



# A deep learning approach in automated detection of schizophrenia using scalogram images of EEG signals

Zülfikar Aslan<sup>1</sup> · Mehmet Akin<sup>2</sup>

Received: 2 September 2021 / Accepted: 15 November 2021 / Published online: 25 November 2021  
© Australasian College of Physical Scientists and Engineers in Medicine 2021

## Abstract

This study presents a method with high accuracy performance that aims to automatically detect schizophrenia (SZ) from electroencephalography (EEG) records. Unlike related literature studies using traditional machine learning algorithms, the features required for the training of the network are automatically extracted from the EEG records in our method. In order to obtain the time frequency features of the EEG signals, the signal was converted into 2D by using the Continuous Wavelet Transform method. This study has the highest accuracy performance in the relevant literature by using 2D time frequency features in automatic detection of SZ disease. It is trained with Visual Geometry Group-16 (VGG16), an advanced convolutional neural networks (CNN) deep learning network architecture, to extract key features found on scalogram images and train the network. The study shows a high success in classifying SZ patients and healthy individuals with a very satisfactory accuracy of 98% and 99.5%, respectively, using two different datasets consisting of individuals from different age groups. Using different techniques [Activation Maximization, Saliency Map, and Gradient-weighted Class Activation Mapping (Grad-CAM)] to visualize the learning outcomes of the CNN network, the relationship of frequency components between SZ and the healthy individual is clearly shown. Moreover, with these interpretable outcomes, the difference between SZ patients and healthy individuals can be distinguished very easily help for expert opinion.

**Keywords** CWT · Deep learning · EEG · Scalogram · Schizophrenia

## Introduction

Schizophrenia (SZ) is a mental disorder that has chronic and severe effects and affects approximately 20 million people worldwide [1, 2]. It is a long-term brain disorder with symptoms such as strange fixed beliefs (delusions) that are not correct, hearing or seeing things that do not exist (hallucinations), irregular emotions, perception or speech [3]. Compared to normal healthy people, SZ patients have a higher mortality rate due to preventable physical illnesses [4].

Detection of SZ disease is usually carried out by a qualified psychiatrist based on patient interviews and monitoring

patient behavior [5]. Since SZ shows symptoms that can be confused with different psychological disorders, the detection phase is not completely reliable, as it is possible for the expert to make mistakes in some cases. Although SZ is usually detected by a specialist, it has recently been used to detect automatically with tools such as electroencephalography (EEG) and neuroimaging techniques such as MRI, CT, FMRI, and PET. Neuroimaging techniques are disadvantageous in that they require high cost, high computation time, and extra recording compared to EEG recordings [3, 6–9]. Since EEG is less costly and more practical, the use of EEG in SZ detection is more widely preferred. The use of EEG data in detecting SZ is a less costly and more practical method.

EEGs are devices that allow the activity of electrical signals in the brain to be recorded. EEG recordings contain data of the signals received from electrodes placed on the scalp that vary depending on periodic or non-stationary time [10, 11]. EEG data are used as an important material that enables the study and analysis of brain activities.

✉ Zülfikar Aslan  
zulfikaraslan@gantep.edu.tr; zulfikaraslan27@gmail.com  
Mehmet Akin  
makin@dicle.edu.tr

<sup>1</sup> Technical Sciences Vocational School, Gaziantep University, Gaziantep 27310, Turkey

<sup>2</sup> Electrical-Electronics Engineering, Faculty of Engineering, Dicle University, Diyarbakir 21280, Turkey

Many linear and nonlinear analysis methods are used in the processing of EEG signals [12–16]. In addition, there are many literature studies presenting signal analysis methods for the removal of EEG artifacts that occur during the recording phase of EEG signals [17–22]. Early detection of SZ disease can help reduce brain impairments, although detection of the disease is difficult, even for a specialist [23]. For these reasons, the development of computer-aided software has been an important research topic to assist the expert in decision making. In the literature, there are many CAD studies using traditional machine learning algorithms [24, 25]. The recent developments in Deep Learning (DL), which gives quite satisfactory results especially in two-dimensional data compared to traditional machine learning algorithms, have been an important reason for researchers to focus on this field.

The signals of EEG data are usually digitized and analyzed with computer software. Thus, this information, which is difficult to interpret directly by any expert, becomes more understandable. For this purpose, there are CAD (Computer Aided Diagnosis) systems developed to assist the expert in the evaluation of disorders that cause brain damage such as SZ or any disease-related abnormality. EEG data are used as input data to the CAD system developed to detect many diseases such as SZ disease [3]. Therefore, CAD systems using EEG signals developed to detect SZ are widely used in the relevant literature. In most of the literature studies in this field, the disease has been tried to be detected with traditional machine learning algorithms by using the features obtained from EEG signals. High accuracy detection of the disease depends on the features obtained from the signal.

It is known that the Wavelet method extracts the important features of the time frequency components of the signal in more detail. For this reason, there are many studies using this method together with traditional machine learning algorithms. It is used extensively in the analysis of many different biomedical signals in the literature and produces satisfactory results. In our study in this direction, we propose a method that tries to detect SZ with much higher classification performance, simple pipeline and interpretable results help for expert opinion. The outputs obtained can be interpreted very easily help for any expert opinion. The proposed study is also important in evaluating the time frequency features of an EEG. There is no need for any processing or material other than EEG recordings to detect SZ.

2D scalogram images are created by applying Continuous Wavelet Transform (CWT) to the raw EEG data. Time frequency features are obtained from the scalogram images. In the relevant literature, scalogram images used as input data to the CNN network for the detection of SZ disease are used for the first time. Important features on the scalogram images that are input to the CNN model are automatically extracted from the next layers of the CNN network. In the

method we propose, it is more advantageous than many literature methods that require extensive knowledge in the field since important features are automatically extracted from CNN layers. In addition, the method we propose has advantages such as more easily interpretable results and learning process when compared to the literature methods using CNN. The proposed method has also verified its success in detecting SZ by using two different datasets, each containing patients from different age groups and healthy controls. The fact that the proposed method reaches high performance values (98% and 99.5%) in both children and adult data in automatic detection of SZ shows that it is a reliable method. Moreover, the classification performance achieved has a higher performance value than most of the literature methods. Another advantage of the method compared to other studies is that since it uses scalogram images as input, the model produces more easily interpretable outputs without requiring expert opinion, with images such as Activation Maximization, Saliency Map and Grad-CAM, which reveal the relationship between disease and frequency components.

To summarize the main contributions of the proposed method:

1. A deep learning-based approach is presented for the detection of schizophrenia with scalogram images obtained from EEG signals.
2. The proposed method distinguishes schizophrenic and healthy individuals with high classification performance.
3. Interpretable outputs that are thought to support expert opinion in the detection of schizophrenia have been obtained.

## Literature review

There are many studies presented for the detection of SZ disease by using different feature extraction methods and using different machine learning algorithms.

Kim et al. [26] selected 5 frequency bands for recordings from a 21-electrode EEG device. By applying the Fast Fourier Transform (FFT) to these 5 selected frequency bands, they calculated the spectral power of these bands with EEGLAB software [27]. Using the delta frequency, they classified SZ and healthy controls with 62.2% accuracy. Additionally, Dvey-Aharon et al. [28] preprocessed the EEG signals and performed feature extraction by Stockwell transform [29]. Using the TFFO (Time–Frequency conversion and then Feature Optimization) method, they achieved a classification performance between 92 and 93.9%. Johannesen et al. [11] Support vector machines (SVM) were used to extract the most important features [23] from EEG recordings to predict the working memory performance of subjects. Using the features, they obtained, they classified them with Support Vector Machines (SVM) and obtained an

accuracy of 87%. Besides, Santos-Mayo et al. [30] studied different machine learning techniques (electrode grouping, feature selection algorithms, and filtering). They performed the classification process with the SVM and Multi-Layer Perceptron (MLP) algorithms with 92.23% and 93.42% accuracy, respectively. Finally, they classified the features they obtained by applying the J5 feature selection algorithm [31] and observed that they showed higher performance. In the study of Aslan and Akın [32], feature extraction from EEG signals was performed using the relative wavelet energy (RWE) method. They trained these features using the KNN (K-Nearest Neighbors) algorithm and reported that they obtained approximately 90% accuracy with this method. Thilakvathi et al. [33] tried to distinguish SZ patients using the Support Vector Machines (SVM) algorithm. They used different methods (such as Higuchi's Fractal Dimension, Spectral Entropy, Hannon Entropy, Information Entropy, and Kolmogorov Complexity values) to obtain the features. By classifying these features with SVM, they obtained 88.5% accuracy. Devia et al. [5] compared ERP (event-related potentials) between healthy and sick individuals. Then they examined the classifier design and performance. In their work, they presented LDA classifiers, then rule-based classifiers and finally, combined approach methods. The best classifier has an overall accuracy of 71%, a sensitivity of 81% for detecting patients and a specificity of 59% for detecting controls. Piryatinska et al. [34] performed the classification of 16-channel EEG recordings of 84 adolescent individuals (39 healthy-45 SZ). Random forest classifier and SVM classifiers were used. Using the obtained features with RF algorithm, the highest accuracy performance was achieved as 83.6%. Shim et al. [24] performed the feature extraction process on the EEG channels of 34 healthy and 34 SZ patients at the sensor level and at the source level, and classified the characteristics they obtained with SVM and identified SZ patients with an accuracy of 88.24%. Sui et al. [35] performed the feature extraction procedure on the EEG records of 48 SZ patients and 53 healthy individuals using the Multiple set canonical correlation analysis method, and classified the obtained features with SVM with an accuracy of 74% and distinguished SZ patients from healthy individuals. Boostani et al. [36] applied autoregressive (AR) model parameters, band power and fractal dimension methods to EEG recordings and classified the feature values they obtained with LDA and achieved 87.5% success. Siuly et al. [37] obtained empirical mode decomposition (EMD) based features from EEG records of 49 SZ patients and 32 healthy individuals, and were classified with Decision Tree (DT), k-NN, SVM and ensemble bagged tree (EBT) classifiers. EBT achieved 89.59% better performance compared to other reported classifiers.

In the literature studies of the related field mentioned above, the features obtained by extracting features from

the EEG records were classified with traditional machine learning algorithms. Although this feature extraction technique has advantages such as successful estimation, it is an important disadvantage that it requires experts with extensive knowledge in this field. With new developments, DL is being investigated as a new alternative to traditional machine learning algorithms. Due to the fact that DL does not require any preprocessing and feature extraction, this area has been intensely studied recently. Convolutional neural network (CNN) is a kind of DL network and the features of an input data are obtained automatically from the relevant layers of the CNN. There are only a few literature studies that classify with CNN using EEG records of SZ patients.

In a study, Phang et al. [38] used brain functional connectivity information as a feature in their method and tested these features with several methods. They obtained the features using methods such as partial directed coherence, vector autoregressive model, and complex network measurements of the network topology. Then, using these features, they fed them into two CNN models. Finally, there is the Fully Connected Neural Network (FCN) that can classify healthy controls and SZ patients. As a result of the classification, they achieved 93.06% performance. They report that their method has reached satisfactory accuracy. However, their methods require additional data such as brain connection features. Oh et al. [3] classified 19-channel EEG recordings of 28 individuals (14 healthy—14 SZ) using a CNN model. The CNN model they use in their work consists of 11 layers in total. They did not preprocess, and fed the CNN model directly with data from raw EEG channels. It is reported that their methods achieved a performance of 98.07% for non-subject based tests and 81.26% for subject based tests. Both methods using DL to detect SZ lack interpretable results. In another study, Aslan and Akın [39] converted 5-s segments they obtained from EEG channels into spectrogram images. CNN, which they trained with the features they obtained, achieved very successful results with 95% and 97% success for two different datasets.

## Methodology

### Dataset A

The first dataset used in this study was, by the Mental Health Research Center (MHRC), the eldest of SZ patients was 14, the youngest was 10 and 8 months old while the eldest healthy controls were 13 years old, 9 months old, and the youngest was 11 years old. It consists of 60-s EEG recordings of 45 children with SZ and 39 healthy children. All SZ patients in the dataset were reviewed by Mental Health Research Center (MHRC) experts, and none of the controls in the dataset received chemical treatment. The mean age of

the SZ patient and healthy controls is 12 years and 3 months [40]. EEG data were taken from 16 channels and sampling frequency was recorded as 128 Hz. During the recording, the subjects was awake and closed eyes. The EEG electrode array is positioned as O1, O2, P3, P4, Pz, T5, T6, C3, C4, Cz, T3, T4, F3, F4, F7, and F8 in accordance with the international standard 10–20.

## Dataset B

The other dataset used in this study consists of 19-channel EEG recordings of a total of 14 SZ patients and 14 healthy controls by the Institute of Psychiatry and Neurology in Warsaw, Poland. EEG recordings are between 12 and 15 min belonging to 14 men and 14 women with a mean age of  $27.3 \pm 3.3$  and  $28.3 \pm 4.1$ . EEG data were recorded with a sampling frequency of 250 Hz, with subject's eyes closed. The recordings were taken from 19 electrodes and arranged as Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2. [41].

## Deep learning

Deep learning (DL) is a new machine learning method that enables the features of a data to be learned hierarchically and includes deeper neural networks. The most important feature that makes DL networks superior to traditional machine learning algorithms is that the extraction process does not need to be done in advance. While feature values need to be extracted in advance in traditional machine learning, important features in DL networks are learned automatically.

DL algorithms require large-scale data and costly hardware for processing data. As such needs become more accessible, the use of DL instead of traditional machine learning algorithms is increasingly common. As access to the needs required for DL has recently increased, it has become widely used as an alternative method for processing, analysis and evaluation of medical images, often with deep learning networks such as CNN [42]. In this study, scalogram images are classified using VGG16 architecture, which has CNN deep networks structure.

## Convolutional neural networks

CNN are a kind of deep learning network consisting of network layers from which features can be extracted hierarchically. It is widely used in multimedia data types such as images [43]. A CNN model is DL networks designed in a structure where multimedia data types (for example images) can be extracted automatically in a hierarchical fashion through network layers [43]. The structure of a CNN model generally consists of two important parts. The first section contains the convolution and pooling layers for

feature extraction. The second part includes the Fully Connected Network layers (FCN) that work like the Traditional Multi-Layer Perceptron (MLP) for the classification stage. A common CNN model consists of several successive convolution layers and pooling layers, each of which is done from the output of the previous layer.

Simple features (such as simple lines on the image) are extracted from the first layers of the network. In the next layers, more complex features (such as an object in an image) are extracted by using these simple features. This kind of hierarchical learning of traits has been carried out, inspired by the human cortex, where cells respond in a hierarchical fashion to visual elements [44].

There are many filters in a convolutional layer. In the convolutional layer, the input image, each used to compute a feature map  $X_k$  of the image, is returned using these filters,  $W = W_1, W_2, W_k$ . Therefore, the number of feature maps and filters in the convolution layer are the same. Each feature map is calculated using the formula in Eq. 1. (1) in the formula  $\sigma(\cdot)$  is a nonlinear transfer function and  $b$  bias [45]:

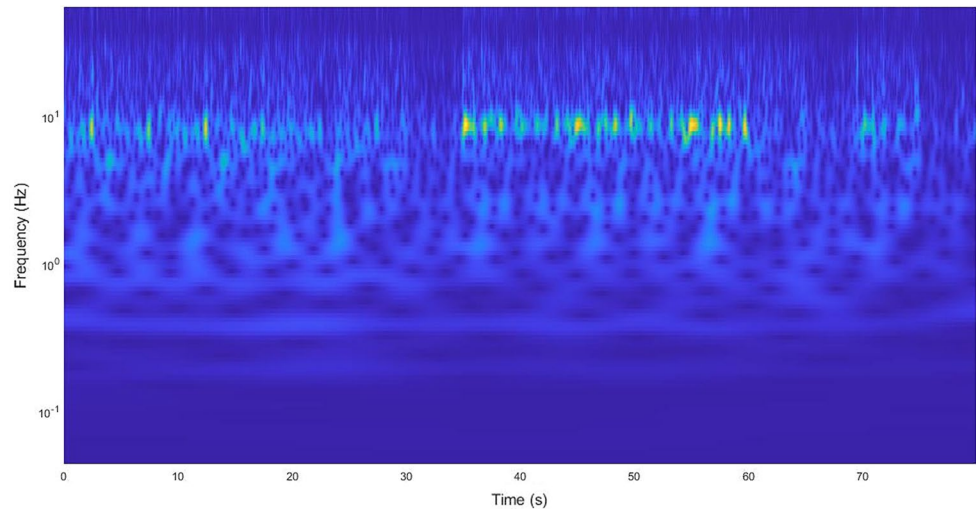
$$X_k^l = \sigma(W_k^{l-1} * X^{l-1} + b_k^{l-1}) \quad (1)$$

The selected pooling function of feature maps is maximum, minimum or average. The Select function is applied to each pixel group in the feature map, and the result of the function (for example, the maximum value in that group) is selected to describe the group in the new subsampled feature map. Generally, a normal CNN deep learning network consists of convolutional layers, pool layers and an FCN [46].

## VGG16 CNN architecture

The VGG16 CNN architecture is a CNN deep network model consisting of 16 layers and designed by the Oxford University Visual Geometry Group for the ILSVRC-2014 competition. Having a deeper learning architecture is the biggest difference from previous deep learning architectures. The VGG16 network is fed with images up to  $224 \times 224 \times 3$  (RGB). Each image is passed through 5 convolutional layers with  $3 \times 3$  filter size. Each block terminates with the maximum pooling layer where input values are down sampled 2 times [47]. The property values obtained in the last step are transferred to an FCN for classification process [34]. The VGG16 deep learning architecture used as a classifier in this study is a state-of-the-art deep learning architecture. It has been observed that this architecture exhibits high classification performance in the classification of many different images. Considering that a good classification performance was obtained as a result of the experimental studies carried out in this aspect, VGG16 deep network architecture was preferred in our study. In this study, the scalogram images were classified using the VGG16 deep learning architecture.

**Fig. 1** A sample scalogram image created using Morlet wavelet with the help of Matlab software



### Creating scalogram images by using CWT method

Continuous Wavelet Transform (CWT) is a method for a time frequency analysis that transforms the signal into 2D by allowing scales to change continuously, with the help of a wavelet function. CWT offers very good frequency and time localization to create a 2D image of the time frequency information of a signal. CWT is a very useful method for mapping the features of the variation of variable signals. CWT is an important time frequency analysis method used to determine whether a signal is constant or variable. If a signal is variable, it can be used to determine fixed parts of the signal [48].

In this study, in order to access the frequency information of different time points of our EEG signal data, the EEG signal is divided into 5-s segments and the vector obtained by combining each segment from all channels has been transformed into a scalogram with the CWT method. These vectors were created using Matlab software. Morlet wavelet was used as the main wavelet while creating scalogram images. Figure 1 shows a sample scalogram image taken from the images we have obtained using the morlet wavelet with the CWT method.

### General architecture of the proposed CAD method for automatic schizophrenia detection

The CAD method we propose does not require any steps such as preprocessing and feature extraction. Figure 2 contains the flowchart of the method we used in our study. Each input data to be trained by the CNN network includes segmentation and then the scalogram stages. Input data trained by the CNN network are classified and finalized by being included in a class as a patient or healthy control.

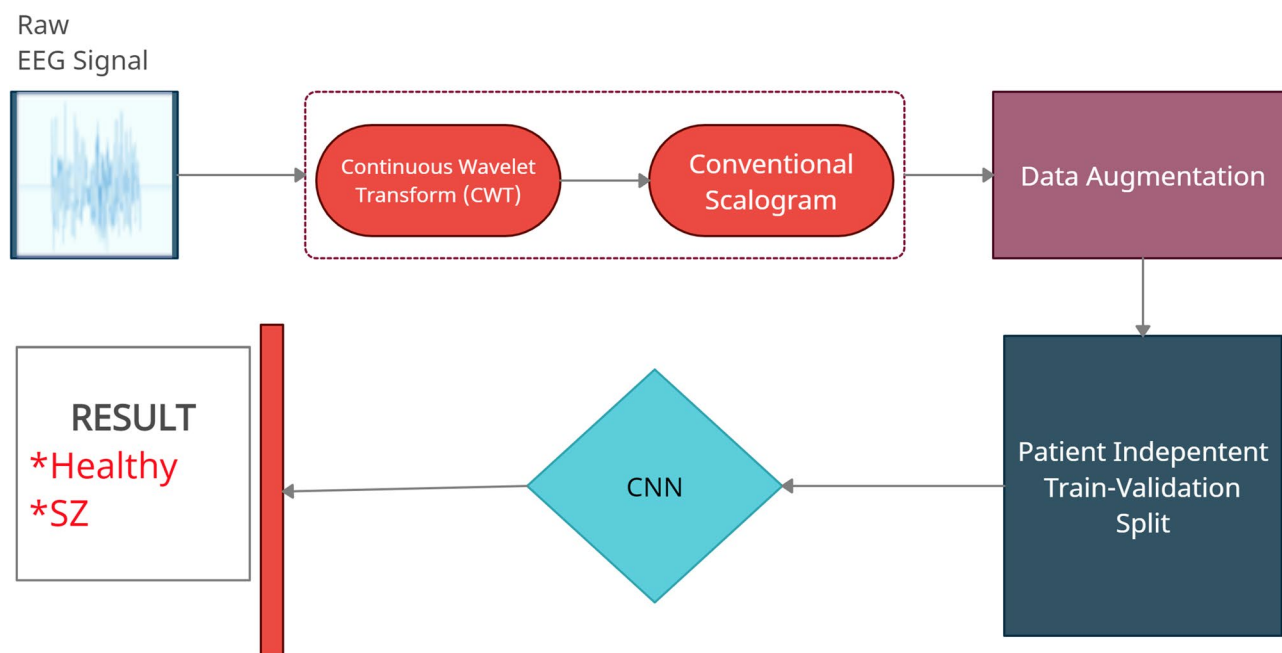
### Results

The success of the method was tested by applying the proposed method to two datasets that include different age groups, including SZ patients and healthy controls.

In the first dataset (A), it consists of 60-s EEG records of 45 children with SZ disease and 39 healthy children. EEG recordings include recordings from 16 electrodes. Each channel is divided into segments of 5 s length, by combining the segments corresponding to each channel, a single-dimensional vector, each of which is 10240 long, has been obtained. The sampling frequency of EEG records is 128 Hz. The length of a vector is calculated as (5 s  $\times$  128 sampling frequency  $\times$  16 channels). The vectors obtained are transformed into scalogram images (224  $\times$  224) using CWT method and morlet wavelet. In this dataset, 1008 scalogram images were obtained as a result of the segmentation process mentioned above. In Fig. 3, sample scalogram images taken from different segments belonging to different people obtained from dataset A are shown.

The scalogram images obtained were divided into 80% training and 20% test datasets. By using the VGG16 deep learning network architecture, the scalogram images obtained to distinguish between SZ patients and healthy controls are given to the network, and the classification process is performed. For the VGG16 deep learning network, hyper parameters were determined such that the image size given as input data was 112  $\times$  112, learning ratio 1.0e  $-$  4, batch size 128 and optimizer Adam [49]. In the experiments conducted in this study, it was determined that the value of 100 epochs is sufficient to train the network. At the end of the training of the network, an average accuracy of 98% was obtained at the 78th epoch.

The second dataset consists of EEG records of 14 healthy controls and 28 adults with 14 SZ patients. EEG recordings consist of recordings taken from 19 electrodes. The EEG

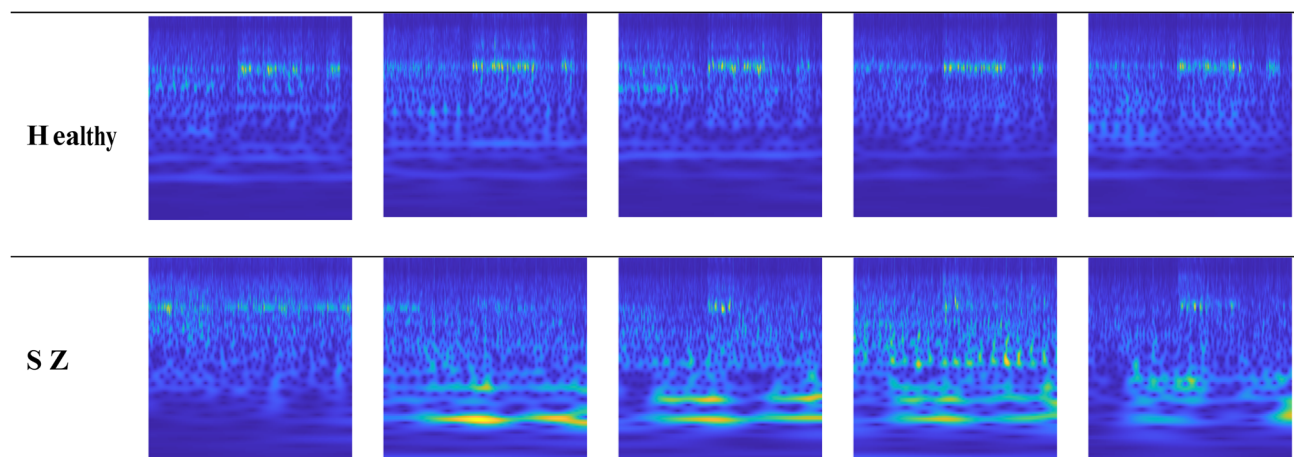


**Fig. 2** Flow chart of the proposed method

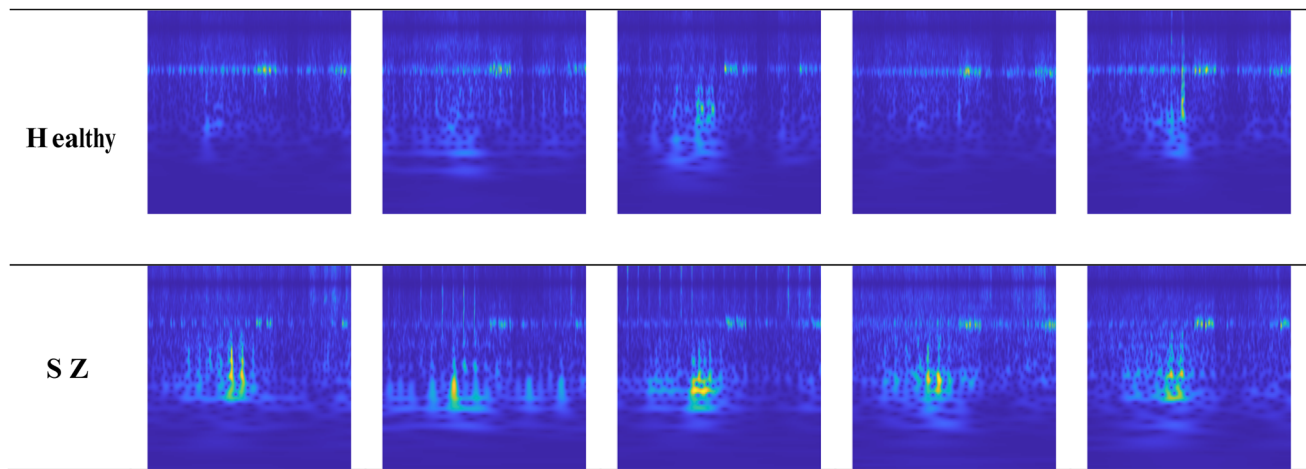
data in dataset B consists of recordings made for periods ranging from 12 to 15 min. Therefore, while performing the segmentation process, the length of the shortest record was taken into consideration. EEG recordings taken from each channel are divided into 5-s segments. Considering the different lengths of EEG recordings as a result of this segmentation, 148 segments were obtained for each SZ patient and 173 segments for each healthy control. In Fig. 4, sample scalogram images taken from different segments belonging to different people obtained from Dataset B are shown. The hyper parameters used to train the VGG16 deep learning network in the data set A were also used in the dataset B. By applying the same method used in dataset A to dataset

B, a total of 4494 scalogram images were obtained. After training the network, an accuracy of 99.5% was achieved at the 25th epoch.

Accuracy values obtained with both of these two datasets achieved higher performance in the literature than studies conducted with these datasets. Figure 5 shows the change of accuracy and error value for dataset A and B according to the epoch training time. Adequate training of the network is an important factor that seriously affects the classification performance of the model. In literature studies, there are no specific values required for how much the network should be trained. Obtaining the best performance values can be determined based on experimental studies.



**Fig. 3** Sample scalogram images taken from different segments belonging to different people from dataset A



**Fig. 4** Sample scalogram images taken from different segments of different people from dataset B

The fact that both the Recall and Sensitivity measurements of the classifier are simultaneously high means that it distinguishes the diseased individuals as well as the healthy controls as disease-free. Performing the classification process in this way shows that it is an ideal classifier. High Sensitivity and Recall measurements are expected in an ideal classification process, and as a result of these two measurement values, the F1 score results in a high value. Evaluation metrics obtained by the proposed method are shown in Table 1 for dataset A and in Table 2 for dataset B.

It is shown in the confusion matrices that the accuracy rate of the classification is higher than 0.97 for the sick and healthy controls and the incorrect evaluation rate of the classification is lower than 0.03 for both datasets (See Fig. 6). In the ROC curves shown in Fig. 7, the AUC value is quite high as 0.98 for dataset A and 0.995 for dataset B, and as shown in the graph, a very successful classification has been performed.

The data in dataset A were studied by Phang et al. [38]. They stated that they achieved 93.06% classification performance in their study. In the method they used, they used a CNN deep learning network where raw EEG signals were given directly. On the other hand, Oh et al. tested the performance of their method with a CNN model fed directly with EEG signals without preprocessing using dataset B [3]. As a result of the method they suggested, they achieved 98.07% success in their work. Aslan and Akin [39], on the other hand, trained the spectrogram images they obtained from both data sets with CNN deep learning network and performed the classification process with 95% accuracy for dataset A and 97% accuracy for dataset B.

In our study, when compared with the literature studies in this field, quite high performance and accuracy values were obtained. In order to prove the accuracy of the method we use, it was evaluated by working with two different datasets

consisting of different age groups. The method we used in our study obtained very high accuracy values in both datasets, proving the reliability of the method. According to other literature studies, one of the most important elements of our study is that it creates images that can be interpreted more easily without the need for any expert opinion.

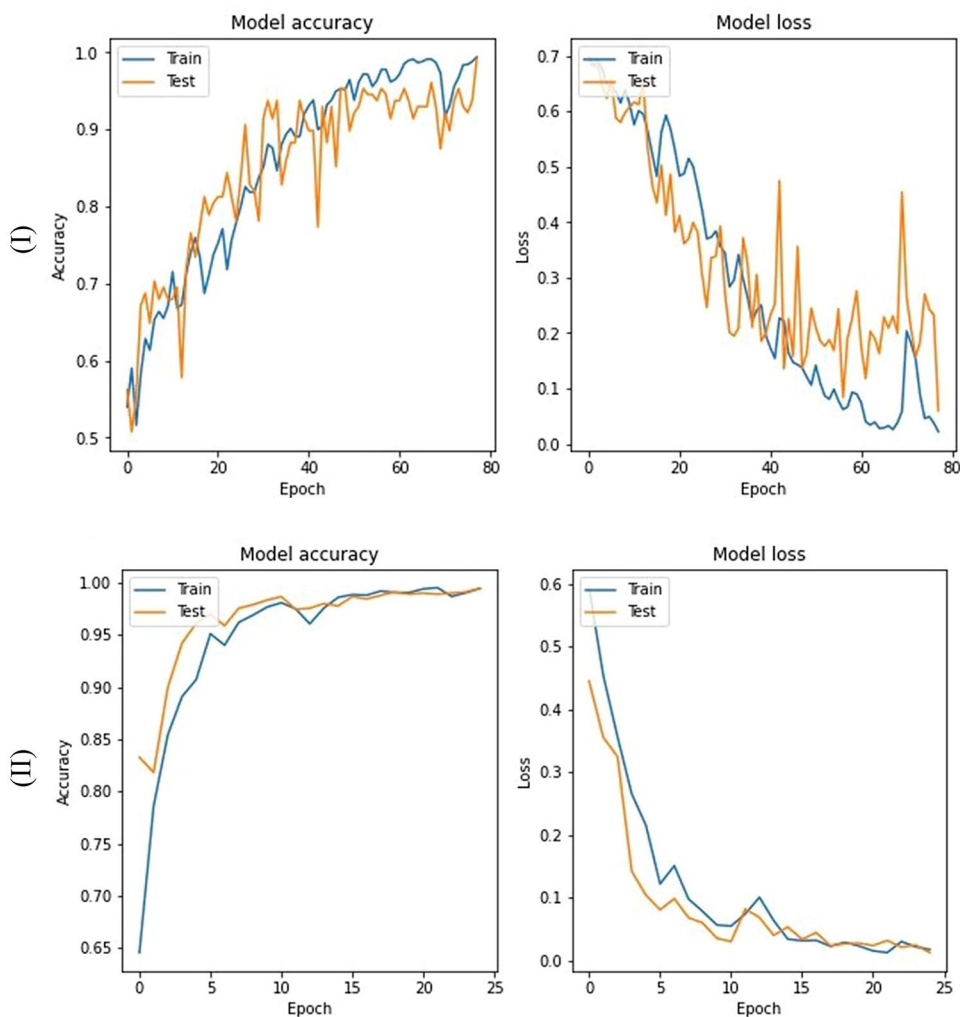
## Discussion

While there are only temporal relationships in the raw EEG signal, the scalogram images we obtain contain more feature information. Scalogram images make a better classification with the CNN network than raw EEG signals, as CNN deep learning networks use spatial features to classify an object that has spatial relationships between pixels. This is because the CNN deep learning network enables the network to learn with higher performance by extracting much more features from the scalogram images.

CWT performs time frequency analysis by using different scales and shifting along the signal during transformation with a master wavelet. This provides long time lapse windowing at low frequencies and short time lapse windowing at high frequencies. By using different sized windows with CWT, high and low frequency information in time series are analyzed in the best way. Scalogram, on the other hand, is a 2D representation of a CWT result signal. Hence, Scalogram images contain more information about the high and low frequency characteristics of the signal. The scalogram is widely used in situations where better frequency localization for low frequency, long duration events and better time localization for short duration, high frequency events are desired.

It is not directly understood from a scalogram image whether the image belongs to the patient or a healthy individual. When the images of scalogram of the healthy and SZ

**Fig. 5** The change of accuracy and error value versus epoch values during network training: (I) For dataset A (II) For dataset B



patients shown in Figs. 3 and 4 are examined, it can be seen that there are changes at different times and frequencies, but it is not clearly interpretable. It is not possible to generalize and formulate for all data in a dataset. We try to demonstrate a clearer interpretation of the differences between the SZ patient and healthy controls using a method called Activation Maximization (AC) [50] of images of the CNN deep learning network output.

Since it creates a synthetic image that is synthesized by finding these values recursively by maximizing the output of the network belonging to a class, the AC image can be considered as an input describing a class. Although there are obvious differences in the AC image of the SZ patient and the healthy control in Fig. 8, the AC images show that the network responds very differently to different scalogram images. It does not indicate the actual scalogram images, as the features specified in the AC images are artificially generated to maximize filter activation.

The literature studies in Table 3 use traditional machine learning algorithms to detect schizophrenia from EEG

**Table 1** Accuracy evaluation criteria obtained with dataset A, using the proposed method

Dataset A	Precision	Recall	F1-score	Support
Healthy	0.98	0.98	0.98	94
SZ	0.98	0.98	0.98	108
Accuracy			0.98	202
Macro average	0.98	0.98	0.98	202
Weighted average	0.98	0.98	0.98	202

signals. Table 3 shows by which feature extraction methods the features were obtained for each study and the classification accuracies obtained by classifying these features with which classifier.

Due to the aforementioned disadvantage of AC images, we used Grad-CAM (Gradient Weighted Activation Maps) [52] and Saliency Map [53] techniques to more clearly reveal the differences of the SZ patient. Grad-CAM is a technique used to obtain images in which the important and relevant



**Table 2** Accuracy evaluation criteria obtained with dataset B, using the proposed method

Dataset B	Precision	Recall	F1-score	Support
Healthy	0.99	1.00	0.99	484
SZ	1.00	0.99	0.99	414
Accuracy			0.99	898
Macro average	0.99	0.99	0.99	898
Weighted average	0.99	0.99	0.99	898

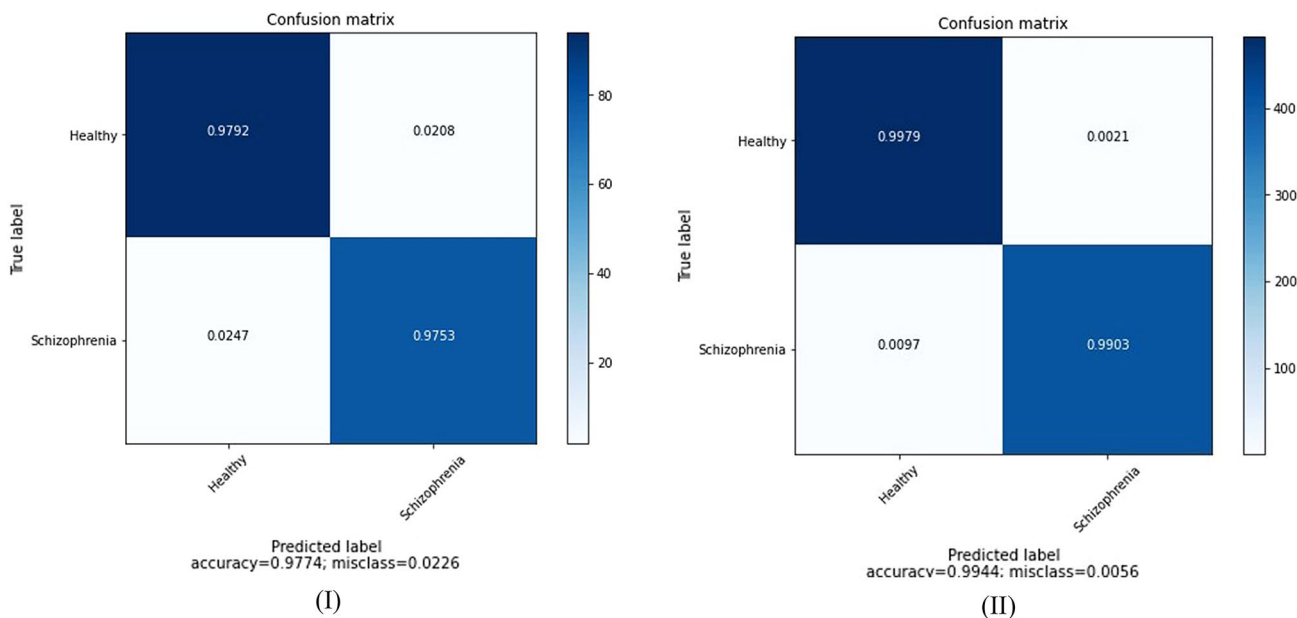
features of a particular class on the image are shown as a color map. The important spatial features of the estimated class are displayed in colors close to red in the grad-CAM image. In Fig. 9, Grad-CAM images obtained from scalogram images belonging to different segments taken from different individuals are shown. When the Grad-CAM images of individuals with SZ are examined, a distinct high density extending from the upper part to the middle is seen, while images of healthy controllers never show such intensity colors.

Another technique we use to make scalogram images interpretable is Saliency Map technique. Saliency Map is an image where the brightness of a pixel represents how prominent the pixel is, that is, the brightness of a pixel is directly proportional to its saliency. Saliency maps are also referred to as a heat map in which the temperature refers to the regions of the image that have a major influence on

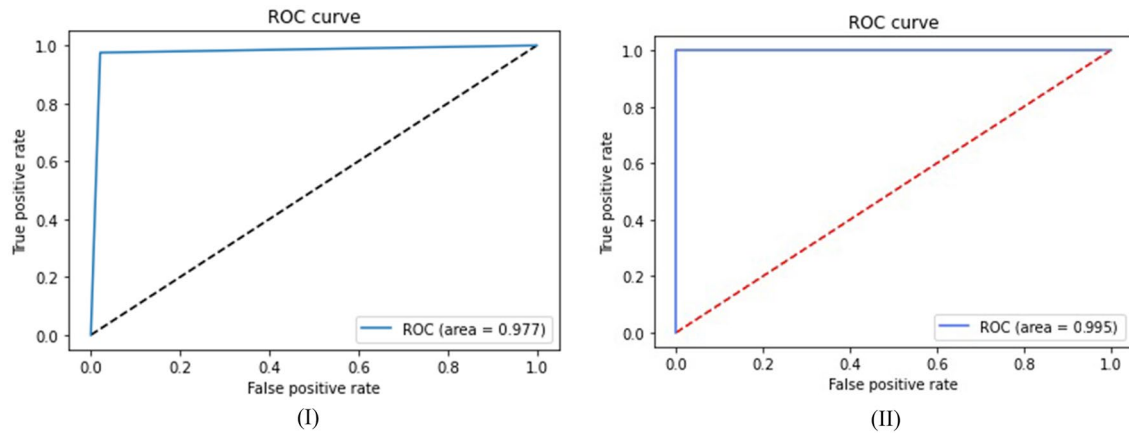
predicting the class to which the object belongs. The purpose of the Saliency Map is to find the areas that stand out in every location in the visual field and to guide the selection of the destinations according to the spatial distribution of the projection [54]. Saliency Map images of healthy and SZ patients taken from different individuals and shown in Fig. 10 are shown. As seen in Fig. 10, Saliency Maps of healthy and SZ patients can be distinguished by obvious differences. When Saliency Maps of SZ patients are examined, it is observed that they create a more intense heat map.

It should be noted that the Grad-CAM and Saliency Map methods shown in Fig. 9 and Fig. 10 are common to others not shown in the figure. As can be clearly seen with these two methods, the frequency components on the scalogram images contain important spatial features in distinguishing SZ patients and healthy controls. In Tables 3 and 4, the literature studies conducted in the relevant field are summarized in terms of the method used and the accuracy values achieved. When these studies in Tables 3 and 4 are examined, it is seen that the proposed method achieves higher accuracy and performance than all studies and has a performance comparable to the proposed method.

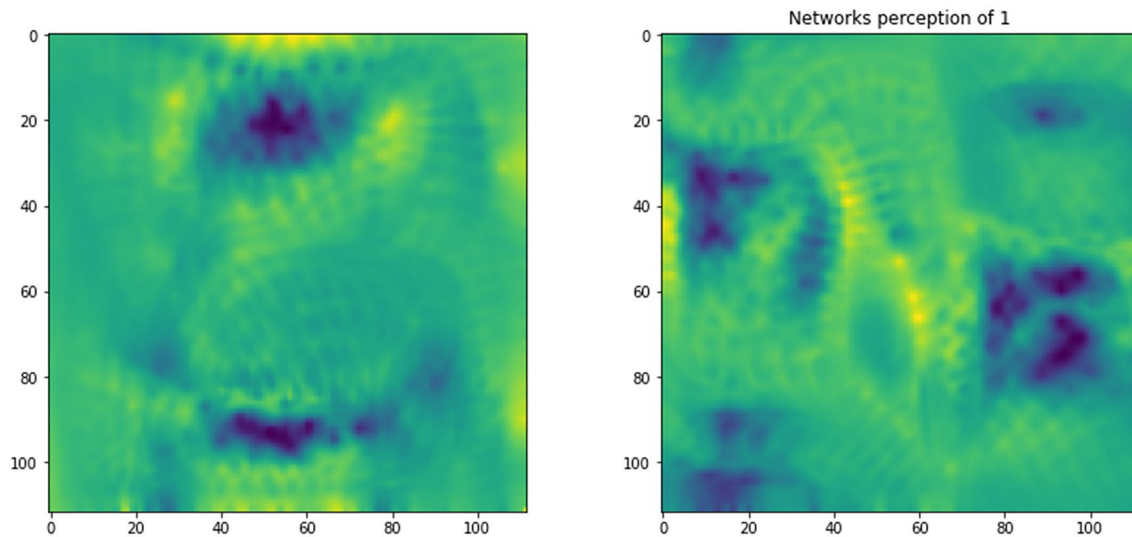
This study is superior to the studies mentioned in Tables 3 and 4 with its simple operation process and producing interpretable outputs help for expert opinion. As far as we know, the highest accuracy value achieved so far has been achieved in SZ determination. Our proposed study has important advantages over previous literature methods beyond its high



**Fig. 6** Confusion matrix: (I) for dataset A (II) for dataset B



**Fig. 7** ROC curves: (I) for dataset A (II) for dataset B



**Fig. 8** Activation maximization images of the healthy control (left) and the SZ patient (right) obtained from the CNN network

performance. First of all, it does not require feature extraction and preprocessing compared to traditional machine learning algorithms. It also includes a very simple but high-performance process. Another advantage is that it offers 3 different techniques to create clearly interpretable images of SZ disease. It makes it possible to distinguish the SZ patient and healthy control very easily from AC, Grad-CAM and Saliency Map images help for expert opinion.

Recently, with the new developments in deep learning networks, there has been an increasing interest in the analysis of biomedical images and signals. There are

some limitations in our study that we think can be compensated in future studies. First of all, it is necessary to select the appropriate values of these parameters for the optimization of the hyper parameters used in deep learning and for the best classification performance. In the next stage, a CNN deep learning network architecture with less computational complexity and high classification performance should be designed instead of a state-of-the-art, pre-trained, multi-layered and high computational deep learning network.

**Table 3** Relevant literature studies detecting schizophrenia using traditional machine learning algorithms

Authors	Classifiers	EEG Dataset	Features extraction	Overall Accuracy
Johannesen et al. [51]	SVM	40 SZ- 12 healthy	Alpha, beta, theta1, theta2, gamma frequency components	87.0%
Devia et al. [5]	LDA	11 SZ- 9 healthy	ERPs (Event-related potentials) features	71.0%
Kim et al. [26]	ROC analysis	90 SZ- 90 healthy	Spectral analysis using frequency bands	62.20%
Piryatinska et al. [34]	RF	45 SZ- 39 healthy	e-complexity coefficients	83.60%
Shim et al. [24]	SVM	34 SZ- 34 healthy	Source- level features And sensor-level features	88.24%
Sui et al. [35]	SVM	48 SZ- 53 healthy	Multi-set canonical correlation analysis	74.0%
Boostani et al. [36]	LDA	13 SZ- 18 healthy	Band power, fractal dimension and autoregressive (AR) model parameters	87.50%
Thilakvathi et al. [33]	SVM	55 SZ- 23 healthy	Dimension, Approximate Entropies, Shannon entropy, Information entropy, Kolmogorov complexity, Spectral entropy, Higuchi's Fractal	88.50%
Siuly et al. [37]	EBT	49 SZ- 32 healthy	EMD (Empirical Mode Decomposition)	89.59%
Aslan and Akın [32]	KNN	45 SZ- 39 healthy	RWT (Relative Wavelet Energy)	90%
Santos-Mayo et al. [30]	SVM MLP	16 SZ- 31 healthy	J5 feature extraction and EEGLAB feature extraction,	MLP: 93.42% SVM: 92.23%
Dvey-Aharon et al. [28]	Time- Frequency transformation followed by Feature Optimization (TFFO)	50 SZ- 25 healthy	Stockwell transformation	Between 92% and 93.9%

## Conclusion

This study proposes a CAD method to automatically distinguish between an SZ patient and a healthy control. The fact that the proposed method has been evaluated with two different datasets containing different age groups clearly reveals the classification success of the method. With the proposed method, the highest accuracy values obtained in this field in the literature were obtained with the VGG16 deep learning algorithm, the highest classification accuracy of 99.5% for dataset B and 98% for dataset A. In addition to its high classification performance, the proposed method has been shown to be successful in terms of computational efficiency by calculating at low epoch values.

This study shows that the analysis of frequency components in EEG data is a robust method for detecting SZ disease, a type of brain deformation. In the proposed method,

images in which SZ patients and healthy controls can be clearly distinguished were created by using 3 different CNN network visualization methods (AC, Grad-CAM, and Saliency Map). The reason why the images obtained by these techniques show obvious differences is that the input scalogram images have important spatial features at the time frequency level. It is thought that the images obtained by CNN network visualization methods can support expert opinion.

The method proposed in this study can serve as an example for CAD studies that can also be used to detect diseases other than EEG recordings. The use of new CNN models is thought to significantly benefit classification performance. In addition, for network training, models with a simpler structure consisting of fewer layers and requiring less computation time can be preferred instead of complex models.

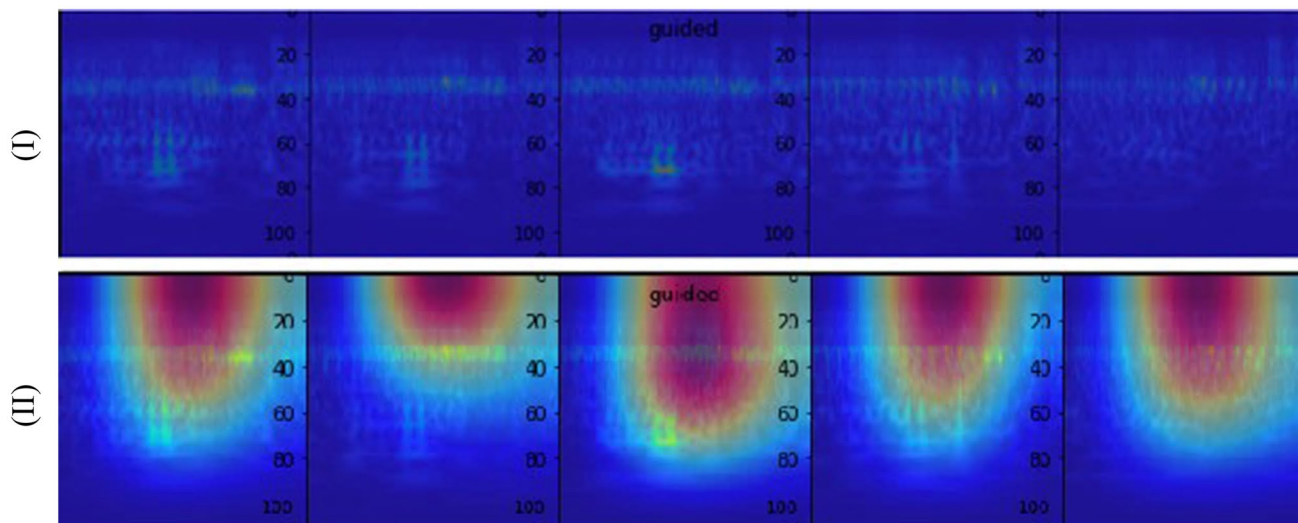


Fig. 9 Grad-CAM images taken from different individuals and different segments: (I) taken from healthy controls (II) SZ patients

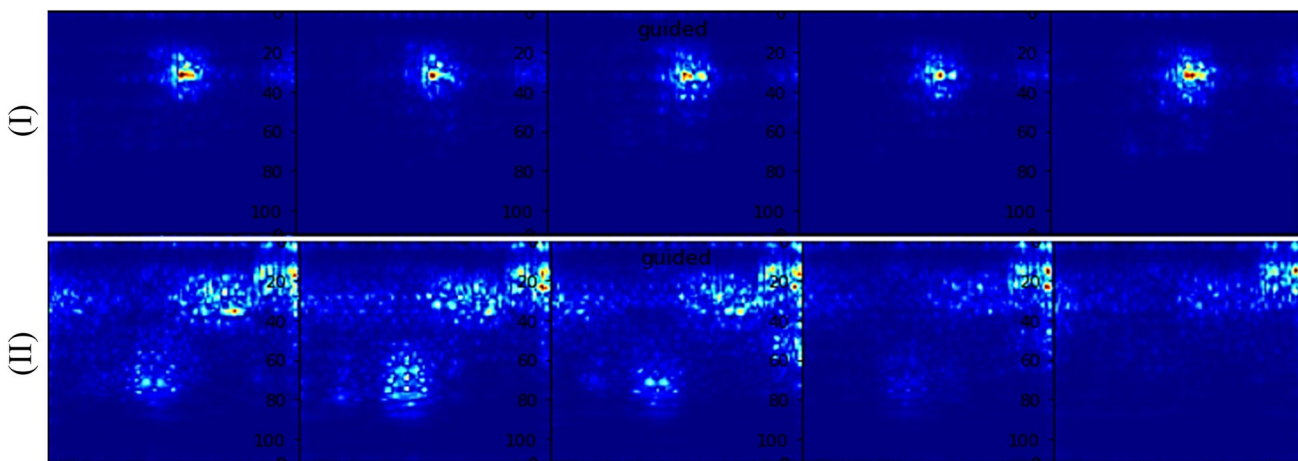


Fig. 10 Saliency Map images taken from different individuals and different segments: (I) taken from healthy controls (II) SZ patients

**Table 4** Relevant literature studies detecting schizophrenia using deep learning methods

Authors	Deep learning method	Overall accuracy (%)	EEG Dataset
Phang et al. [38]	CNN	93.06	45 SZ–39 healthy
Shu Lih Oh et al. [3]	CNN	98.07	14 SZ–14 healthy
Aslan et al. [39]	CNN	95	45 SZ–39 healthy
		97	14 SZ–14 healthy
<b>This study</b>	<b>CNN</b>	<b>98</b>	<b>45 SZ–39 healthy</b>
		<b>99.5</b>	<b>14 SZ–14 healthy</b>

Bold font indicates the prominent aspects of the proposed method when compared to other methods

**Data availability** The datasets used in this study are publicly accessible data including EEG data of healthy and SZ patient individuals. The datasets used in this study can be accessed from: (dataset A) [http://brain.bio.msu.ru/eeg\\_schizophrenia.htm](http://brain.bio.msu.ru/eeg_schizophrenia.htm); (dataset B) <https://repositorio.icm.edu.pl/dataset.xhtml?persistentId=doi:10.18150/repo.0107441> web addresses.

**Declarations**

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

## References

- WHO. "Schizophrenia\_." [https://www.who.int/mental\\_health/management/schizophrenia/en/](https://www.who.int/mental_health/management/schizophrenia/en/). Accessed on 24 Sept 2020
- James SL et al (2018) Global, regional, and national incidence, prevalence, and years lived with disability for 354 Diseases and Injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. [https://doi.org/10.1016/S0140-6736\(18\)32279-7](https://doi.org/10.1016/S0140-6736(18)32279-7)
- Oh SL, Vicnesh J, Ciaccio EJ, Yuvaraj R, Acharya UR (2019) Deep convolutional neural network model for automated diagnosis of schizophrenia using EEG signals. *Appl Sci* 9(14):2870. <https://doi.org/10.3390/app9142870>
- Laursen TM, Nordentoft M, Mortensen PB (2014) Excess early mortality in schizophrenia. *Annu Rev Clin Psychol* 10:425
- Devia C et al (2019) Eeg classification during scene free-viewing for schizophrenia detection. *IEEE Trans Neural Syst Rehabil Eng* 27(6):1193–1199. <https://doi.org/10.1109/TNSRE.2019.2913799>
- Siuly S, Alcin OF, Bajaj V, Sengur A, Zhang Y (2019) Exploring Hermite transformation in brain signal analysis for the detection of epileptic seizure. *IET Sci Meas Technol*. <https://doi.org/10.1049/iet-smt.2018.5358>
- Talo M, Baloglu UB, Yıldırım Ö, Rajendra Acharya U (2019) Application of deep transfer learning for automated brain abnormality classification using MR images. *Cogn Syst Res*. <https://doi.org/10.1016/j.cogsys.2018.12.007>
- Gudigar A, Raghavendra U, San TR, Ciaccio EJ, Acharya UR (2019) Application of multiresolution analysis for automated detection of brain abnormality using MR images: a comparative study. *Future Gener Comput Syst*. <https://doi.org/10.1016/j.future.2018.08.008>
- Acharya UR, Sree SV, Ang PCA, Yanti R, Suri JS (2012) Application of non-linear and wavelet based features for the automated identification of epileptic EEG signals. *Int J Neural Syst* 22(02):1250002. <https://doi.org/10.1142/s0129065712500025>
- Supriya S, Siuly S, Wang H, Zhang Y (2018) EEG sleep stages analysis and classification based on weighed complex network features. *IEEE Trans Emerg Top Comput Intell*. <https://doi.org/10.1109/TETCI.2018.2876529>
- Yin J, Cao J, Siuly S, Wang H (2019) An integrated MCI detection framework based on spectral-temporal analysis. *Int J Autom Comput*. <https://doi.org/10.1007/s11633-019-1197-4>
- Tuncer T, Dogan S, Naik GR, Pławiak P (2021) Epilepsy attacks recognition based on 1D octal pattern, wavelet transform and EEG signals. *Multimed Tools Appl*. <https://doi.org/10.1007/s11042-021-10882-4>
- Gao Z et al (2019) EEG-based spatio-temporal convolutional neural network for driver fatigue evaluation. *IEEE Trans neural networks Learn Syst* 30(9):2755–2763. <https://doi.org/10.1109/TNNLS.2018.2886414>
- Chai R et al (2016) Driver fatigue classification with independent component by entropy rate bound minimization analysis in an EEG-based system. *IEEE J Biomed Heal Inform* 21(3):715–724. <https://doi.org/10.1109/JBHI.2016.2532354>
- Liu G et al (2017) Complexity analysis of electroencephalogram dynamics in patients with Parkinson's disease. *Park Dis*. <https://doi.org/10.1155/2017/8701061>
- Al-Ani A, Koprinska I, Naik G (2017) Dynamically identifying relevant EEG channels by utilizing channels classification behaviour. *Expert Syst Appl* 83:273–282. <https://doi.org/10.1016/j.eswa.2017.04.042>
- Acharyya A, Jadhav PN, Bono V, Maharatna K, Naik GR (2018) Low-complexity hardware design methodology for reliable and automated removal of ocular and muscular artifact from EEG. *Comput Methods Programs Biomed* 158:123–133. <https://doi.org/10.1016/j.cmpb.2018.02.009>
- Butkevičiute E et al (2019) Removal of movement artefact for mobile EEG analysis in sports exercises. *IEEE Access* 7:7206–7217. <https://doi.org/10.1109/ACCESS.2018.2890335>
- Nejedly P et al (2019) Intracerebral EEG artifact identification using convolutional neural networks. *Neuroinformatics* 17(2):225–234. <https://doi.org/10.1007/s12021-018-9397-6>
- Bhardwaj S, Jadhav P, Adapa B, Acharyya A, Naik GR (2015) Online and automated reliable system design to remove blink and muscle artefact in EEG. In: The 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, Milan, pp 6784–6787
- Jadhav PN, Shanmugan D, Chourasia A, Ghole AR, Acharyya AA, Naik G (2014) Automated detection and correction of eye blink and muscular artefacts in EEG signal for analysis of Autism Spectrum Disorder. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, Chicago, pp 1881–1884. <https://doi.org/10.1109/embc.2014.6943977>
- Kilicarslan A, Vidal JLC (2019) Characterization and real-time removal of motion artifacts from EEG signals. *J Neural Eng* 16(5):56027. <https://doi.org/10.1088/1741-2552/ab2b61>
- Zhang L (2019) EEG signals classification using machine learning for the identification and diagnosis of schizophrenia. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, Berlin, pp 4521–4524. <https://doi.org/10.1109/EMBC.2019.8857946>
- Shim M, Hwang H-J, Kim D-W, Lee S-H, Im C-H (2016) Machine-learning-based diagnosis of schizophrenia using combined sensor-level and source-level EEG features. *Schizophr Res* 176(2–3):314–319. <https://doi.org/10.1016/j.schres.2016.05.007>
- Cao B et al (2018) Treatment response prediction and individualized identification of first-episode drug-naive schizophrenia using brain functional connectivity. *Mol Psychiatry*. <https://doi.org/10.1038/s41380-018-0106-5>
- Kim JW, Lee YS, Han DH, Min KJ, Lee J, Lee K (2015) Diagnostic utility of quantitative EEG in un-medicated schizophrenia. *Neurosci Lett* 589:126–131. <https://doi.org/10.1016/j.neulet.2014.12.064>
- Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 134(1):9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Dvey-Aharon Z, Fogelson N, Peled A, Intrator N (2015) Schizophrenia detection and classification by advanced analysis of EEG recordings using a single electrode approach. *PLoS ONE* 10(4):e0123033. <https://doi.org/10.1371/journal.pone.0123033>
- Stockwell RG, Mansinha L, Lowe RP (1996) Localization of the complex spectrum: the S transform. *IEEE Trans signal Process* 44(4):998–1001. <https://doi.org/10.1109/78.492555>
- Santos-Mayo L, San-José-Revueña LM, Arribas JI (2016) A computer-aided diagnosis system with EEG based on the P3b wave during an auditory odd-ball task in schizophrenia. *IEEE Trans Biomed Eng* 64(2):395–407. <https://doi.org/10.1109/tbme.2016.2558824>
- Devijver PA, Kittler J (1982) Pattern recognition: a statistical approach. Prentice hall, Hoboken
- Aslan Z, Akin M (2019) Detection of schizophrenia on EEG signals by using relative wavelet energy as a feature extractor. *Proceedings Book*
- Thilakvathi B, Shenbaga Devi S, Bhanu K, Malaippan M (2017) EEG signal complexity analysis for schizophrenia during rest and mental activity. *Biomed Res—India* 28:1–9
- Piryatinska A, Darkhovsky B, Kaplan A (2017) Binary classification of multichannel-EEG records based on the e-complexity of

- continuous vector functions. *Comput Methods Programs Biomed* 152:131–139. <https://doi.org/10.1016/j.cmpb.2017.09.001>
35. Sui J et al (2014) Combination of fMRI-SMRI-EEG data improves discrimination of schizophrenia patients by ensemble feature selection. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014. IEEE, Chicago. <https://doi.org/10.1109/embc.2014.6944473>
  36. Boostani R, Sadatnezhad K, Sabeti M (2009) An efficient classifier to diagnose of schizophrenia based on the EEG signals. *Expert Syst Appl*. <https://doi.org/10.1016/j.eswa.2008.07.037>
  37. Siuly S, Khare SK, Bajaj V, Wang H, Zhang Y (2020) A computerized method for automatic detection of schizophrenia using EEG signals. *IEEE Trans Neural Syst Rehabil Eng*. <https://doi.org/10.1109/tnsre.2020.3022715>
  38. Phang CR, Ting CM, Noman F, Ombao H (2019) Classification of EEG-based brain connectivity networks in schizophrenia using a multi-domain connectome convolutional neural network,” *arXiv*. <https://arxiv.org/ct?url=https%3A%2F%2Fdx.doi.org%2F10.1109%2FJBHI.2019.2941222&v=18340120>
  39. Aslan Z, Akin M (2020) Automatic detection of schizophrenia by applying deep learning over spectrogram images of EEG signals. *Trait du Signal*. <https://doi.org/10.18280/ts.370209>
  40. Borisov SV, Kaplan AY, Gorbachevskaya NL, Kozlova IA (2005) Analysis of EEG structural synchrony in adolescents with schizophrenic disorders. *Hum Physiol*. <https://doi.org/10.1007/s10747-005-0042-z>
  41. Olejarczyk E, Jernajczyk W (2017) Graph-based analysis of brain connectivity in schizophrenia. *PLoS ONE* 12(11):e0188629. <https://doi.org/10.1371/journal.pone.0188629>
  42. Aslan Z (2019) On the use of deep learning methods on medical images. *Int J Energy Eng Sci* 3(2):1–15
  43. Hubel DH, Wiesel TN (1968) Receptive fields and functional architecture of monkey striate cortex. *J Physiol* 195(1):215–243. <https://doi.org/10.1113/jphysiol.1968.sp008455>
  44. Min S, Lee B, Yoon S (2017) Deep learning in bioinformatics. *Brief Bioinform*. <https://doi.org/10.1093/bib/bbw068>
  45. Litjens G et al (2017) A survey on deep learning in medical image analysis. *Med Image Anal*. <https://doi.org/10.1016/j.media.2017.07.005>
  46. Goodfellow BYI (2016) Courville a—deep learning-MIT (2016). *Nature*
  47. Tindall L, Luong C, Saad A (2015) Plankton classification using vgg16 network
  48. WEISANG, “Continuous Wavelet Transform (CWT),” 2020. [Online]. Available at [https://www.weisang.com/en/documentation/timefreqspectralalgorithmcwt\\_en/](https://www.weisang.com/en/documentation/timefreqspectralalgorithmcwt_en/). Accessed on 26 Oct 2021
  49. Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. *arXiv Prepr. arXiv1412.6980*
  50. Kotikalapudi R et al (2017) Keras-vis. GitHub
  51. Johannesen JK, Bi J, Jiang R, Kenney JG, Chen C-MA (2016) Machine learning identification of EEG features predicting working memory performance in schizophrenia and healthy adults. *Neuropsychiatr Electrophysiol*. <https://doi.org/10.1186/s40810-016-0017-0>
  52. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D (2017) Grad-cam: visual explanations from deep networks via gradient-based localization. In: *Proceedings of the IEEE international conference on computer vision*. pp 618–626. <https://doi.org/10.1007/s11263-019-01228-7>
  53. Simonyan K, Vedaldi A, Zisserman A (2013) Deep inside convolutional networks: visualising image classification models and saliency maps. *arXiv Prepr. arXiv1312.6034*
  54. D\_believer (2020) What is Saliency Map? <https://www.geeksfors.org/what-is-saliency-map/>. Accessed 24 Sept 2020

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.