyavanandi, and S. A. Ludwig, "Performance evaluation of deep neural networks applied to speech recognition: RNN, LSTM and GRU," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 9, no. 4, pp. 235–245, 2019, doi: 10.2478/jaiscr-2019-0006.
[29] K. Smagulova and A. P. James, "A survey on LSTM memristive neural network architectures and applications," *European Physical Journal: Special Topics*, vol. 228, no. 10, pp. 2313–2324, 2019, doi: 10.1140/epjst/e2019-900046-x.
[30] G. Zou, "The recognition of action idea EEG with deep learning," *Complexity*, vol. 2022, no. 3, 2022, doi: 10.1155/2022/5308885.
[31] P. T. Yamak, L. Yujian, and P. K. Gadosey, "A comparison between ARIMA, LSTM, and GRU for time series forecasting," in *ACM International Conference Proceeding Series*, 2019, pp. 49–55. doi: 10.1145/3377713.3377722.
[32] H. Zhao, Z. Chen, H. Jiang, W. Jing, L. Sun, and M. Feng, "Evaluation of three deep learning models for early crop classification using Sentinel-1A imagery time series-a case study in Zhanjiang, China," *Remote Sensing*, vol. 11, no. 22, 2019, doi: 10.3390/rs11222673.

BIOGRAPHIES OF AUTHORS

Sukemi was born in Palembang in 1966. He studied Bachelor of Electrical Engineering at UNSRI (1992). Then continued at the Bandung Institute of Technology Bandung for a Masters degree (1999) and a Doctorate in Engineering in 2016 at the University of Indonesia. He has taught at the Department of Computer Systems of Sriwijaya University since 2006 until now. The first research regarding the paradigm shift in computing systems was submitted at the Cyberneticscom2012 IEEE Indonesia section, then continued at TelSaTech2013 with the same topic. The last publication was about signal processing in the JNTE SINTA 2 journal, and image processing in the JSI SINTA 4 journal. He can be contacted at email: sukemi@unsri.ac.id.

Gabriel Ekoputra Hartono Cahyadi S.SI., M.Kom. is a lecturer in the Department of Information Systems at the Faculty of Computer Science (2023), Sriwijaya University. He achieved bachelor's degree (2018) and master's degree (2022) from Sriwijaya University, with a focus on deep learning for medical applications in the thesis. With a particular interest in image processing and neural networks, he has successfully authored two papers in these areas. As an academic and researcher, he is dedicated to integrating his knowledge and skills in deep learning to tackle real-world challenges in the field of deep learning. He can be contacted at email: ekoputra2695@gmail.com.

Samsuryadi is currently a lecturer at the Department of Computer Sience Faculty of Computer Science, Universitas Sriwijaya, Indonesia. He received a Bachelor of Science in Mathematics (1995) from Universitas Sriwijaya, Indonesia, Magister of Computer Science (2002) from Gadjah Mada University, Indonesia, and Ph.D. in Computer Science (2013), Universiti Teknologi Malaysia, Malaysia. His research interest is image processing, machine learning, and pattern recognition. He can be contacted at email: samsuryadi@unsri.ac.id.

Muhammad Agung Akbar is a Doctor of Nursing candidate from Faculty of Nursing, Universitas Indonesia. He completed his education at Poltekkes Kemenkes Palembang (nursing diploma), Universitas Andalas (bachelor and nursing profession), and Universitas Padjadjaran (Master of Nursing). His research interests in nursing and health sciences. He can be contacted at email: magungakbar24@gmail.com.