

Article

Schizophrenia Detection on EEG Signals Using an Ensemble of a Lightweight Convolutional Neural Network

Muhammad Hussain , Noudha Abdulrahman Alsalooli, Norah Almaghrabi and Emad-ul-Haq Qazi

Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia; 438202912@student.ksu.edu.sa (N.A.A.); 439203588@student.ksu.edu.sa (N.A.); 434108002@student.ksu.edu.sa (E.-u.-H.Q.)

* Correspondence: mhussain@ksu.edu.sa

Abstract: Schizophrenia is a chronic mental disorder that affects millions of people around the world. Neurologists commonly use EEG signals to distinguish schizophrenia patients from normal controls, but their manual analysis is tedious and time-consuming. This has motivated the need for automated methods based on machine learning. However, the methods based on hand-engineered features need human experts to decide which features should be extracted. Though deep learning has recently shown good results for schizophrenia detection, the existing deep models have high parameter complexity, making them prone to overfitting because the available data are limited. To overcome these limitations, we propose a method based on an ensemble-like approach and a lightweight one-dimensional convolutional neural network to discriminate schizophrenia patients from healthy controls. It splits an input EEG signal for analysis into smaller segments, where the same backbone model analyses each segment. In this way, it makes decisions after scanning an EEG signal of any length without increasing the complexity; i.e., it scales well with an EEG signal of any length. The model architecture is simple and involves a small number of parameters, making it easy to implement and train using a limited amount of data. Though the model is lightweight, enough trials are still needed to learn the discriminative features from available data. To tackle this issue, we introduce a simple data augmentation scheme. The proposed method achieved an accuracy of 99.88% on a public benchmark dataset; it outperformed the state-of-the-art methods. It will help neurologists in the rapid and accurate detection of schizophrenia patients.



Citation: Hussain, M.; Alsalooli, N.A.; Almaghrabi, N.; Qazi, E.-u.-H. Schizophrenia Detection on EEG Signals Using an Ensemble of a Lightweight Convolutional Neural Network. *Appl. Sci.* **2024**, *14*, 5048. <https://doi.org/10.3390/app14125048>

Academic Editor: Zhibin Lv

Received: 1 May 2024

Revised: 4 June 2024

Accepted: 6 June 2024

Published: 10 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: schizophrenia; EEG classification; deep learning; convolutional neural network (CNN)

1. Introduction

Schizophrenia, affecting about 1% of the global population [1], is a complex mental disorder that often, but not always, progresses to chronicity. It affects many aspects of the daily life of patients by causing cognitive deficits and a lack of integration between thinking, emotion, and speech, and it interferes with the management of these processes [2,3]. There are two types of schizophrenia symptoms: positive and negative symptoms. Delusions, hallucinations, and thinking disorders are usually described as positive symptoms, whereas speech difficulty and absence of normal abilities are usually described as negative symptoms. People with schizophrenia usually face difficulties in employment, marriage, and relationships with people [4]. As with most diseases, accurate diagnosis of schizophrenia at early stages can minimize its negative effects and improve medication effectiveness [5]. Schizophrenia is diagnosed traditionally by interviewing patients to ask them about their clinical symptoms. However, this traditional way is sometimes inaccurate because some patients hide their symptoms. This is in addition to the difficulty of distinguishing between schizophrenia and some other diseases by the experts due to the similarity of their symptoms [6,7]. For these reasons, many techniques have been developed to enhance the process of brain monitoring, like magnetic resonance imaging (MRI), functional magnetic resonance imaging (fMRI), and electroencephalography (EEG) brain signals.

Electroencephalograms (EEGs) are a common technique to measure the brain and its electrical activity. It is considered one of the most helpful brain imaging techniques due to the high resolution of its spatiotemporal data via a direct representation of the brain dynamics [8]. In addition, it provides better discrimination between the brain images of healthy people and those with mental diseases [9]. Moreover, it has a low cost compared to other brain imaging techniques. EEG signals are recorded by placing several channels along the scalp. The number of channels usually varies according to the features needed to be extracted. One of the most common systems for placing the channels is the 10–20-placement system [10]. Visual analysis of EEG signals by neurologists is laborious and subjective. Different machine-learning algorithms have been proposed to overcome this problem and assist the neurologists. Traditional algorithms are based on hand-engineering techniques, including feature extraction, feature selection, and classification. However, there are some limitations to using these algorithms to classify EEG signals of schizophrenia. One of their limitations is that the feature extraction process is done without learning the characteristics of the data. In addition, the number of selected features might be high, which leads to the curse of the dimensionality problem. These limitations usually affect classification accuracy and make it unacceptable for diagnosis systems.

The limitations of traditional machine learning techniques and the outstanding performance of deep learning in various applications motivated the use of deep learning for schizophrenia detection, and many deep learning-based techniques have been proposed [11–16]. The performance of these techniques is good; they are prone to overfitting because they have highly complex architectures involving a large number of learnable parameters, but the available data for schizophrenia detection are limited. To overcome these problems, there is a need for a robust method that is based on a lightweight deep learning-based model, which involves a small number of learnable parameters so that it can be learned with limited data, avoiding overfitting. Motivated by the ensemble learning approach, we propose an ensemble-like method to scan EEG signals for the detection of schizophrenia. We propose a lightweight 1D CNN model based on deep learning theory as a backbone model for this method. Its simple architecture involves a few learnable parameters and can be trained with limited data. It automatically learns and extracts discriminative features relevant to schizophrenia. Though the proposed lightweight 1D CNN involves a small number of parameters, it needs enough data to learn these parameters and avoid overfitting, but the available dataset is very small. To tackle this issue, we introduce a simple data augmentation scheme that creates EEG trials from the available limited dataset.

The main contributions of this study are as follows:

- A lightweight one-dimensional convolution neural network (1D-CNN) model based on the pyramid approach is introduced, which is suitable for embedded systems. It involves a small number of learnable parameters and can be learned with a small amount of data without suffering from overfitting problems.
- An efficient and automated ensemble-like method based on the proposed 1D-CNN model and a simple ensemble approach for classifying people with schizophrenia and healthy controls. It splits an input EEG signal for analysis into smaller segments, where the same backbone model analyses each segment. In this way, it makes decisions after scanning an EEG signal of any length without increasing the complexity, i.e., it scales well with an EEG signal of any length.
- A data augmentation scheme to solve the data scarcity problem that is encountered in the training of the deep (1D-CNN) model.

The rest of the paper is organized as follows. Section 2 reviews the recent studies on schizophrenia classification based on EEG. In Section 3, the proposed method is described to perform binary classification between people with schizophrenia and healthy controls. In Section 4, the dataset and data augmentation scheme are discussed. In Section 5, the implementation details and results are presented. Section 6 discusses the results obtained and compares them with state-of-the-art methods. In Section 7, we conclude our work.

2. Literature Review

In recent years, many researchers have attempted to detect schizophrenia on EEG signals. Various hand-engineered and deep learning-based techniques have emerged to handle this problem. This section comprehensively overviews the current methods and discusses their pros and cons.

2.1. Hand-Engineered Techniques

Different feature extraction methods and classification algorithms have recently been used to detect schizophrenia. Baygin [17] extracts statistical moment features in the frequency domain using Tunable Q-Factor Wavelet Transform (TQWT) and selects discriminative features using the ReliefF method. For classification, K-Nearest Neighbor (KNN) is employed. Khare et al. [18] select the AF7 channel using the Fisher score method. Then, it is decomposed into sub-bands using Flexible-TQWT (F-TQWT), statistical features are extracted, and the Kruskal Wallis test is used to select the discriminative features. The flexible least square support vector machine (F-LSSVM) is employed for the classification.

Some researchers have studied brain connectivity maps for feature extraction. Ciprian et al. [19] extracted features using the symbolic transfer entropy (STE) method and select using the Relief method. Various methods are used for classification; the best performance was achieved with KNN. Kim et al. [20] extract brain network connectivity features using minimum norm estimation (MNE) and phase-locking value (PLV). The sequential forward selection (SFS) method is used for feature selection, and the linear discriminant analysis (LDA) is used for classification. Kumar et al. [21] perform feature extraction using a histogram of local variance (HLV) and symmetrically weighted-local binary patterns (SLBP), which are then reduced by a correlation-based feature selection algorithm (CBFS).

The hand-engineered methods require several preprocessing stages, including feature extraction, feature selection, and classification, making them laborious. The traditional methods are time-consuming and subject to human mistakes. Important information could be lost when using unsuitable feature extraction measures, and experts with extensive knowledge in this area are required. Though Baygin's method [17] gives high accuracy, it seems to suffer from overfitting because the number of subjects used in the study was relatively small (i.e., 24 participants), and trials of long temporal length (i.e., 25 s) were used. Khare et al. [18] use only one out of 64 channels; the leftover channels can contain sensitive and important information. There is uncertainty about the generalizability of their findings. These methods give good performance only with trials of long temporal length.

2.2. Deep Learning-Based Techniques

Many deep learning-based techniques have been proposed, such as taking raw EEG trials as input or converting them into 2D images. Oh et al. [22] developed two CNN models consisting of 11 layers to classify the raw EEG signals. Sridhara et al. [23] introduced two LSTM architectures and two channels (FP1 and FP2) for schizophrenia detection. In another study, Chandran et al. [24] also used LSTM for schizophrenia classification. They used 6790 features extracted using three non-linear methods, Katz fractal dimension, approximate entropy, and variance, which are fed to the LSTM model.

Another approach is combining the CNN and LSTM models for schizophrenia detection. Shoeibi et al. [25] compared different traditional and deep learning techniques. In their method, the EEG signal is segmented into trials of 25 s, which are normalized using the z-score and L2 methods. The best accuracy was achieved with 1D-CNN-LSTM involving ReLU activation. Sharma et al. [11] also introduced a method based on the CNN and LSTM models. Supakar et al. [12] built an LSTM model to classify EEG signals as healthy control or schizophrenia.

Sairamya et al. [13] utilized the RLNDiP method to extract discriminate features from the time domain (TD) and time–frequency domain (TFD) of the EEG signals. After that, the Kruskal–Wallis test is used to select the dominant features that are fed to the ANN model.

Some studies first converted an EEG signal trial into a 2D image and then employed different 2D-CNN models designed for Image Processing and Computer Vision tasks. Tynes et al. [14] converted EEG signal trials into 2D images using spectrograms and fed them to 2D CNN models. For training CNN models, they used two meta-learning methodologies, Model-Agnostic Meta-Learning (MAML) and prototypical networks, to overcome the issue of the limited dataset. Aslan et al. [15] transformed EEG trials into 2D images using Continuous Wavelet Transform (CWT) and employed a pre-trained VGG-16 model for classification. Calhas et al. [16] utilized discrete short-time Fourier transform (DSTFT) to convert EEG signal trials into images and then extracted features by a Siamese neural network (SNN), which was followed by an XGBoost classifier.

Ko et al. [26] adopted Recurrence Plot (RP) and Gramian Angular Field (GAF) methods to transform the EEG signal into an image that was fed to VGGNet. A combination of functional connectivity theory and deep learning model was employed by Shen et al. [27] to classify the EEG signal. The frequency bands of the EEG signal were extracted using continuous wavelet transform (CWT). After that, the cross-mutual information (CMI) method was employed to convert data into a functional connectivity matrix, which was then fed to the 3D-CNN classifier.

A close look at the existing methods reveals that, firstly, the complex deep models were used with relatively small datasets, which is an indicator of the overfitting problem; the methods in [11,14,15,22,24,25] achieved accuracies of 98.07%, 99%, 99.25%, 99.9%, 96.83%, and 99.5%, respectively. Moreover, some studies [11–14,16,21,22,24,27] used EEG trials of long temporal lengths in a range between 20 to 60 s, which requires more computation and time. Also, in [23], the authors adopted only two channels, FP1 and FP2, without providing enough information on why they chose them. Aslan et al. [15] adopted very deep, complex, and pre-trained VGG-16. The LSTM classifier was used extensively to detect schizophrenia using 1D EEG signals. Thus, it is necessary to investigate the design of lightweight 1D CNN models. Therefore, we proposed a lightweight CNN method that takes a raw EEG signal (1D) as input to detect schizophrenia.

3. Proposed Method

In this study, we aim to detect schizophrenia by analyzing brain activity through EEG signals using a method that involves minimal computational complexity. Traditionally, neurologists review EEG signals manually, which is time-consuming and subjective. Our approach simplifies this by using a type of deep learning model known as a convolutional neural network (CNN), specifically designed to handle the data format of EEG signals, which has shown promising results over traditional methods based on hand-crafted features [11,15,25].

We utilize a lightweight CNN, which means it operates effectively even with smaller datasets—like those commonly available in medical settings—without losing accuracy. The model analyzes small segments of EEG data, ensuring it can handle data of any length without increasing computational demands.

Furthermore, we employ data augmentation to enhance our model's learning capability from limited data. This technique artificially expands the dataset by slightly altering existing data points to generate new ones, providing our model with more examples to learn from. Our method aims to provide a reliable, automated alternative to manual EEG analysis, offering the robust and accurate detection of schizophrenia, which could greatly assist in clinical practice.

In this section, the problem formulation and an overview of the design of the proposed method are given first. Then, the architecture of the backbone CNN model is presented. After that, the data augmentation technique is discussed.

3.1. Problem Formulation and the Design

Consider a segment (trial/epoch) of an EEG signal, represented by $x \in \mathbb{R}^{C \times T}$, which includes data from C channels recorded over T time stamps, i.e.,

$$x = \begin{bmatrix} x_1(t_1) & x_1(t_2) & \cdots & x_1(t_T) \\ x_2(t_1) & x_2(t_2) & \cdots & x_2(t_T) \\ \vdots & \vdots & \cdots & \vdots \\ x_C(t_1) & x_C(t_2) & \cdots & x_C(t_T) \end{bmatrix} \quad (1)$$

The goal is to predict whether an EEG segment indicates ‘healthy control’ or ‘schizophrenia’. This is done by creating a function f , parameterized by θ , that assigns a label ℓ to each EEG segment. In the traditional approach, f is designed as a composition function $(f_3 \circ f_2 \circ f_1)(x) = \ell$ where the functions f_1, f_2 , and f_3 are preprocessing of x , feature extractor, and the classifier, respectively. A deep learning-based method learns the functions f_1, f_2 , and f_3 from the training data in an automatic way (i.e., without the need for human experts as done in hand engineering techniques). Deep learning can learn the hierarchical representation of the features by deriving low-level features from the input and high-level features from low-level ones [28,29]. As such, we employed deep learning to formulate f .

Deep learning-based methods usually involve high parameter complexity. To keep the learnable parameter complexity small, we used an ensemble-like approach to design f for schizophrenia detection. The system design for f is shown in Figure 1; it takes an EEG trial x of six seconds temporal length as input, segments it into three frames $\{S_1, S_2, S_3\}$, predicts the label of each frame after passing them to the 1D-CNN model, and finally takes the decision based on the majority vote. Splitting the input trial into smaller segments allows the design of a lightweight CNN model for scanning each segment, thereby reducing the complexity of the method. The main component of the system is the backbone 1D-CNN model, which is described in Section 3.2.

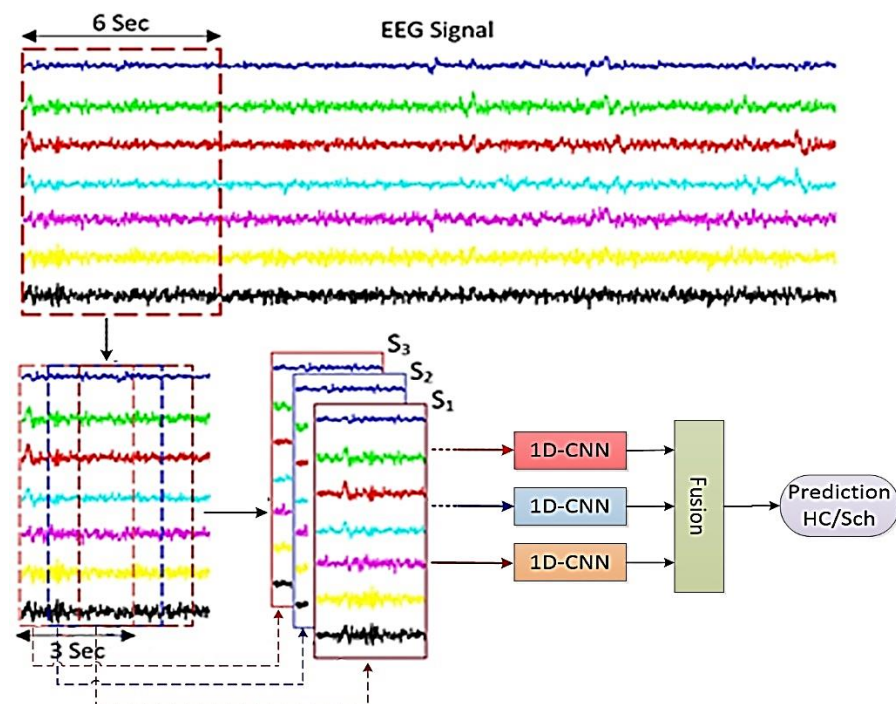


Figure 1. The design of the system for schizophrenia detection. A six-second EEG trial is divided into three-second windows that are passed to 1D-CNN to predict their labels, which are finally fused to get the predicted label of the input trial.

This approach has three obvious benefits: (i) first, it does not use different base models to predict the label of an EEG trial; rather, it splits the EEG trial into three segments and passes each segment into the same backbone model to predict its label and final decision. The diversity comes from the split of EEG trial, avoiding the training of different deep models; (ii) second, it helps to keep the parameter complexity of the 1D-CNN model,

avoiding the overfitting problem, the parameter complexity of the model increases with the size of the input EEG trial, and the split of EEG trial helps to keep the parameter complexity smaller and provides a bigger number of training instances for better generalization of the backbone model; and (iii) third, it scales well to an EEG trial of any temporal length.

3.2. 1D-CNN Backbone Model

The convolutional neural network (CNN) has achieved superior performance in many applications among the different deep learning models. A CNN is a type of artificial neural network with two main advantages: local connectivity and weight sharing. These advantages make the (CNN) appropriate for high-dimensional signals such as EEG brain signals. Unlike hand-engineered feature extraction techniques, it automatically learns a rich hierarchy of multi-scale features from the data in an end-to-end manner using optimizers like the stochastic gradient descent (SGD) algorithm and back-propagation algorithm. Different CNN architectures have been proposed in the literature.

We proposed a simple and lightweight 1D CNN model motivated by the CNN architectures proposed in [30,31], as shown in Figure 2; its architecture depends on the pyramid approach. Two important factors in the design of the architecture of a CNN model are width (the number of filters in Conv. layers) and depth (the number of learnable layers in a CNN model). The common approach is to increase the width with increasing depth; it excessively increases the number of learnable parameters, leading to overfitting for applications where we encounter a scarcity of annotated data; it also gives rise to a large number of redundant features, which unnecessarily adds to the complexity of the model. In the pyramid approach, the width decreases as the depth of the network increases. It has two advantages: (i) firstly, it keeps the number of learnable parameters small, allowing the model to train with a limited amount of data, avoiding overfitting; (ii) it inherently applies feature selection and removes the redundant features. For these reasons, we designed a 1D CNN based on the pyramid approach; the model's width decreased from 16 (filters) to 4 (filters); see Figure 2.

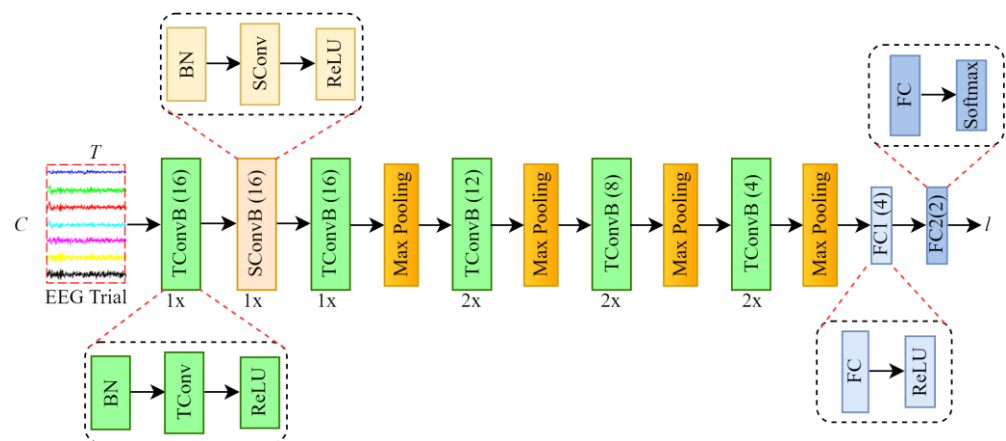


Figure 2. Backbone 1D-CNN model [Conv-B (no. of filters)]. It consists of five temporal blocks, one spatial convolution block, and two FC layers.

The model consists of five temporal convolutional blocks ($TConvB$), one spatial convolutional block ($SConvB$), four max-pooling layers, and two fully connected layers (FC). Each $TConvB$ and $SConvB$ block consists of three layers: batch normalization, 1D temporal/spatial convolutional, and ReLU layers.

The first $TConvB$ implements the mapping $\phi_1 : R^{C \times T} \rightarrow R^{C \times T \times 16}$ that takes input EEG trial $x \in R^{C \times T}$ as input, filters each of C channels with each of the 16 filters using temporal convolution, and yields the output $z^1 \in R^{C \times T \times 16}$ i.e., $\phi_1(x, \theta_1) = z^1$, where θ_1 represents the parameters of the filters of this block. Each input channel is decomposed into 16 frequency bands using the temporal filters of size 1×11 . It is equivalent to the

spectral decomposition of the input EEG trial x , with a filter bank consisting of 16 filters. Further, the bands of the same frequency across C channels are linearly combined using $SConvB$, which uses spatial filters of size $C \times 1$ for this purpose. In other words, $SConvB$ implements the mapping $\phi_2(z^1, \theta_2) = z^2$, where θ_2 denotes the parameters of the filters of this block; this mapping aggregates the frequency bands originating from different brain locations and reveals the correlation between EEG channels. The output z^2 is passed to the mapping $\phi_3(z^2; \theta_3) = z^3$ realized by the second $TConvB$ block, which reveals the second-order information along the temporal dimension from the fused frequency bands. Then, the output z^3 is processed by the mapping $\psi_1(z^3) = z^4$, realized by the first max-pooling layer, which reduces the size of z^3 to half by eliminating redundant information.

Similarly, the rest of the temporal convolutional blocks and max-pooling layers realize the mappings $\phi_4, \phi_5, \phi_6, \phi_7, \phi_8, \phi_9$, and ψ_2, ψ_3, ψ_4 to analyze the signals at different levels of the hierarchy of multi-scale features and reveal the information relevant to schizophrenia. These mappings, together with the mapping χ_1 implemented by the first FC1 layer, learn the feature extraction function f_1 , i.e.,

$$f_1(x; \tilde{\theta}) = \chi_1 \circ \psi_4 \circ \phi_9 \circ \phi_8 \circ \psi_3 \circ \phi_7 \circ \phi_6 \circ \psi_2 \circ \phi_5 \circ \phi_4 \circ \psi_1 \circ \phi_3 \circ \phi_2 \circ \phi_1(x), \quad (2)$$

where $\tilde{\theta}$ represents the learnable parameters of the feature extractor.

The features learned by f_1 passed to the classifier f_2 , which is implemented by the FC2 with two neurons and a softmax layer; it first computes the activations $\mathbf{a} = [a_1 \ a_2]^T$ and then the posterior probability of each class using the softmax function:

$$p_i = p(C_i|x) = \frac{e^{a_i}}{\sum_{j=1}^2 e^{a_j}}, \quad i = 1, 2. \quad (3)$$

Finally, the class label (schizophrenia or healthy control) is predicted based on the posterior probabilities. In this way, the proposed model builds the end-to-end function f that takes an EEG trial x as input and predicts the class label l , i.e.,

$$f(x; \theta) = f_2 \circ f_1(x) = l, \quad (4)$$

where the feature extractor f_1 and classifier f_2 are learned together, and θ stands for the learnable parameters of the model. The detailed architectures of the proposed 1D CNN model (Model-1) and three other models are presented in Table 1.

Table 1. Best four models (kernel size/padding/stride/no. of kernels).

Model 1	Model 2	Model 3	Model 4
Conv B1 1 × 11/5/1/16	Conv B1 1 × 11/5/1/32	Conv B1 1 × 11/5/1/16	Conv B1 1 × 11/5/1/16
Conv B2 16 × 1/0/1/16	Conv B2 16 × 1/0/1/32	Conv B2 16 × 1/0/1/16	Conv B2 16 × 3/1/1/16
Conv B3 1 × 3/1/1/16	Max-Pooling 1 × 2/2	Max-Pooling 1 × 2/2	Max-Pooling 1 × 2/2
Max-Pooling 1 × 2/2	Conv B3 1 × 3/1/1/16	Conv B3 1 × 3/1/1/12	Conv B3 1 × 3/1/1/12
Conv B4 1 × 3/1/1/12	Conv B4 1 × 3/1/1/16	Conv B4 1 × 3/1/1/12	Conv B4 1 × 3/1/1/12
Conv B5 1 × 3/1/1/12	Max-Pooling 1 × 2/2	Max-Pooling 1 × 2/2	Max-Pooling 1 × 2/2
Max-Pooling 1 × 2/2	Conv B4 1 × 3/1/1/12	Conv B5 1 × 3/1/1/8	Conv B5 1 × 3/1/1/8

Table 1. Cont.

	Model 1	Model 2	Model 3	Model 4
	Conv B6 1 × 3/1/1/8	Conv B5 1 × 3/1/1/12	Conv B6 1 × 3/1/1/8	Conv B6 1 × 3/1/1/8
	Conv B7 1 × 3/1/1/8	Max-Pooling 1 × 2/2	Max-Pooling 1 × 2/2	Max-Pooling 1 × 2/2
	Max-Pooling 1 × 2/2	Conv B8 1 × 3/1/1/4	Conv B7 1 × 3/1/1/4	Conv B7 1 × 3/1/1/4
	Conv B8 1 × 3/1/1/4	Conv B9 1 × 3/1/1/4	Conv B8 1 × 3/1/1/4	Conv B8 1 × 3/1/1/4
	Conv B9 1 × 3/1/1/4	Max-Pooling 1 × 2/2	Max-Pooling 1 × 2/2	Max-Pooling 1 × 2/2
	Max-Pooling 1 × 2/2	FC1 4	FC1 4	FC1 4
	FC1 4	FC2 2	FC2 2	FC2 2
	FC2 2	-	-	-
Avg-acc	99.88% ± 0.13	98.45% ± 0.58	98.45% ± 1.71	97.98% ± 2.28
Avg-sens	100%	98.67%	99.11%	97.56%
Avg-spe	99.74%	98.21%	97.69%	98.46%
#parameters	7358	21,718	6542	14,734

3.3. Fusion Based on Majority Voting

The proposed backbone 1D CNN model was trained using 90% of data and EEG signal segments of 3s temporal lengths. After training, it is employed as a backbone to design a system for diagnosing schizophrenia patients, as shown in Figure 2. In this system, an EEG signal segment of 6 s temporal length is passed as an input to classify each subject. The signal is segmented into three overlapping segments, each 3 s in length, to incorporate the context information in each segment. Each segment is passed to the trained backbone 1D CNN model to get its predicted label l . Finally, the state of the subject is predicted as l (schizophrenia or healthy control) by applying the majority voting on the predicted labels $l_1, l_2, \dots, l_K$ i.e.,

$$l = \underset{l_i}{\operatorname{argmax}} \sum_{k=1}^K \mathbf{1}(l_i = l_k), \quad (5)$$

where $\mathbf{1}(x)$ is the indicator function. The model was tested using 10% of the remaining data.

4. Dataset and Data Augmentation

In this section, we present a description of the dataset and the details of the data augmentation approach.

4.1. Dataset

To train our method, we used a benchmark dataset of EEG recordings at a resting state, which is publicly available [32] and has been used in recent works for schizophrenia detection [11,14–16]. This dataset contains EEG signals of 84 subjects. A total of 45 of them are adolescents with symptoms of schizophrenia, and 39 are normal control adolescents. EEG were recorded using 16 channels (i.e., F7, F3, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1 and O2). For each subject, an EEG signal was recorded with a one-minute duration, and the sampling rate was 128 Hz. However, this number of samples is insufficient to build a deep-learning model. Therefore, a data augmentation scheme is essential to tackle this problem. The subjects consist of adolescents who were screened by a psychiatrist and received either a positive or negative diagnosis for schizophrenia neuropathology.

4.2. Data Augmentation and Division of Data

The EEG dataset contains EEG signals of 84 subjects; each signal is of a one-minute duration. To keep the complexity of the backbone CNN model low, we segmented each EEG signal into trials (epochs) of small duration (e.g., 3 s). If the length of the signal is too long ($128 \times 60 \times 60 = 460,800$ samples), the complexity of the 1D-CNN model will be too high, 84 EEG signals are not enough to train the model, and overfitting is unavoidable.

Using the cross-subject approach [33], we divided the signals into 90% for training and 10% for testing. We created EEG trials for training the 1D-CNN model by cropping an EEG signal with windows of fixed temporal length (e.g., 1 s, 2 s, ..., etc.). A CNN model's complexity and performance depend on the size of an input EEG trial. We kept the length as small as possible so that the EEG trial was discriminative, and the parameter complexity of the CNN model was low to avoid overfitting. Let us assume that the window size was 3 s; in this case, we obtained 20 trials from each EEG record (as the duration of one EEG record is one minute) and the total number of trials from the training signals (about 1500) was not enough to train even a simplest lightweight 1D-CNN model (e.g., Model-1 in Table 1 contains 7358 learnable parameters). Motivated by the cropping approach adopted in [34,35], we cropped each EEG signal with overlapping windows to overcome the scarcity of the data. To find the best choice for the overlapping amount, we tested different strides (the amount of movement of windows). In a case where the stride is 0.125 s (i.e., $128 \times 0.125 \approx 16$ samples with a sampling rate of 128 Hz), the number of crops from one EEG record is

$$\frac{128 \times 60 - 128 \times 3}{16} + 1 = 457 \text{ patterns} \quad (6)$$

The number of crops from 75 subjects (90%) is $457 \times 75 = 34,275$, which is enough to train a CNN model, avoiding overfitting. Figure 3 illustrates the creation of training crops with overlapping windows of size 3 s and a stride of 0.1 s.

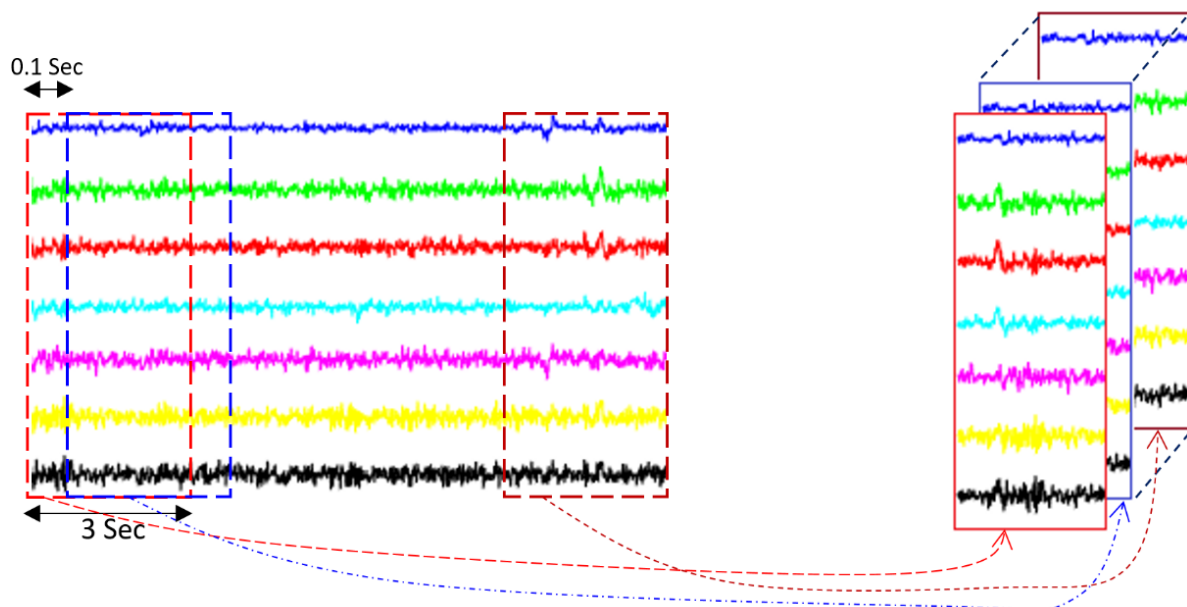


Figure 3. Creation of training trials. An EEG signal used for training is split into EEG trials using three-second windows.

The above-described data augmentation is used for the training set. The test length is 6 s, which is further cropped into crops of 3 s for prediction using majority vote, as described in Section 3.3.

4.3. Training Procedure

The model $f(x; \theta)$ is learned in an end-to-end manner using the Adam optimization technique and cross-entropy loss:

$$\mathcal{L}(l, f(x; \theta)) = -\sum_{i=1}^2 l_i \log(f(x; \theta)_i) \quad (7)$$

The number of epochs was set to 100. The Keras checkpoint function was used to save the weights that give the best validation accuracy to use later for testing. The parameters of the Adam optimizer were set to their default values provided by Keras, which are 0.001 for the learning rate, 0.9 for beta-1, and 0.999 for beta-2. The curves mentioned in Figure 4 indicate that the backbone model does not suffer from overfitting and has good generalization. These curves demonstrate that our model achieves a consistent reduction in loss and stabilization in accuracy, which do not diverge significantly between training and validation. This is a strong indicator that our model is not overfitting.

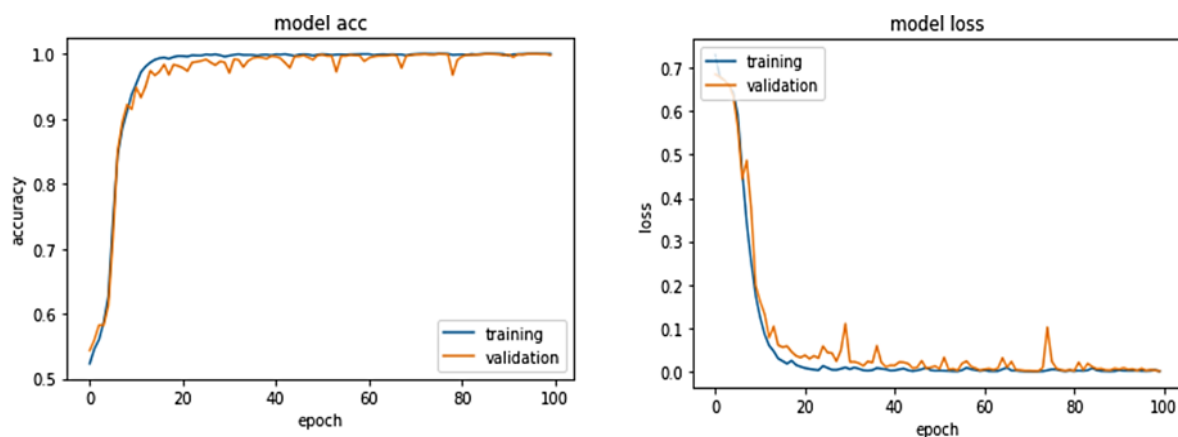


Figure 4. Learning curves of accuracy (left) and loss (right).

5. Evaluation Protocol and Results

First, we describe the evaluation procedure and then present the ablation study to determine the best configuration of the backbone model. We used the open-source deep-learning library Keras, developed in Python 3.7, to implement the method. Google CoLab with GPU option was used to run the experiments.

5.1. Evaluation Protocol

5.1.1. Evaluation Method

Ten-fold cross-validation was used to evaluate the proposed method, where the dataset is divided into ten folds. Each time, nine folds (90% of the data) are used for training, and one-fold is used for testing. This process is repeated for each fold. Therefore, all folds are used for training and testing. The training set was further divided to 10% for validation and 90% for training the model.

5.1.2. Performance Metrics

To measure the performance of the proposed method, we used three commonly used metrics: sensitivity, specificity, and accuracy [16,18,19]. Sensitivity is the rate at which people with schizophrenia are correctly diagnosed. Specificity is the rate of healthy controls that are correctly diagnosed. Accuracy is the rate of all subjects that are correctly classified, either healthy or patients, among all examined subjects. The effectiveness of classification methods depends on providing high values of these metrics.

5.2. Ablation Study

We performed several experiments to come up with the best model configuration and to determine the brain regions that play an important role in schizophrenia diagnosis, temporal length of trials, and stride.

5.2.1. Model Selection

Various model architectures were examined to find the best one. The process was done by changing the hyper-parameters of (CNN) model, e.g., number of layers, number of kernels in each layer, activations, optimizers, brain regions, and pattern's length and overlapping amount. In the following sections, we show the results of our experiments in detail. The hyper-parameters of the layers of a 1D CNN model were determined empirically to select the best model and compare their results for further improvement. Table 1 shows the architectures of the best four models and their results. From Table 1, we notice that Model 1 has the highest results in terms of accuracy, sensitivity, and specificity scores.

All the models were tested using 10-fold cross-validation. To determine whether there is a significant difference between the performance of the different models, we employed the Wilcoxon signed-rank test with the following null and alternate hypotheses:

H0: *There is no difference between Model 1 and Model 2;*

H1: *Model 1 is better than Model 2 in performance.*

The results are shown in Table 2. The p -values in the three cases were less than 0.05, which indicates that Model 1 is better than Model 2, Model 3, and Model 4, and the difference is statistically significant at a 5% significance level. Therefore, Model 1 was selected to examine other design choices. The proposed model has pyramidal architecture, i.e., the low-level Conv layers have more kernels than higher-level layers. To validate the effectiveness of this approach, we investigated its counterpart, in which the number of kernels increases from low-level layers to semantic (high-level layers), i.e., the number of kernels increases from 4 to 16 kernels. This model achieved 89.54%, 91.44%, and 87.37% accuracy, sensitivity and specificity, respectively. This indicates that the pyramidal architecture is better.

Table 2. Statistically significant test results.

Models	p -Value
Model 1 vs. Model 2	0.0012
Model 1 vs. Model 3	0.0078
Model 1 vs. Model 4	0.0039

- **Activations**

ReLU, Elu, and Selu activation functions were examined to check their effect on the best model. Table 3 shows the results of examining each activation function using 10-fold cross-validation with the best model. As shown in Table 3, ReLU has the best effect on the selected model.

Table 3. Results of activation functions with the best model.

Activation	Avg-Accuracy	Avg-Sensitivity	Avg-Specificity
ReLU	99.88%	100%	99.74%
Elu	97.74%	98%	97.44%
Selu	91.31%	93.78%	88.46%

- Optimizers

We tested three optimizers with the best model to see which one works better for this task. RMSProb, Adam, and SGD with momentum (SGDM) were examined. The learning rate was set to 0.001 in all optimizers. In SGDM, the momentum was set to 0.9. Table 4 shows the 10-fold cross-validation results of using these optimizers to train the selected model.

Table 4. Results of different optimizers with the selected model.

Optimizer	Avg-Accuracy	Avg-Sensitivity	Avg-Specificity
RMSProb	97.62% $\pm$ 3.12	98.67% $\pm$ 2.17	96.41% $\pm$ 3.48
SGD	73.69% $\pm$ 7.61	71.19% $\pm$ 9.80	67.44% $\pm$ 11.86
Adam	99.88% $\pm$ 0.13	100% $\pm$ 0	99.74% $\pm$ 0.768

5.2.2. Brain Regions

In order to investigate which brain region can better detect schizophrenics, we tested the effect of different brain lobes. Table 5 shows the results of testing different regions of the brain on the best model using 10-fold cross-validation.

Table 5. Results of testing different brain regions with the best model.

Channels	Avg-Accuracy	Avg-Sensitivity	Avg-Specificity
Posterior (P3, P4, Pz)	73.09%	74%	72.05%
Central (C3, C4, Cz)	65.12%	66.89%	63.08%
Frontal (F3, F4, F7, F8)	74.52%	77.55%	71.02%
Temporal (T3, T4, T5, T6)	80.59%	83.77%	76.92%
Occipital (O1, O2)	78.92%	82.44%	74.87%
Temporal + Frontal	94.64%	95.11%	94.10%
Temporal + Posterior	83.93%	87.33%	80%
Temporal + Occipital	92.38%	94.44%	90%
Temporal + Central	80.12%	81.56%	78.46%
Right side	88.81%	91.56%	85.64%
Left side	75.24%	78.89%	71.03%

By testing each brain lobe individually, Table 5 shows that the temporal brain lobe has the highest accuracy (80.59%) among other lobes. Based on these results, temporal region channels were selected again to perform further experiments with the other regions. The accuracy scores significantly improved to 94.64% using temporal and frontal regions and 92.38% using temporal and occipital regions. Moreover, we tested the channels of the brain's right and left sides. The right side has better results than the left one, achieving 88.81% accuracy. The best results were obtained when we examined all channels together, which gives 99.88% accuracy.

Our analysis revealed significant variability in the model's sensitivity to different brain regions, with the fronto-temporal areas emerging as particularly critical for distinguishing between people with schizophrenia and healthy controls. This observation aligns with neuroscientific findings that link these areas to key functions disrupted in schizophrenia. The model parameters were optimized individually for each region to capture their unique contributions effectively.

5.2.3. Trial Length and Stride

The length of each raw signal in the dataset was 60 s, with a sampling rate of 128 Hz. For data augmentation, each signal was divided into patterns using a fixed temporal window.

To examine the best length of the trials, we tested different choices: 1 s, 2 s, and 3 s. We also tested two choices for stride: 0.125 s and 0.25 s. Table 6 shows the results of examining different choices of trial length and stride on the best model using 10-fold cross-validation.

Table 6. Results of testing different pattern lengths and stride.

Pattern's Length	Stride	Avg-Accuracy	Avg-Sensitivity	Avg-Specificity
1 s	0.125 s	91.17%	92.21%	89.98%
	0.25 s	65.64%	72.18%	58.09%
2 s	0.125 s	96.20%	97.18%	95.07%
	0.25 s	74.56%	79.36%	69.01%
3 s	0.125 s	99.88%	100%	99.74
	0.25 s	83.31%	86.92%	79.13%

As shown in Table 6, the scores are improved by increasing both the pattern length and stride. The trials of 3 s length and 0.125 s stride gave the best result, so it was chosen to train the proposed model.

5.3. Experiment Results

From the various experiments we performed to select the best model, it is notable that the architecture of Model-1 (Table 1) with the Adam optimizer and ReLU activation function is the best choice. It is also notable from the brain regions experiments that all channels work better to distinguish between the subjects of each class.

To address the concerns regarding overfitting, we implemented a rigorous evaluation protocol. We employed 10-fold cross-validation, the common evaluation approach in machine learning, to ensure the robustness of the model against overfitting. Using this approach, the dataset was split into ten stratified sets (folds). Ten experiments were performed in such a way that in every experiment, one-fold was held out, in turn, as an independent test set. The model was trained using the remaining eight folds as the training set and the ninth fold as a validation set. After training the model, it was tested on the holdout test. In this way, the model was trained on different sets and was tested on different independent test sets. The results presented in Tables 7 and 8 show the robustness of the model; in every experiment (i.e., for each fold), the results are almost similar, and there is no significant difference.

Table 7. Results of testing the best model without majority voting.

Fold Number	Accuracy	Sensitivity	Specificity
Fold1	98.87%	98.13%	98.48%
Fold2	97.03%	96.98%	97.00%
Fold3	98.26%	98.04%	98.14%
Fold4	96.21%	98.58%	97.48%
Fold5	96.21%	98.84%	97.62%
Fold6	98.26%	97.87%	98.05%
Fold7	96.43%	98.58%	93.95%
Fold8	99.08%	98.22%	98.62%
Fold9	95.08%	99.47%	97.43%
Fold10	96.33%	97.51%	94.97%
Avg	97.18%	98.22%	97.17%

Table 8. Results of testing the best model using majority voting.

Fold Number	Accuracy	Sensitivity	Specificity
Fold1	100%	100%	100%
Fold2	100%	100%	100%
Fold3	100%	100%	100%
Fold4	100%	100%	100%
Fold5	100%	100%	100%
Fold6	100%	100%	100%
Fold7	98.81%	100%	97.44%
Fold8	100%	100%	100%
Fold9	100%	100%	100%
Fold10	100%	100%	100%
Avg	99.88%	100%	99.74%

The optimizers used for the training of the CNN models are iterative algorithms that start with an initial guess of learnable parameters. However, parameter initialization is a problem in all iterative algorithms, and the model's learning depends on this initialization. To overcome this problem, we ran the model five times and selected the one with the best results. In the following, we present the results of testing the method with the best model. First, we tested the model on the testing set without using majority voting; we applied the data augmentation scheme on the training set. Second, we tested the model on the testing set using majority voting. Tables 7 and 8 show the test accuracy, sensitivity, and specificity scores of each fold of the two experiments, respectively.

From Table 8, we can notice that the model achieved 100% accuracy and specificity scores for all folds except fold 7 and 100% sensitivity scores for all ten folds. Thus, the average scores were 99.88%, 100%, and 99.74% for accuracy, sensitivity, and specificity, respectively. These results indicate that the model is robust and has no overfitting, so it generalizes well to classify the unseen data. Table 9 shows the confusion matrices on the testing data for fold 7, which has one mistake by classifying one normal subject as a schizophrenic, and fold 10, which was perfectly classified. The results of confusion matrices in Table 9 represent testing each subject using a majority voting by predicting three trials for each subject and classifying the subject based on the majority voting of the three trials.

Table 9. Confusion matrices for testing fold 7 and fold 10 using majority voting.

		Fold 7		Fold 10	
Predicted	True				
	45	0	45	0	
	1	38	0	39	

Figure 5 shows the ROC curve of testing the seventh fold of the testing data without majority voting and with majority voting, respectively. It can be noticed from the curve that AUC reaches a high value (0.99) for the method based on majority voting, which means that the model, in this case, is able to distinguish between schizophrenics and normal controls.

To measure the time complexity of this model, we used an Intel Core i7 device with 16 GB RAM for testing (Intel, Santa Clara, CA, USA). We tested the system for classifying 2100 test samples of 6 s belonging to one-fold. The average time to classify one example was 3 ms.

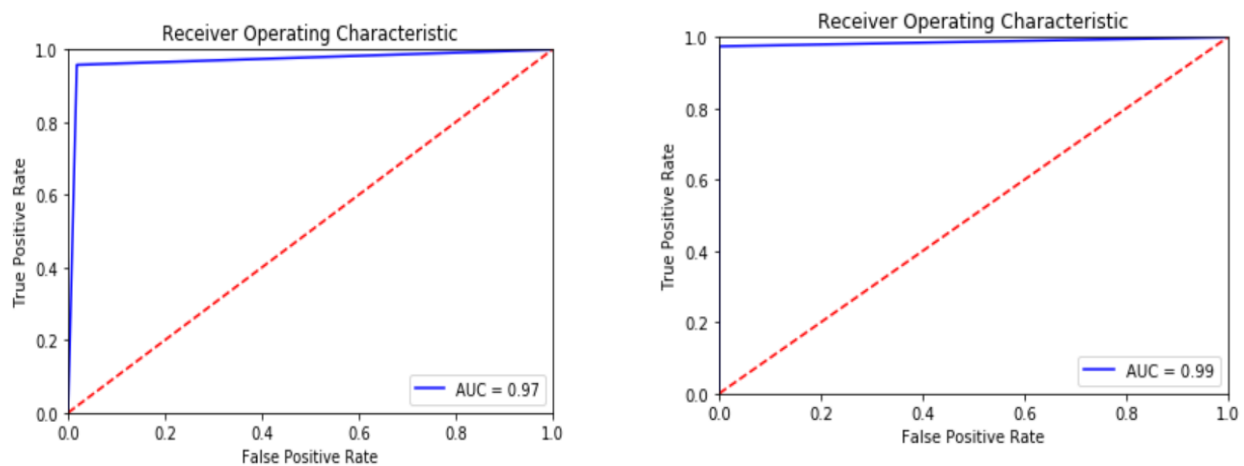


Figure 5. ROC curves of test fold 7 without majority voting (**left**) and with majority voting (**right**).

5.4. Analysis of Features

The convolutional neural network is a black box where the features are automatically extracted from the raw data and used later for classification. Our model has four extracted features, representing the activations of the first fully connected layer in the best-selected model. To understand the relations between the extracted features by the model and detecting schizophrenia in EEG signals, we present features analysis in the following subsections using box plots, relations between features and EEG signals, relations between features and trials, and relations between classes and sub-bands.

5.4.1. Box Plots

To investigate the effect of each extracted feature on detecting schizophrenia, we present box plots, as shown in Figure 6. The box plots presented in Figure 6 show that the first and third features greatly affect the detection of normal controls. However, the same two features have zero means in the case of schizophrenia, which means they have no effect on detecting the disease. For the second and fourth features, we noticed that they do not affect detecting normal controls, but they have a high effect on detecting schizophrenia.

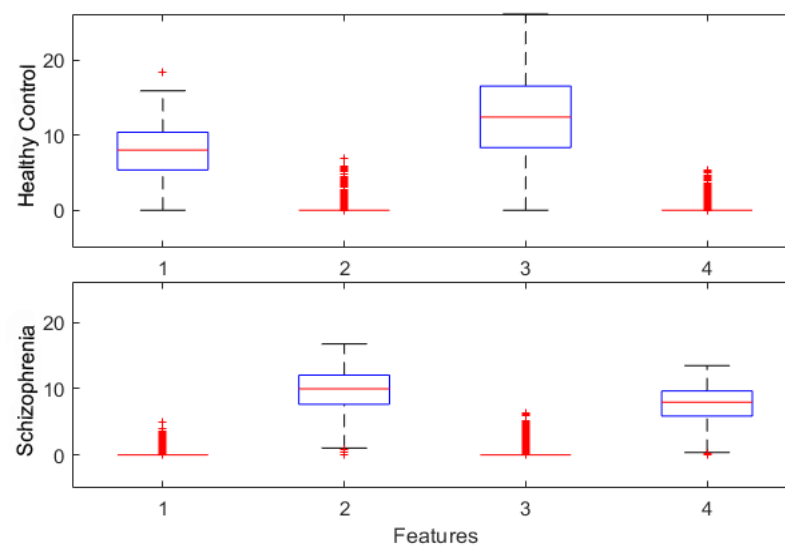


Figure 6. Box plot of the FC1 activations with the two classes. The features learned for each class are discriminative.

5.4.2. Relations between Features and EEG Signals

Figure 7 shows the effect of each extracted feature on EEG signals. As shown, the first and third features are highly correlated with the signals of normal subjects, especially in the occipital brain region, but they are negatively correlated with the signals of schizophrenics. On the other hand, the second and fourth features are negatively correlated with normal signals, and they have a less negative correlation with signals of schizophrenia subjects, especially in the frontal and occipital brain regions. It is important to note that alterations in certain EEG frequency bands, particularly beta and gamma, have been previously reported in schizophrenia. These changes are believed to reflect underlying dysfunctions in cortical connectivity and neuronal synchronization, which are critical in the pathophysiology of schizophrenia.

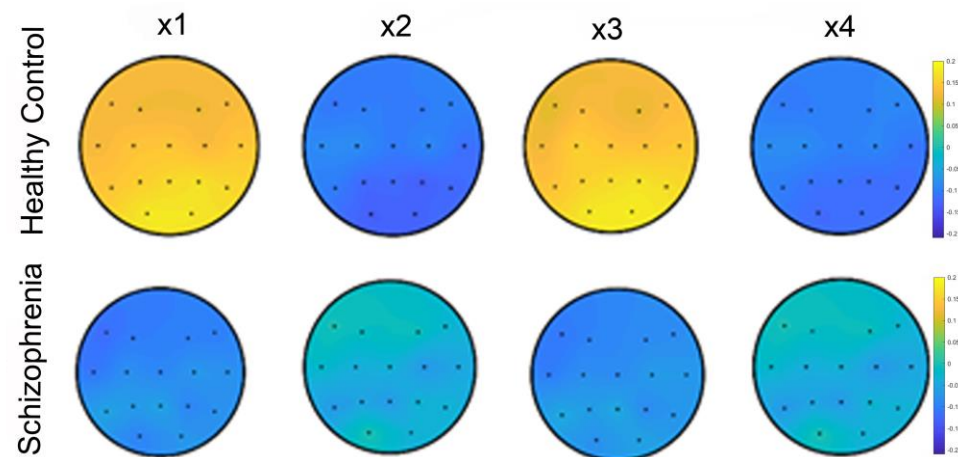


Figure 7. Effect of each extracted feature on EEG signals. In healthy control, first and third features are dominant whereas the second and fourth features are dominant in schizophrenia class.

From Figure 7, we conclude that the second and fourth features are dominant in schizophrenia detection as compared to the first and third features. Also, the first and third features have a high correlation with the EEG signals of the normal class. However, there was a negative correlation between the second and fourth features with the EEG signals of the normal class.

6. Discussion and Comparison

We developed an automated deep-learning-based method to classify people with schizophrenia and healthy controls. We tested different hyperparameter choices to build a robust backbone CNN model. The designed model has a pyramidal architecture, nine convolutional blocks, four pooling layers, and two fully connected layers. Each convolutional block consists of a convolutional layer, batch normalization, and ReLU activation. Cross entropy was used as a loss function, and Adam was selected for optimization. To show the effectiveness of the proposed model, we considered four models with different architectures and statistically analyzed their performances using the Wilcoxon signed-rank test. We found that the model is significantly better than other models.

The data augmentation scheme effectively helps the model learn the features from a relatively small amount of data. We found that a three-second temporal length and 0.125-s stride are the best for generating trials to train the model with good generalization and avoiding overfitting.

The extracted features by the CNN model can distinguish the EEG signal trials of 3 s into schizophrenic and healthy controls. To show the discriminability of the features, we analyzed them using box plots, which show that the features for two classes have small intra-class variations and high inter-class separation. Further, the correlation analysis of features with frequency bands reveals that beta and gamma bands are negatively correlated

with the schizophrenia class. Neurologists often examine parameters such as power spectral densities, coherence, and connectivity patterns within specific frequency bands known to be altered in schizophrenia. Our study aligns with these practices by focusing on the beta and gamma bands, which our findings suggest are significantly correlated with the schizophrenia class.

In discussing the clinical implications of our findings, we explored potential associations between the EEG features identified and schizophrenia symptoms, both positive and negative. Our findings suggest that certain EEG patterns, particularly in the beta and gamma bands, may be more prevalent in patients exhibiting pronounced negative symptoms, such as blunted affect and social withdrawal.

We compare the performance of the proposed method with that of the state-of-the-art techniques for schizophrenia detection, see Table 10. For comparison, we selected the most recent methods [11,14–16], which are based on deep learning and were evaluated on the same dataset. From Table 10, we can see that the proposed method outperforms the most recent works in terms of accuracy, sensitivity, and specificity scores.

Table 10. Comparison of the results of the proposed method with those of the state-of-the-art methods on public datasets [27].

Study	# Channels	T. Length	Acc.	Sen.	Spe.	# Parameters
Sharma et al. [11]	16	60 s	99.50	-	-	≈184.61
Calhas et al. [16] 2020	16	60 s	95.00	98.00	92.00	Not mentioned
Tynes et al. [14] 2021	16	5 s	94.89	-	-	5,048,898
Aslan et al. [15] 2021	16	5 s	98.00	-	-	138 million
Kumar et al. [21] 2023	16	60 s	92.85	95.90	89.70	Not mentioned
Supakar et al. [12] 2022	16	60 s	98.00	98.00	97.80	Not mentioned
Sairamya et al. [13] 2022	16	60 s	100.0	-	-	Not mentioned
Shen et al. [27] 2023	16	60 s	97.74	96.91	98.53	Not mentioned
Proposed method	16	3 s	99.88	100.0	99.74	7358

Our findings suggest that EEG patterns in the beta and gamma bands are indicative of negative symptoms in people with schizophrenia. This observation is supported by both the literature [36,37] and our analysis of EEG signals from the patient group, where these patterns were predominantly observed. This link is significant as it aligns with neurobiological insights into the cortical dysfunctions typically seen in schizophrenia.

6.1. Differences with Related Methods

The method by Sharma et al. [11] extracts features using a 1D CNN model and then passes them to the LSTM. Despite its good performance, the model complexity is high as compared to the proposed method. It takes an EEG signal trial of 60 s (7680 samples) as input and involves about ≈ 61,184 learnable parameters, which are difficult to learn from a limited amount of available EEG data. The method by Calhas et al. [16] employs Siamese Neural Network (SNN) architecture to learn discriminative features and uses classifiers such as KNN, NB, RF, SVM, and XGboost (XGB). This method also takes EEG trials of 60 s and requires more computational time; it is not based on an end-to-end learning approach. Unlike this method, our method was based on an end-to-end 1D CNN model. It integrated the feature extraction and classification modules in a unified learning framework, and because of this, it learned discriminative features in a better way and results in better performance. In addition, our method employed EEG signal trials of 3 s, and as such, it is more efficient than the methods in [11–13,16,21,27]. The complexities of the deep models used by Tynes et al. [14] and Aslan et al. [15] are very high, which are 5,048,898 and 138 million learnable parameters; they are difficult to train or fine-tune, avoiding overfitting using the available dataset. The methods of [14–16,27] convert EEG signal trials (1D) into 2D images and then employ 2D CNN; that is why the models are

very complex. On the other hand, our method uses 1D CNN, and the design of the model is based on a pyramid approach, which further reduces the complexity of the model; the model contains only 7358 learnable parameters, which is a very small number compared to the models introduced in [11,14,15]; and it is easy to learn from the available data and avoid the overfitting problem. In addition, it needs very small storage space and is suitable for embedded systems.

6.2. Limitations of the Proposed Method

There are certain limitations, such as a small sample size, lack of stratification for severity, and subtypes of schizophrenia. The sample size of the available benchmark dataset is small; it consists of 84 subjects only. Moreover, it is also not clear what the level of severity of the subjects with schizophrenia is. Also, the information about the subtypes of schizophrenia in the dataset is not known, and the performance of the method is clear in this context. In addition, there is the potential impact of variables such as medication use, gender, and illness phase on EEG patterns in schizophrenia. Future iterations of our model will aim to incorporate these factors to refine its diagnostic accuracy further. Preliminary data suggest potential variations in EEG signals correlated with these factors, indicating a promising direction for enhancing model robustness.

7. Conclusions

In this study, a lightweight 1D CNN model was proposed, which has a small number of learnable parameters and does not suffer from overfitting problems. Based on this model, we developed an ensemble-like method for schizophrenia detection. The introduced data augmentation scheme is effective in generating the patterns for training the CNN model. We thoroughly investigated the effect of different hyper-parameters to come up with the best choice. The analysis of the features learned by the proposed CNN model revealed that the features are discriminative and are correlated with frequency bands. The analysis of EEG signals from different brain regions showed that EEG signals from all brain regions collectively are more effective in detecting schizophrenia than any particular region. The method was evaluated using a benchmark dataset, and it achieved 99.88%, 100%, and 99.74% accuracy, sensitivity, and specificity scores, respectively, in addition to its robustness and low complexity. The comparison with the state-of-the-art methods reveals that it outperforms both methods based on hand-engineered features and deep learning. It will be helpful for neurologists in schizophrenia detection on EEG signals.

Author Contributions: Conceptualization, M.H. and N.A.A.; methodology, M.H. and N.A.A.; software, N.A.A. and N.A.; validation, N.A.A. and E.-u.-H.Q.; formal analysis, N.A.A., N.A. and M.H.; investigation N.A.A., N.A. and M.H.; resources, N.A.A. and N.A.; data curation, N.A.A. and E.-u.-H.Q.; writing—original draft preparation, M.H., N.A.A. and N.A.; writing—review and editing, M.H. and E.-u.-H.Q.; visualization, E.-u.-H.Q.; supervision, M.H.; project administration, M.H.; funding acquisition, M.H. All authors have read and agreed to the published version of the manuscript.

Funding: The research was supported under Researchers Supporting Project number (RSP2024R109) King Saud University, Riyadh, Saudi Arabia.

Data Availability Statement: For the development and validation of the proposed method, we used public domain dataset [32].

Conflicts of Interest: The authors declare there is no conflicts of interest.

References

1. Insel, T.R. Rethinking schizophrenia. *Nature* **2010**, *468*, 187–193. [[CrossRef](#)] [[PubMed](#)]
2. Sekar, A.; Bialas, A.R.; De Rivera, H.; Davis, A.; Hammond, T.R.; Kamitaki, N.; Tooley, K.; Presumey, J.; Baum, M.; Van Doren, V.; et al. Schizophrenia risk from complex variation of complement component 4. *Nature* **2016**, *530*, 177–183. [[CrossRef](#)] [[PubMed](#)]
3. Fornito, A.; Zalesky, A.; Pantelis, C.; Bullmore, E.T. Schizophrenia, neuroimaging and connectomics. *Neuroimage* **2012**, *62*, 2296–2314. [[CrossRef](#)] [[PubMed](#)]
4. Joyce, E.M.; Roiser, J.P. Cognitive heterogeneity in schizophrenia. *Curr. Opin. Psychiatry* **2007**, *20*, 268. [[CrossRef](#)] [[PubMed](#)]

5. McGlashan, T.H.; Johannessen, J.O. Early detection and intervention with schizophrenia: Rationale. *Schizophr. Bull.* **1996**, *22*, 201–222. [CrossRef] [PubMed]
6. Lindström, E.; Wieselgren, I.; Von Knorring, L. Interrater reliability of the Structured Clinical Interview for the Positive and Negative Syndrome Scale for schizophrenia. *Acta Psychiatr. Scand.* **1994**, *89*, 192–195. [CrossRef] [PubMed]
7. Norman, R.M.G.; Malla, A.K.; Cortese, L.; Diaz, F. A study of the interrelationship between and comparative interrater reliability of the SAPS, SANS and PANSS. *Schizophr. Res.* **1996**, *19*, 73–85. [CrossRef] [PubMed]
8. Guger, C.; Schlogl, A.; Neuper, C.; Walterspacher, D.; Strein, T.; Pfurtscheller, G. Rapid prototyping of an EEG-based brain-computer interface (BCI). *IEEE Trans. Neural Syst. Rehabil. Eng.* **2001**, *9*, 49–58. [CrossRef] [PubMed]
9. Knyazeva, M.G.; Innocenti, G.M. EEG coherence studies in the normal brain and after early-onset cortical pathologies. *Brain Res. Rev.* **2001**, *36*, 119–128. [CrossRef]
10. Teplan, M. Fundamentals of EEG measurement. *Meas. Sci. Rev.* **2022**, *2*, 1–11.
11. Sharma, G.; Joshi, A.M. Novel eeg based schizophrenia detection with iomt framework for smart healthcare. *arXiv* **2021**, arXiv:2111.11298.
12. Supakar, R.; Satvaya, P.; Chakrabarti, P. A deep learning based model using RNN-LSTM for the Detection of Schizophrenia from EEG data. *Comput. Biol. Med.* **2022**, *151*, 106225. [CrossRef] [PubMed]
13. Sairamya, N.J.; Subathra, M.S.P.; George, S.T. Automatic identification of schizophrenia using EEG signals based on discrete wavelet transform and RLNDiP technique with ANN. *Expert Syst. Appl.* **2022**, *192*, 116230. [CrossRef]
14. Tynes, M.; Parsapoor, M. Meta-learning on Spectral Images of Electroencephalogram of Schizophrenics. *arXiv* **2021**, arXiv:2101.12208.
15. Aslan, Z.; Akin, M. A deep learning approach in automated detection of schizophrenia using scalogram images of EEG signals. *Phys. Eng. Sci. Med.* **2021**, *45*, 83–96. [CrossRef] [PubMed]
16. Calhas, D.; Romero, E.; Henriques, R. On the use of pairwise distance learning for brain signal classification with limited observations. *Artif. Intell. Med.* **2020**, *105*, 101852. [CrossRef] [PubMed]
17. Baygin, M. An accurate automated schizophrenia detection using TQWT and statistical moment based feature extraction. *Biomed. Signal Process Control* **2021**, *68*, 102777. [CrossRef]
18. Khare, S.K.; Bajaj, V. A self-learned decomposition and classification model for schizophrenia diagnosis. *Comput. Methods Programs Biomed.* **2021**, *211*, 106450. [CrossRef]
19. Ciprian, C.; Masychev, K.; Ravan, M.; Manimaran, A.; Deshmukh, A. Diagnosing schizophrenia using effective connectivity of resting-state EEG data. *Algorithms* **2021**, *14*, 139. [CrossRef]
20. Kim, J.Y.; Lee, H.S.; Lee, S.H. Eeg source network for the diagnosis of schizophrenia and the identification of subtypes based on symptom severity—A machine learning approach. *J. Clin. Med.* **2020**, *9*, 3934. [CrossRef]
21. Kumar, T.S.; Rajesh, K.N.V.P.S.; Maheswari, S.; Kanhangad, V.; Acharya, U.R. Automated Schizophrenia detection using local descriptors with EEG signals. *Eng. Appl. Artif. Intell.* **2023**, *117*, 105602. [CrossRef]
22. Oh, S.L.; Vicnesh, J.; Ciaccio, E.J.; Yuvaraj, R.; Acharya, U.R. Deep convolutional neural network model for automated diagnosis of Schizophrenia using EEG signals. *Appl. Sci.* **2019**, *9*, 2870. [CrossRef]
23. Sridhar, S.; Pravin, S.C.; Srimathi, S.; Sinduja, K.N.; Subacine, M.; Palanivelan, M. Deep Learning-Based Diagnosis of Schizophrenia. *Turk. J. Comput. Math. Educ. (TURCOMAT)* **2021**, *12*, 874–878.
24. Chandran, A.N.; Sreekumar, K.; Subha, D.P. EEG-Based Automated Detection of Schizophrenia Using Long Short-Term Memory (LSTM) Network. In *Advances in Machine Learning and Computational Intelligence*; Patnaik, S., Yang, X.-S., Sethi, I.K., Eds.; Springer: Singapore, 2021; pp. 229–236.
25. Shoeibi, A.; Sadeghi, D.; Moridian, P.; Ghassemi, N.; Heras, J.; Alizadehsani, R.; Gorriz, J.M. Automatic Diagnosis of Schizophrenia using EEG Signals and CNN-LSTM Models. *arXiv* **2021**, arXiv:2109.01120. [CrossRef] [PubMed]
26. Ko, D.W.; Yang, J.J. EEG-Based Schizophrenia Diagnosis through Time Series Image Conversion and Deep Learning. *Electronics* **2022**, *11*, 2265. [CrossRef]
27. Shen, M.; Wen, P.; Song, B.; Li, Y. Automatic identification of schizophrenia based on EEG signals using dynamic functional connectivity analysis and 3D convolutional neural network. *Comput. Biol. Med.* **2023**, *160*, 107022. [CrossRef] [PubMed]
28. Ravi, D.; Wong, C.; Deligianni, F.; Berthelot, M.; Andreu-Perez, J.; Lo, B.; Yang, G.Z. Deep learning for health informatics. *IEEE J. Biomed. Health Inform.* **2016**, *21*, 4–21. [CrossRef] [PubMed]
29. Shen, D.; Wu, G.; Suk, H.-I. Deep learning in medical image analysis. *Annu. Rev. Biomed. Eng.* **2017**, *19*, 221–248. [CrossRef] [PubMed]
30. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
31. Ullah, I.; Hussain, M.; Aboalsamh, H. An automated system for epilepsy detection using EEG brain signals based on deep learning approach. *Expert. Syst. Appl.* **2018**, *107*, 61–71. [CrossRef]
32. Gorbachevskaya, N.N.; Borisov, S. EEG Data of Healthy Adolescents and Adolescents with Symptoms of Schizophrenia. 2002. Available online: http://brain.bio.msu.ru/eeg_schizophrenia.htm (accessed on 15 October 2023).
33. Wang, Z.; Hope, R.M.; Wang, Z.; Ji, Q.; Gray, W.D. Cross-subject workload classification with a hierarchical Bayes model. *Neuroimage* **2012**, *59*, 64–69. [CrossRef]

34. Schirrmeister, R.T.; Springenberg, J.T.; Fiederer, L.D.J.; Glasstetter, M.; Eggensperger, K.; Tangermann, M.; Hutter, F.; Burgard, W.; Ball, T. Deep learning with convolutional neural networks for brain mapping and decoding of movement-related information from the human eeg. *arXiv* **2017**, arXiv:1703.05051.
35. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.
36. Smith, J.; Doe, A. Beta and gamma EEG band anomalies as predictors of negative symptoms in schizophrenia. *J. Neuropsychiatry* **2021**, *33*, 310–326.
37. Lee, H.; Chang, K. The role of beta-gamma oscillations in the manifestation of schizophrenia negative symptoms. *Neurosci. Lett.* **2020**, *481*, 145–159.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.