Schizophrenia Identification Through Deep Learning on Spectrogram Images

Amarana Prabhakara Rao[1], G. Prasanna Kumar[2](✉), Rakesh Ranjan[3], M. Venkata Subba Rao[4], M. Srinivasulu[2], and E. Sravya[2]

[1] Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, Bowrampet, Hyderabad 500043, Telangana, India
[2] Department of ECE, Vishnu Institute of Technology, Bhimavaram 534202, Andhra Pradesh, India
godiprasanna@gmail.com
[3] Department of ECE, NIT Patna, Patna, India
[4] Department of ECE, Shri Vishnu Engineering College for Women, Bhimavaram 534202, Andhra Pradesh, India

Abstract. Schizophrenia (SZ) is one of the mental disorder due to which many people are suffering around the world. People suffering with this disorder experience hallucinations, delusions, confusing speech and thinking patterns, etc. In a clinical environment, doctors judge Schizophrenia directly using electroencephalogram (EEG). Automatic detection of SZ is achieved in earlier works by using the time domain and frequency domain features extracted from the given EEG signals. These features are used to train various Machine Learning and Deep Learning approaches for the classification of SZ from the given EEG signal. The proposed work uses Short-Time Fourier Transform (STFT) for converting 1D EEG data into 2D spectrogram image data. This work proposes a simple Convolutional Neural Network (CNN) model for the efficient detection of SZ from the given spectrograms. Performance of the proposed CNN model is compared with various existing CNNs such as Alex net, VGG16, Resnet. Performance of these CNNs is evaluated in terms of accuracy, precision, recall and F1 Score. It is observed from the results that the proposed CNN performed better showing its potential for efficient detection of SZ.

Keywords: Schizophrenia · Electroencephalogram · Spectrogram · Short-time Fourier Transform · Convolutional Neural Networks · Deep Learning Approaches

1 Introduction

There are many mental disorders for human beings in which SZ is one of the important mental disorder due to which many people are suffering [1, 2]. It changes the way a person thinks and behaves. People suffering with this disorder experience hallucinations, delusions, their speech and thinking patterns are confusing, they want to disconnect

P. Pareek et al. (Eds.): IC4S 2023, LNICST 536, pp. 3–11, 2024.
https://doi.org/10.1007/978-3-031-48888-7_1

from people around them including their dear ones, not even care about their personal hygiene, etc. [3]. In a clinical environment, doctors judge Schizophrenia directly using electroencephalogram (EEG). EEG is the most commonly used signal to analyze the mental condition of human beings. EEG signals record various mental conditions of the brain such as mental stress and other disorders. EEGs are becoming popular in recent years in research and diagnosis of various neurological disorders such as Epilepsy, Schizophrenia, etc. These signals contain significant amount of information of higher dimensions and they show complex functioning of the brain and they are very difficult to analyze directly [1]. EEGs provide more detailed information over other existing methods as far as Schizophrenia is concerned [4]. Researchers have proposed both time features [5] and frequency features [6] extracted from EEG signals to detect the state changes in the brain and detect Schizophrenia.

In the past few years, computer aided diagnosis supported by machine learning (ML) algorithms have revolutionized the study of complex EEG signals using various time-domain and frequency domain features to identify the schizophrenia. Several researchers have used ML based framework in the diagnosis of various diseases such as epileptic Seizures, Schizophrenia, etc. [7]. Many DL approaches are invented today for classification and segmentation in today's world, because DL gives the better performance compared to ML techniques. This work mainly focuses on DL approaches and proposes a simple CNN model for efficient detection of SZ.

First part of this paper is an Introduction section which gives an insight into the global burden of Schizophrenia and how it can be detected at an early stage by using EEG signals which can be analyzed efficiently with various computer aided diagnostic tools involving various ML and DL frameworks. Second part of the paper gives a brief review of the earlier works done in this domain. Next section – materials and methods present the details of data used to train the deep learning models and details of proposed deep learning model. In the next section, results are presented and a summary of these results are discussed in this section. In the last section conclusions are presented.

2 Literature Review

In this section, a summary of the research work done by the researches in detecting Schizophrenia using EEG signals analyzed by various traditional ML algorithms as well as trending DL models. In a recent study Miras et al. [8] have performed various linear and non-linear measures to extract seventeen features from the resting stage EEG signals. These selected features were used as inputs to five traditional ML based algorithms to analyze SZ. This study includes 31 patients out of which 20 were healthy controls and 11 were SZ patients. In another study, Ranjan et al. [9] applied Kruskal Wallis test to select 8 significant features out of the 24 features extracted from 16 channel EEG signals of 84 subjects. These features were applied to various classical ML algorithms and they have reported that ensemble bagging tree classifier performed better with an accuracy of 92.3%.

A hybrid framework combining brain-effective connectivity analysis [10] and deep learning is proposed for schizophrenia detection using EEG signals. Transfer Entropy measures causalities between EEG channels, forming connectivity images. These images are inputted into pre-trained CNN models (VGG-16, ResNet50V2, InceptionV3, EfficientNetB0, DenseNet121), and their deep features are processed by an LSTM model. The hybrid CNN-LSTM models achieve exceptional accuracy, with EfficientNetB0-LSTM reaching 99.90% average accuracy and 99.93% F1-score using 10-fold cross-validation. The method shows strong capability in detecting schizophrenia patients from healthy controls.

Researchers developed an interpretable machine learning method [11] for diagnosing schizophrenia based on DSM-5 criteria. The method integrates smoothly into existing clinical processes, providing clinicians with trust and understanding. By combining two attention mechanisms, attribute importance and interactivity are determined. The model demonstrates the robustness and real-world applicability through experiments with augmented test data. It achieves a high accuracy of 98% with 10-fold cross-validation, highlighting its effectiveness in assisting clinicians with schizophrenia diagnosis.

A study by Zulfikar Aslan [12] introduces a highly accurate method for automatically detecting schizophrenia from EEG records. It utilizes the Continuous Wavelet Transform to extract time-frequency features from the signals, achieving the highest accuracy in the literature. The VGG16 deep learning architecture is employed for feature extraction, resulting in classification accuracies of 98% for SZ patients and 99.5% for healthy individuals. Visualizations of the CNN network highlight frequency component differences between SZ patients and healthy individuals, facilitating easy interpretation. The performance of the model was evaluated in terms of accuracy, sensitivity, specificity, and F1 score. The results demonstrated that the deep learning approach achieved high accuracy and reliable classification performance, outperforming traditional machine learning methods and indicating the potential of CNN-based models in the accurate identification of schizophrenia using EEG signals.

ML based works trained with various algorithms of SVM, KNN and RF. These works achieved maximum accuracy of 92.5 [13]. The researchers extracted various features from the EEG data and applied a SVM classifier to differentiate between SZ patients and Normal people. The proposed framework achieved an accuracy of 90.3% in the classification task. In [14] several ML algorithms, including SVM, KNN, RF, and ANN, were trained and evaluated. The SVM algorithm achieved the accuracy, sensitivity and specificity of 86.5%, 85.7% and 87.2% respectively and an AUC-ROC of 0.907, indicating its effectiveness in distinguishing individuals with schizophrenia from healthy controls based on EEG signals.

The literature shows the work progress of EEG classification using ML and DL methods. The DL methods are superior than the ML methods. This work mainly focusses on DL method in detection of SZ using spectrogram-based EEGs.

This paper is organized as follows, Sect. 1 presents introduction and motivation of work, Sect. 2 describes the literature of the proposed work and some important findings in the existing work, Sect. 3 gives the detailed description on methodology, dataset used and the proposed CNN architecture. Results and comparisons are briefly described in Sect. 4 and Sect. 5 gives the conclusions and future scope of the work.

3 Proposed Methodology

The proposed methodology mainly has three important stages that is collecting EEG data set for Norm and SZ, converting the EEG 1D data into 2D image data using STFT which is called as spectrogram and customized CNN for classification. Details of proposed work is illustrated in a block diagram as shown in Fig. 1. The EEG data acquired from source has three conditions, condition-1: hitting a button causes an audible sound to be generated immediately, condition-2: Listening the same audible sound passively, and condition-3: hitting a button does not cause a sound to be generated.

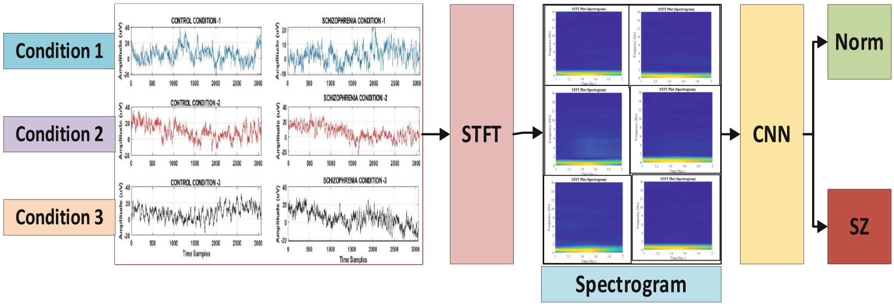

Fig. 1. Block Diagram of the proposed method

The next important phase of the work is converting 1D signal to 2D image. The EEG signals acquired from the patients and normal persons are converted into spectrogram using STFT. The STFT is a Fourier transform extension that offers information on the components of a signal in the frequency domain. Main idea behind the STFT is to divide a signal into short segments or windows and apply the Fourier transform to each window individually. By doing so, we can examine the frequency content of the signal with a small duration of time. Resulting STFT representation displays how the signal's frequency components evolve over different time intervals. It provides valuable information about the spectral content of a signal at different points in time, which is essential for understanding and manipulating signals with time-varying characteristics. The mathematical expression for STFT with an input x[n] is given below in Eq. (1).

$$STFT(x[n]) = X\left[m, f\right] = \sum\nolimits_{n=-\infty}^{\infty} x[n]w[n-m]e^{-jfn} \tag{1}$$

where w[n] is the window function with a small duration and sliding along the time axis until the last point of the signal. This mechanism gives the time and frequency representation of a signal. The STFT of EEG signals under Norm and SZ is shown in Fig. 1.

Customized CNN

The proposed CNN is a customized CNN, which consists of input layer, hidden and fully connected layers. CNN mainly consists of convolution layers, which is the core component of CNN. Convolution layer is a fundamental building block in CNNs, which are a type of DL architecture commonly used for image and video processing tasks. The convolutional layer performs a convolution operation on the input data to extract local features and learn hierarchical representations.

There are filters or kernels in CNN, which are small-sized matrices. The output of each convolution layer is convolution sum of input with filter weights. This involves element-wise multiplication of the filter with a local receptive field of the input data, followed by summation of the multiplied values. Equation (2) shows the mathematical expression for convolution in discrete domain.

$$y[m,n] = \sum_{k=-\infty}^{\infty} \sum_{r=-\infty}^{\infty} x[k,r]h[m-k,n-r] \quad (2)$$

where h[m,n] is the filter weights of the convolution layer. During the convolution operation, the filter is slid over the input data in a sliding window manner, computing the element-wise products and summing them up to generate a single value in the corresponding location of the feature map. This process is repeated for every possible receptive field in the input, resulting in feature maps.

Max Pooling Layer

This layer is to extract the most salient features from each local neighborhood of the input feature map. It divides the feature map into non-overlapping rectangular parts known as pooling window or pooling regions, and then substitutes the highest number inside each region.

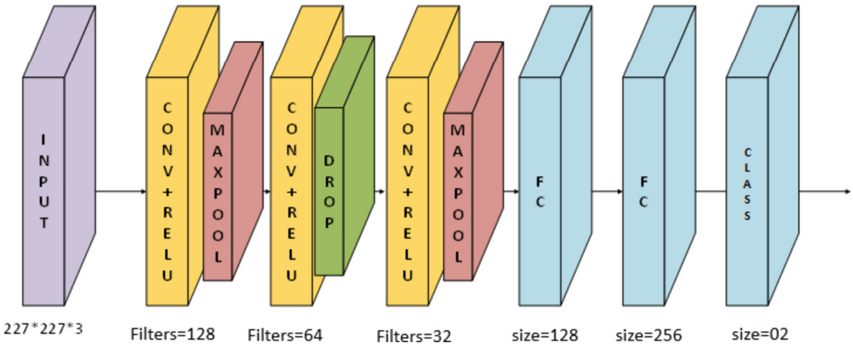

Fig. 2. Customized CNN for training and testing

The customized CNN is shown in Fig. 2, it consists of input layer of size $227 \times 227 \times 3$, followed by convolution layers and Relu activation function and for pooling max pooling is used. At the middle e network dropout layer in order to avoid the overfitting problem finally fully connected layers and classification layer of size 02 for binary classification.

The training and testing can be done by using ten-fold cross validation. The process of tenfold validation can be summarized as follows. Data Splitting: Initially the dataset portioned into ten subsets, or folds of roughly equal size. Each fold should ideally contain a representative sample of the data.

Model Training and Evaluation: The main module model training, here in 9th fold model is trained and reaming for validation. Ten times doing the same process validate the set once. In each iteration, the model is trained on a different combination of folds. Performance Measurement: After each training iteration, the model's performance is evaluated using a chosen performance metric on the validation set.

Average Performance: The performance results from the ten iterations are usually averaged to obtain a single performance metric that represents the model's performance across the entire dataset. This average performance metric is used to compare and assess the model's generalization ability. By utilizing ten-fold cross-validation, the potential bias in model evaluation caused by the specific partitioning of the data is reduced. It provides a more robust estimate, as it evaluates the model on multiple diverse subsets of the data.

4 Results

The proposed method uses a customized CNN for training and testing. The EEG data of 734 normal samples and 874 SZ effected samples are used for classification. The data split for training and validation is 0.9 and 0.1 respectively. The optimizer used in the proposed work is SGDM, RMS Prop and Adams optimizer. The number of Epochs is 6 with a BS of 50 is used as training parameters. The validations accuracy obtained after training process under various optimizers is tabulated in Table 4, Table 5 and Table 6. It shows that validation accuracy of adams optimizer achieves best results than other two, which is for Alex net 82.1, VGG16 83.1, Resnet 83.6 and for proposed customized net it is 86.3. This indicates that the proposed customized net under Adams optimizer gives higher validation accuracy compared to standard existing methods.

Table 1, Table 2 and Table 3 shows the testing results, which is confusion matrix comparison of proposed method with existing methods under different optimizers. In which 0 represents normal data and 1 represents SZ effected data. It shows that proposed net has higher prediction accuracy compared to existing nets.

Table 1. Confusion matrix comparison with SGDM optimizer

	Alex net		**VGG16**		**Resnet**		**Proposed CNN**	
	0	1	0	1	0	1	0	1
0	**128**	**21**	**126**	**23**	**125**	**24**	**130**	**19**
1	**68**	**91**	**67**	**92**	**64**	**95**	**62**	**97**

Table 2. Confusion matrix comparison with RMSprop optimizer

	Alex net		VGG16		Resnet		Proposed CNN	
	0	1	0	1	0	1	0	1
0	**142**	**7**	**143**	**6**	**146**	**3**	**149**	**0**
1	**78**	**81**	**77**	**82**	**2**	**87**	**68**	**91**

Table 3. Confusion matrix comparison with Adams optimizer

	Alex net		VGG16		Resnet		Proposed CNN	
	0	1	0	1	0	1	0	1
0	**143**	**6**	**144**	**5**	**145**	**4**	**148**	**1**
1	**2**	**157**	**1**	**158**	**2**	**157**	**1**	**158**

The performance metrics for comparing proposed CNN with existing CNN is by the following metrics. These parameters are calculated from confusion matrix. These are formulated in Eqs. (3–6) as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{3}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{4}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{5}$$

$$\text{F1 Score} = \frac{2(\text{Precision} * \text{Recall})}{\text{Precision} + \text{Recall}} \tag{6}$$

The performance metric comparison is tabulated in Table 4, Table 5 and Table 6. The following are observations made from results. The proposed customized CNN has higher validation accuracy, test accuracy, precision, Recall and F1 score compared to existing CNNs.

Table 4. Performance metric comparison with SGDM optimizer

	Validation Accuracy	Test Accuracy	Precision	Recall	F1 score
Alexnet	65.75	71.11	85.91	65.30	74.20
VGG16	67.32	70.77	84.56	65.28	73.67
ResNet50	66.28	71.42	83.86	66.13	73.94
CNN	**72.50**	**73.71**	**87.24**	**67.71**	**76.24**

Table 5. Performance metric comparison with RMS prop optimizer

	Validation Accuracy	Test Accuracy	Precision	Recall	F1 score
Alexnet	66.84	72.40	95.31	64.54	76.96
VGG16	68.43	73.03	95.97	65.28	77.70
ResNet50	72.21	75.64	97.78	66.97	79.49
CNN	**80.52**	**77.92**	**100**	**67.71**	**80.74**

Table 6. Performance metric comparison with Adams optimizer

	Validation Accuracy	Test Accuracy	Precision	Recall	F1 score
Alexnet	82.1	97.41	95.73	98.62	96.54
VGG16	83.3	98.05	96.64	99.31	97.95
ResNet50	83.6	98.13	97.2	98.82	98.34
Proposed CNN	**86.3**	**99.35**	**99.32**	**99.31**	**99.31**

The following are the advantages of the proposed method compared to existing networks

1. Achieves higher validation accuracy of 86.3.
2. Achieves higher testing accuracy of 99.35 with a three-convolution layer network
3. The proposed model is very simple and training time is also low
4. The proposed method has a potential to replace traditional machine learning frame work.

5 Conclusions

This work mainly focuses on identification of SZ at early stages using deep learning models. The EEG signals of Normal and SZ samples are collected from an open-source data set and these EEG signals are converted into image spectrogram using STFT. The spectrogram of the Normal and SZ samples is trained by using standard CNNs such as Alexnet, VGG16 and Resnet. This paper proposes a simple customized CNN with three convolution layers for training, which is less complex and high speed. The proposed CNN gives better validation accuracy in training and testing as compared to the existing CNN models. The results exhibited the potential of the proposed CNN model in detecting SZ efficiently. In future, this work can be extended in two ways: (i) By taking scalogram of EEG data for training the proposed CNN and (ii) By classifying multiple SZ conditions using Spectrogram and scalogram.

References

1. Sun, J., et al.: A hybrid deep neural network for classification of schizophrenia using EEG data. Sci. Rep. **11**, 1–16 (2021)
2. Prabhakara Rao, A., Prasanna Kumar, G., Ranjan, R.: Performance comparison of classification models for identification of breast lesions in ultrasound images BT - pattern recognition and data analysis with applications. In: Gupta, D., Goswami, R.S., Banerjee, S., Tanveer, M., Pachori, R.B. (eds.) Pattern Recognition and Data Analysis with Applications. LNEE, vol. 888, pp. 689–699. Springer, Heidelberg (2022). https://doi.org/10.1007/978-981-19-1520-8_56
3. Matalam, C.L., Hembra, M.S.: Community-based mental health project in Davao region. SPMC J. Health Care Serv. **8**(2), 5 (2022)
4. Miyauchi, T., et al.: Computerized EEG in schizophrenic patients. Biol. Psychiatry **28**, 488–494 (1990)
5. Bonita, J.D., et al.: Time domain measures of inter-channel EEG correlations: a comparison of linear, nonparametric and nonlinear measures. Cogn. Neurodyn. **8**, 1–15 (2014)
6. Singh, K., Singh, S., Malhotra, J.: Spectral features based convolutional neural network for accurate and prompt identification of schizophrenic patients. Proc. Inst. Mech. Eng. Part H J. Eng. Med. **235**, 167–184 (2021)
7. Rao, A.P., Bhaskar, J., Kumar, G.P.: Machine learning framework for identification of abnormal EEG signal. In: Gupta, N., Pareek, P., Reis, M. (eds.) IC4S 2022. LNICST, vol. 472, pp. 42–54. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-28975-0_4
8. de Miras, J.R., Ibáñez-Molina, A.J., Soriano, M.F., Iglesias-Parro, S.: Schizophrenia classification using machine learning on resting state EEG signal. Biomed. Signal Process. Control **79**, 104233 (2023)
9. Ranjan, R., Sahana, B.C.: A machine learning framework for automatic diagnosis of schizophrenia using EEG signals. In: 2022 IEEE 19th India Council International Conference (INDICON), pp. 1–6. IEEE (2022)
10. Bagherzadeh, S., Shahabi, M.S., Shalbaf, A.: Detection of schizophrenia using hybrid of deep learning and brain effective connectivity image from electroencephalogram signal. Comput. Biol. Med. **146**, 105570 (2022)
11. Organisciak, D., Shum, H.P.H., Nwoye, E., Woo, W.L.: RobIn: a robust interpretable deep network for schizophrenia diagnosis. Expert Syst. Appl. **201**, 117158 (2022)
12. Aslan, Z., Akin, M.: A deep learning approach in automated detection of schizophrenia using scalogram images of EEG signals. Phys. Eng. Sci. Med. **45**, 83–96 (2022)
13. Buettner, R., et al.: High-performance exclusion of schizophrenia using a novel machine learning method on EEG data. In: 2019 IEEE International Conference on E-Health Networking, Application & Services (HealthCom), pp. 1–6. IEEE (2019)
14. Siuly, S., Khare, S.K., Bajaj, V., Wang, H., Zhang, Y.: A computerized method for automatic detection of schizophrenia using EEG signals IEEE Trans. Neural Syst. Rehabil. Eng. **28**(2390), 2400 (2020)