SchizoNET: a robust and accurate Margenau–Hill time-frequency distribution based deep neural network model for schizophrenia detection using EEG signals

IOP Publishing Physiol. Meas. 44 (2023) 035005 https://doi.org/10.1088/1361-6579/acbc06

Physiological Measurement

IPEM Institute of Physics and Engineering in Medicine

OPEN ACCESS

RECEIVED
12 October 2022

REVISED
28 January 2023

ACCEPTED FOR PUBLICATION
14 February 2023

PUBLISHED
8 March 2023

PAPER

SchizoNET: a robust and accurate Margenau–Hill time-frequency distribution based deep neural network model for schizophrenia detection using EEG signals

Smith K Khare[1], **Varun Bajaj**[2] and **U Rajendra Acharya**[3,4,5,6,7]

[1] Electrical and Computer Engineering Department, Aarhus University, Denmark
[2] Discipline of Electronics and Communication Engineering, Indian Institute of Information Technology, Design, and Manufacturing (IIITDM) Jabalpur, India
[3] School of Mathematics, Physics, and Computing, University of Southern Queensland, Springfield, Australia
[4] Department of Biomedical Engineering, School of Science and Technology, University of Social Sciences, Singapore
[5] Department of Biomedical Informatics and Medical Engineering, Asia University, Taiwan
[6] Distinguished Professor, Kumamoto University, Japan
[7] Adjunct Professor, University of Malaya, Malaysia

E-mail: smith7khare@gmail.com, varunb@iiitdmj.ac.in and aru@np.edu.sg

Keywords: electroencephalogram classification, schizophrenia detection, convolutional neural networks, decision support system

Abstract

Objective. Schizophrenia (SZ) is a severe chronic illness characterized by delusions, cognitive dysfunctions, and hallucinations that impact feelings, behaviour, and thinking. Timely detection and treatment of SZ are necessary to avoid long-term consequences. Electroencephalogram (EEG) signals are one form of a biomarker that can reveal hidden changes in the brain during SZ. However, the EEG signals are non-stationary in nature with low amplitude. Therefore, extracting the hidden information from the EEG signals is challenging. *Approach.* The time-frequency domain is crucial for the automatic detection of SZ. Therefore, this paper presents the SchizoNET model combining the Margenau–Hill time-frequency distribution (MH-TFD) and convolutional neural network (CNN). The instantaneous information of EEG signals is captured in the time-frequency domain using MH-TFD. The time-frequency amplitude is converted to two-dimensional plots and fed to the developed CNN model. *Results.* The SchizoNET model is developed using three different validation techniques, including holdout, five-fold cross-validation, and ten-fold cross-validation techniques using three separate public SZ datasets (Dataset 1, 2, and 3). The proposed model achieved an accuracy of 97.4%, 99.74%, and 96.35% on Dataset 1 (adolescents: 45 SZ and 39 HC subjects), Dataset 2 (adults: 14 SZ and 14 HC subjects), and Dataset 3 (adults: 49 SZ and 32 HC subjects), respectively. We have also evaluated six performance parameters and the area under the curve to evaluate the performance of our developed model. *Significance.* The SchizoNET is robust, effective, and accurate, as it performed better than the state-of-the-art techniques. To the best of our knowledge, this is the first work to explore three publicly available EEG datasets for the automated detection of SZ. Our SchizoNET model can help neurologists detect the SZ in various scenarios.

1. Introduction

Schizophrenia (SZ) is a complex, neuropsychiatric and cognitive syndrome that appears to result from a disruption in brain development caused by hereditary or environmental factors, or both. SZ disturbs the thinking, behaviour, and feeling of an individual. According to the reports published by the World Health Organization (WHO), about 21 million people accounting for 1% of the global population are suffering from SZ (WHO 2022). The onset of SZ typically occurs between late adolescence to the beginning of early adulthood. It emerges earlier in males (early 20 s—late adolescence) than in females (early 20 s—early 30 s)

(Bromet and Fennig 1999). It is one of the top 25 leading causes of worldwide disability (Jin and Mosweu 2017). The symptoms of SZ are heterogeneous that leading to reduced quality of life and functional impairments (Jin and Mosweu 2017). It is characterized by cognitive deficits, and negative and positive symptoms (Green and Horan 2015). The cognitive deficits involve language (difficult to understand for others), difficulty performing routine activities, lack of attention, and trouble with thinking (deviating from one subject to another with no logical reason). The negative symptoms (associated with negative SZ) are abnormal memory while the positive symptoms (associated with positive SZ) are hallucinations, delusions, and confused speech (Oh *et al* 2019, Lai *et al* 2021, Sadeghi *et al* 2022). The epidemiological characteristics of SZ have three lows (low visit rate, low detection rate, and low compliance) and three highs (high disability rate, heavy disease burden, and high recurrence rate). SZ could result in damage to various brain tissues as well as mental deterioration, resulting in severe mental disability. As a result, SZ has a negative impact on educational and occupational performance. The possibility of death in SZ is higher than that of healthy people due to physically preventable diseases (cardiovascular disorder, infections, and metabolic disease) (Siuly *et al* 2020). Suicide attempts among SZ patients are about 50%, with a mortality rate due to suicide being about 4%–6% (Caldwell and Gottesman 1990, Hettige *et al* 2017). About 69% of SZ patients do not get enough care and treatment resulting in an increased death rate, disability rate, and suicide rate (Baygin 2021). According to the WHO, timely detection of SZ may help experts to identify the stage and severity of SZ (WHO 2022). These factors demand a need for timely and accurate detection of SZ. Various resources such as interviews, imaging, and signaling techniques have been used to detect SZ. Interviewing by a qualified expert takes time which is susceptible to errors and biased in some cases (Lloyd *et al* 2017). Imaging tools (magnetic resonance imaging and computed tomography) are time-consuming, more expensive, and necessitate extra recordings (Talo *et al* 2019). Electroencephalogram (EEG) signals can reveal changes in brain activity to identify various states of the brain (Khare and Bajaj 2021c, Khare *et al* 2022). During EEG recording, sensors placed at the appropriate location on the scalp extract secret information about changes during SZ. In addition, researchers are well accepted by EEG signals in the automated identification of brain disorders such as Alzheimer's disease, seizures, and Parkinson's disease (Kumar and Bhuvaneswari 2012, Khare *et al* 2022).

2. Related work

Recently, many studies have been developed to get insights into the automatic classification of SZ using EEG as a biomarker. The summary of the existing models developed for SZ detection is shown in table 1.

3. Findings and motivation

The summary of our findings by screening various literature is shown in table 2. It reveals that most of the automated SZ detection models have been developed on one EEG dataset either in a resting state or performing some task. Also, we noted that models like LSTM and 1D CNN help to extract Spatio-temporal information but with reduced performances compared to 2D CNN models (Cho and Jang 2020, Vareka 2021). CNN allows automatic feature extraction and classification but applying non-stationary EEG signals directly to CNN may not reveal desired performance. Over the last decade, many CNN models have been developed whose architecture varied from tens to hundreds of layers. But there is no standard CNN model specified for a particular application. Therefore, the selection and development of the CNN model depend on the user and applications. Even with deep models like visual geometry group and ResNet, the desired performance is not obtained (Smith *et al* 2021). Also, these techniques involve handcrafted feature extraction, empiric selection of tuning parameters, rigorous statistical analysis for feature selection, and user-dependent classification techniques resulting in the lower performance of models.

Therefore, from the identified research gaps we have been motivated to develop a SchizoNET model comprised of Margenau–Hill time-frequency distribution (MH-TFD) and CNN. The TFD helps to study detailed insight into EEG signals by capturing minute details in terms of time-frequency-amplitude contents. A CNN model with a simple architecture is developed using multiple validation techniques including holdout, five-, and ten-fold cross-validation (FCV) techniques to extract and classify the deep features obtained from TFD. The proposed model is tested and evaluated on three public EEG datasets of SZ. The working steps of the SchizoNET model are as follows: (i) the temporal information of EEG signals about time-frequency-amplitude is extracted from MH-TFD, (ii) the TFD is fed to the developed CNN model for automated feature extraction and classification, (iii) different performance parameters are evaluated and compared them with the current state-of-the-art techniques. The contribution of the proposed SchizoNET model is listed as follows:

Table 1. Summary of the existing automated SZ decision models using EEG signals.

Author and Year	Subjects	Tasks	Analysis technique	Classification	Performance
Yin *et al* (2017)	SZ (Negative): 14 SZ (Positive): 14 HC: 14	Resting state eye closed	mutual information using Shannon and joint entropy	Statistical analysis	—
Alimardani *et al* (2018)	SZ: 23 BP: 23	Steady state evoked potential	Statistical and nonlinear features	k-nearest neighbor (KNN)	ACC: 91.30%
Hiesh *et al* (2013)	SZ: 5 HC: 5	Tasks (auditory simulations)	Wavelet transform (WT) with statistical (mean, minima, maxima, and standard deviation (STD)) and nonlinear features	Support vector machine (SVM)	ACC: 88.24% SEN: 89.48% SPE: 87%
Begić *et al* (2000)	SZ (Negative): 22 SZ (Positive): 25 HC: 50	Resting state eye closed	fast Fourier transform (FFT) for rhythmic analysis	Statistical analysis	—
Namazi *et al* (2019)	SZ: 45 HC: 39	Resting-state from adolescents	Fractal dimension (FD) and approximate entropy	Statistical analysis	—
Akar *et al* (2016)	SZ: 22 HC: 22	Resting state eye closed	Entropy and complexity features	Statistical analysis	—
Dvey-Aharon *et al* (2015)	SZ: 25 HC: 25	Tasks (stimuli of triangle)	Time-frequency feature optimization-based feature extraction	Linear discriminant analysis	ACC: 88.7% SEN: 77.4% SPE: 100%
Parvinnia *et al* (2014)	SZ: 13 HC: 18	Resting state eye open	Band power, autoregressive model (ARM), and FD	Weighted distance nearest neighbor	ACC: 95.3%
Sabeti *et al* (2009)	SZ: 20 HC: 20	Resting state eye open	Higuchi dimensions, Lempel Ziv complexity, and entropy	AdaBoost	ACC: 90%
Sui *et al* (2014)	SZ: 48 HC: 53	Resting state eye open	Multi-set canonical correlation analysis	SVM	ACC: 74% SEN: 72% SPE: 75%
Phang *et al* (2019)	SZ: 45 HC: 39	Resting-state from adolescents	Vector-autoregression model (VAM)-based directed connectivity, and graph-theoretical complex network	Deep neural network	ACC: 95%

Table 1. (Continued.)

Author and Year	Subjects	Tasks	Analysis technique	Classification	Performance
Piryatinska *et al* (2017)	SZ: 45 HC: 39	Resting-state from adolescents	ϵ-complexity of continuous vector functions of original EEG signals and their finite differences	Random forest (RF)	ACC: 85.3% SEN: 88.6% SPE: 82.6%
Singh *et al* (2021)	SZ: 45 HC: 39	Resting-state from adolescents	Mean spectral amplitude, spectral power and Hjorth descriptors	Convolutional neural network (CNN)	ACC: 94.08% SEN: 92.7% SPE: 95.31%
	SZ: 14 HC: 14	Resting state eye closed	Mean spectral amplitude, spectral power and Hjorth descriptors	CNN	ACC: 98.96% SEN: 99.05% SPE: 98.88%
Dimitriadis (2021)	SZ: 45 HC: 39	Resting-state from adolescents	Dynamic correlation of the envelope (corrEnv)	SVM	ACC: 100% SEN: 100% SPE: 100%
Calhas (2019)	SZ: 45 HC: 39	Resting-state from adolescents	Discrete short-time Fourier transform (DSTFT)	RF	ACC: 84% SEN: 87% SPE: 82%
Phang *et al* (2020)	SZ: 45 HC: 39	Resting-state from adolescents	VAM, and partial directed coherence (PDC)	CNN	ACC: 91.69% SEN: 91.11% SPE: 92.5%
Siuly *et al* (2020)	SZ: 49 HC: 32	Tasks (push-button task)	Empirical mode decomposition (EMD) and statistical features	SVM	ACC: 89.59% SEN: 89.76% SPE: 89.32%
Krishnan *et al* (2020)	SZ: 14 HC: 14	Resting state eye closed	Multivariate EMD	SVM	ACC: 93% SEN: 94% SPE: 92%
Khare and Bajaj (2021a)	SZ: 49 HC: 32	Tasks (push-button task)	Optimized variational mode decomposition (OVMD)	Optimized extreme learning machine (OELM)	ACC: 92.93% SEN: 97.15% SPE: 91.06%

Table 1. (Continued.)

Author and Year	Subjects	Tasks	Analysis technique	Classification	Performance
Aslan and Akin (2020)	SZ: 45 HC: 39	Resting-state from adolescents	Spectrograms obtained using short time Fourier transform (STFT)	CNN	ACC: 95% SEN: 95%
	SZ: 14 HC: 14	Resting state eye closed	STFT	CNN	ACC: 97% SEN: 97%
Nikhil *et al* (2021)	SZ: 14 HC: 14	Resting state eye closed	FD, entropy, variance-based features	long short-term memory (LSTM)	ACC: 99% SEN: 98.9
Shalbaf *et al* (2020)	SZ: 14 HC: 14	Resting state eye closed	Continuous wavelet transform (CWT)	ResNet-SVM	ACC: 95.3% SEN: 96.45% SPE: 94.5%
Prabhakar *et al* (2020b)	SZ: 14 HC: 14	Resting state eye closed	Partial least squares (PLS), expectation-maximization-based principal component analysis (EM-PCA), nonlinear regression, and isometric mapping (Isomap)	RF	ACC: 96.77% SEN: 96.77% SPE: 96.77%
Khare *et al* (2020)	SZ: 49 HC: 32	Tasks (push-button task)	Empirical wavelet transform (EWT)	Ensemble bagged tree (EBT)	ACC: 88.7% SEN: 91.13% SPE: 89.29%
Prabhakar *et al* (2020a)	SZ: 14 HC: 14	Resting state eye closed	Nonlinear features	SVM	ACC: 89.25%
Oh *et al* (2019)	SZ: 14 HC: 14	Resting state eye closed	CNN	CNN	ACC: 98.07% SEN: 97.32% SPE: 98.17%
Jahmunah *et al* (2019)	SZ: 14 HC: 14	Resting state eye closed	Nonlinear features	SVM	ACC: 92.91% SEN: 93.45% SPE: 98.17%
Khare and Bajaj (2021b)	SZ: 49 HC: 32	Tasks (push-button task)	Flexible tunable Q wavelet transform (FTQWT)	Flexible least square SVM (FLSSVM)	ACC: 91.39% SEN: 92.65% SPE: 93.22%

Table 1. (Continued.)

Author and Year	Subjects	Tasks	Analysis technique	Classification	Performance
Smith *et al* (2021)	SZ: 49 HC: 32	Tasks (push-button task)	Smoothed pseudo-Wigner Ville distribution (SPWVD)	CNN	ACC: 93.36% SEN: 94.25% SPE: 92.03%
Sharma and Acharya (2020)	SZ: 14 HC: 14	Resting state eye closed	L1 norm features obtained with optimal wavelet	KNN	ACC: 99.21% SEN: 98.84% SPE: 99.42%
Sun *et al* (2021)	SZ: 54 HC: 55	Resting state eye open	Different rhythms using FFT and fuzzy entropy features	CNN	ACC: 99.22%
Racz *et al* (2020)	SZ: 14 HC: 14	Resting state eye closed	Dynamic functional connectivity (DFC)	RF	ACC: 89.29% SEN: 78.57% SPE: 100%
Goshvarpour and Goshvarpour (2020)	SZ: 14 HC: 14	Resting state eye closed	Complexity, Higuchi FD (HFD), and Lyapunov exponents	Probabilistic neural network (PNN)	ACC: 100% SEN: 100% SPE: 100%
Masychev *et al* (2021)	SC: 57 HC: 66	Tasks (auditory odd-ball paradigm)	Symbolic transfer entropy (STE)	SVM	ACC: 92.68% SEN: 92.98% SPE: 92.42%
Ravan *et al* (2015)	SC: 47 HC: 66	Tasks (auditory odd-ball paradigm)	Brain source localization	Statistical analysis	—
Li *et al* (2019)	SZ: 23 HC: 25	Visual P300 tasks Resting state eye closed	Spatial pattern of network (SPN)	SVM	ACC: 90.48% SEN: 89.47% SPE: 91.3%
Ciprian *et al* (2021)	SZ: 62 HC: 70	Resting state eye closed	Effective connectivity and STE	KNN	ACC: 96.92% SEN: 95% SPE: 98.57%

Table 1. (Continued.)

Author and Year	Subjects	Tasks	Analysis technique	Classification	Performance
Liu *et al* (2017)	SZ: 10 HC: 10	Resting state eye closed	Statistical features	SVM	ACC: 91.16%
Kumar *et al* (2023)	SZ: 45 HC: 39	Resting-state from adolescents	Correlation-based feature selection (CBFS)	AdaBoost	ACC: 92.85% SEN: 93.3% SPE: 92.3%
	SZ: 14 HC: 14	Resting state eye closed	CBFS	AdaBoost	ACC: 99.36% SEN: 99.2% SPE: 99.4%
Aydemir *et al* (2022)	SZ: 14 HC: 14	Resting state eye closed	Cyclic group of prime order pattern	KNN	ACC: 99.82% SEN: 99.84% SPE: 99.81%
Siuly *et al* (2022)	SZ: 49 HC: 32	Tasks (push-button task)	Average filtering	GoogLeNet and SVM	ACC: **98.84**% SEN: **99.02**% SPE: **98.58**%
Baygin *et al* (2023)	SZ: 49 HC: 32	Tasks (push-button task)	Tunable Q wavelet transform (TQWT)	KNN	ACC: 95.84% SEN: 97.01% SPE: 94.06%

*HC (Healthy control)

Table 2. Summary of findings obtained from the state-of-the-art techniques.

Parameters	Findings	Advantages	Limitations
		Dataset	
Signals	The available studies have mostly explored EEG signal analysis either in a resting state with eyes closed/open or with tasks	The acquisition of EEG is fast, portable, and non-invasive	SZ may attack individuals at any age affecting their auditory and motor abilities. Therefore, it is desired to study EEG signals acquired during resting and tasks. But models tested on one dataset do not guarantee desired performance with others
		Signal analysis	
Direct feature analysis	Extraction of statistical features and nonlinear features requires the selection of scaling parameters. Nonlinear (FD, entropy, ARM, and Hjorth parameters) and statistical features are affected by noise	Extracting features directly eliminates additional signal analysis tools	The extraction of features directly from EEG signals fails to find representative information. The performance of the system is degraded due to noise and inappropriate selection of scaling range
Frequency analysis (FFT, filtering, STFT, DSTFT) and rhythm separation	EEG signals are composed of multi-frequency bands that help to study the changes in EEG during SZ	Allows the analysis of exact frequency contents of EEG which may help to reveal hidden characteristics during SZ	Frequency domain analysis provides analysis of exact frequency content but fails to reveal at what time the frequency contents occurred. Also, it suffers time-frequency localization and sharp filter boundaries. DSTFT and STFT assume signals to be stationary and require a choice of window (type and length)
Nonlinear decomposition (EWT, optimal wavelet, TQWT, CWT, EMD and MEMD)	Nonlinear decomposition extracts instantaneous information about time-frequency. It requires to define the basis function with which the signal is represented	It decomposes the signal into multicomponent that provide representative characteristics of the signals	EMD and MEMD are experimental and lack mathematical modelling. EWT, optimal wavelet, and CWT require tuning parameters to get multi-bands
		Decision-making	
Machine learning (SVM, KNN, decision tree, ensemble, neural networks, etc.)	Machine learning models require tuning of hyperparameters and selection of kernel	Decision-making is fast. Availability of different models with distinct cost functions	An inappropriate selection of hyperparameters and kernels may result in decreased performance or overfitting
Deep learning (LSTM, CNN, auto encoder, etc.)	The models like LSTM and 1D CNN necessitate the extraction of Spatiotemporal information but have degraded performances when compared to 2D or image-based CNN models	Enables simultaneous extraction of features and classification. The vast availability of models	No standard model is available for decision-making. Some models are so deep that require a large number of training time and learning parameters. CNN allows automatic feature extraction and classification but applying non-stationary EEG signals directly to CNN may not reveal desired performance

- To the best of our knowledge, we are the first group to develop automated SZ detection on three EEG datasets. Therefore, the SchizoNET model has good generalization ability on different datasets.
- Visual inspection of EEG signals is very tedious and prone to human error. Hence, the study of temporal and spatial information of EEG signals about time-frequency amplitude is performed by MH-TFD.

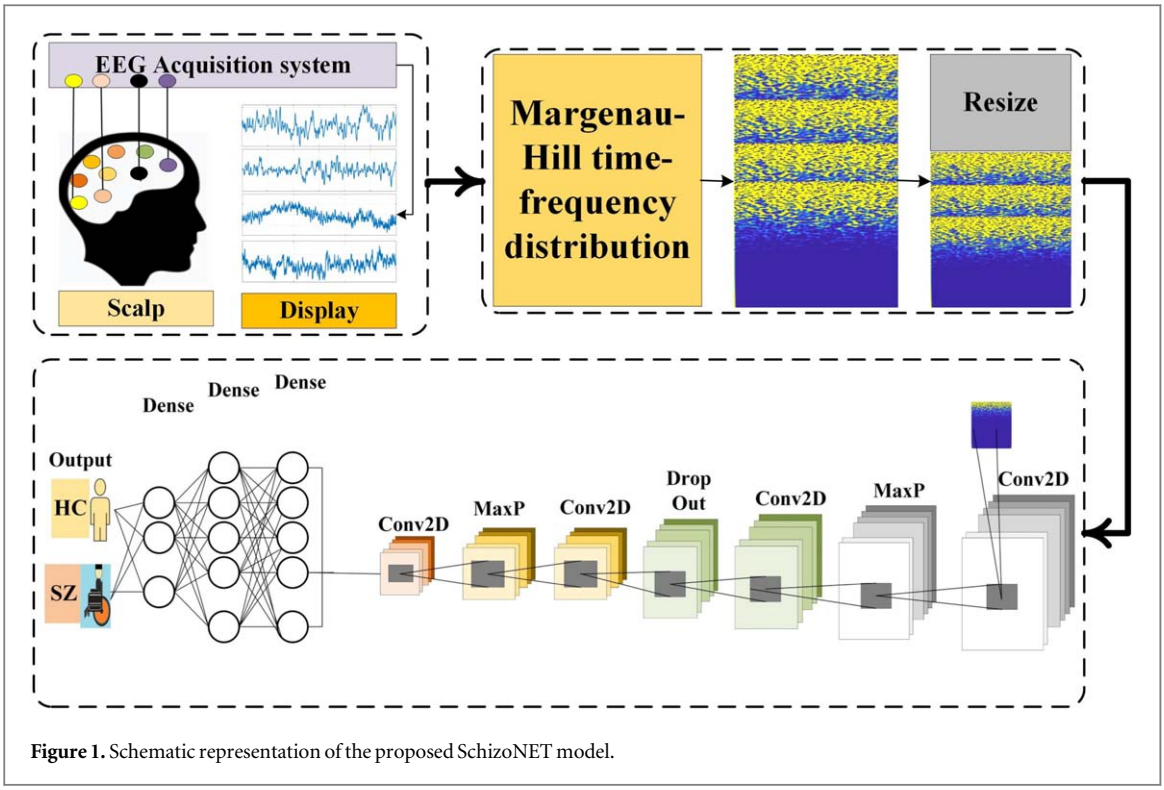

Figure 1. Schematic representation of the proposed SchizoNET model.

- Traditional techniques require extensive parameter tuning, the selection of handcrafted features is time-consuming, and the appropriate choice of classifiers is difficult. Therefore, we developed a simple CNN architecture using fewer layers and tested it with multiple validation techniques to evaluate performance metrics (PM).

The paper is structured as sections 1 and 2 presented the introduction and related work. Findings of literature and motivation are covered in section 3, and details about materials and methods are covered in section 4. Results are presented in section 5, performance comparison with current state-of-the-art is provided in section 6, a discussion is covered in section 7, and conclusions are given in section 8.

4. Materials and methods

The steps of the proposed SchizoNET model involve details of EEG datasets, the extraction of simultaneous temporal and spatial information using MH-TFD, automatic feature extraction and classification using the CNN model. The schematic of the SchizoNET is shown in figure 1.

4.1. Datasets

The proposed method uses three publicly available EEG datasets to test the SchizoNET model. The first dataset is acquired from adolescents during resting state, the second dataset is comprised of resting-state EEG acquired from adults, and the third dataset is recorded during the press button task. The demographic details of these datasets are discussed below and presented in table 3.

4.1.1. Dataset 1

The EEG dataset of Lomonosov Moscow State University has 84 adolescent subjects (Borisov *et al* 2005, Sergey and Gorbachevskaya). The resting-state eyes-closed EEG data is captured for 1 minute from 16 channels (T3, T4, T5, T6, F7, F3, F4, F8, P3, Pz, P4, C3, Cz, C4, O1, and O2 referenced to coupled ear electrodes). The SZ patients (including schizotypical, childhood SZ, and schizoaffective disorders) were diagnosed as per SZ diagnostic criteria F20, F21, and F25 of the International Classification of Diseases-10 (ICD-10) of Mental and Behavioural Disorders, set by the International Statistical Classification of Diseases and Related Health Problems. Specialists of the Mental Health Research Center have confirmed the diagnoses of SZ patients. Patients did not undergo any chemotherapy during the examination at the Mental Health Research Center. It is noted that only artefact-free EEG recordings are used for the analysis.

Table 3. Details of the datasets used.

Features	Values		
	Dataset 1	Dataset 2	Dataset 3
Total subjects	84	28	81
HC	39	14	32
SZ	45	14	49
Males (SZ)	—	7	26
Females (SZ)	—	7	6
Males (HC)	—	7	41
Females (HC)	—	7	8
Mean age (SZ)	12 years 3 months	28.1 ± 3.7 years	38.37 ± 13.91 years
Mean age (HC)	12 years 3 months	27.75 ± 3.15 years	40.02 ± 13.48 years
Mean age (Male SZ)	—	27.9 ± 3.3 years	40.21 ± 12.93 years
Mean age (Male HC)	—	26.8 ± 2.9 years	38.15 ± 12.97 years
Mean age (Female SZ)	—	28.3 ± 4.1 years	39 ± 16.98 years
Mean age (Female HC)	—	28.7 ± 3.4 years	39.33 ± 18.91 years
EEG segment	60 s	25 s	3 s
No. of segments	1344	21 702	493 824
No. of channels	16	19	64
Sampling Freq. (Hz)	128	250	1024

4.1.2. Dataset 2

Dataset 2 comprises 14 subjects with paranoid SZ hospitalized at the Institute of Psychiatry and Neurology in Warsaw, Poland, and 14 HC subjects (Olejarczyk and Jernajczyk 2017). All patients met the ICD-10 criteria for paranoid SZ (F20). The criteria for inclusion of SZ subjects: are ICD-10 diagnosis F20, a minimum medication washout period of seven days, and a minimum age of 18. Exclusion criteria: organic brain pathology, presence of a general medical condition, pregnancy, first episode of SZ, and neurological diseases. The EEG was recorded for 15 minutes in an eyes-closed resting-state condition. The 19-channel (P3, Fp1, Fp2, Pz, C4, F7, F3, Fz, P4, C3, Cz, F4, F8, O1, O2, T3, T4, T5, T6) EEG montage built in-accordance to international 10–20 system was used.

4.1.3. Dataset 3

It is obtained from Kaggle which contains EEG signals of 81 subjects (button–tone https://www.kaggle.com/broach/button-tone-sz n.d.). The diagnosis criteria for SZ patients were by the Structured Clinical Interview for DSM-IV. Subjects of both groups i.e. SZ and HC, were age, handedness (right), and gender-matched. Exclusion criteria for SZ included no dependence on substances for the past year while for HC subjects no history of substance dependence, current or past history of having a first-degree relative with a psychotic disorder, or DSM-IV Axis I disorder. The data were band-pass filtered between 0.5 and 15 Hz and a baseline was corrected at -0.6 to -0.5 s. The EEG epochs were artefact rejected for voltages exceeding $\pm 100\ \mu V$ at all scalp sites. The details about the dataset and acquisition steps can be found in Ford *et al* (2013). From the previous studies it has been found that pressing a button to generate a tone immediately is helpful in the detection of SZ and HC hence, it is used for analysis in the current work (Khare *et al* 2020, Siuly *et al* 2020, Smith *et al* 2021). The examples of EEG signals for HC and SZ subjects for Datasets 1, 2, and 3 are shown in figure 2.

4.2. Margenau–Hill time-frequency distribution (MH-TFD)

The information provided by signals about frequency-domain and time-domain components helps to study the characteristics of any signal. Since time-based representations use the entire frequency span in which the signal is defined, they ignore some hidden characteristics along with frequency. Similar limitations are also true to frequency-based representations (Advanced Time-Frequency Signal and System Analysis 2016). To address this, transformations based on TFD are the best way to represent a time-dependent spectrum of non-stationary EEG signals. The linear TFD like STFT and wavelets uses a window to localize behaviour in time and frequency. But to satisfy the Heisenberg-Gabor inequality, the resolution in time-frequency of this transformation is limited by localizing window parameters like duration and bandwidth. The choice of smaller time duration results in greater bandwidth and vice versa due to a compromise between time and frequency in linear TFD (Advanced Time-Frequency Signal and System Analysis 2016). The CWT-based TFR requires an appropriate selection of mother wavelet; which is again tedious. The MH-TFD helps to overcome the limitation of linear TFD as it does not use localizing windows or wavelet selection. MH-TFD uses autocorrelation of a signal rather than windows thus, it does not restrict resolutions in time frequency. MH-TFD provides better representation and decomposes EEG signal components into TFD. The time-frequency representation obtained using MH-TFD is denoted by

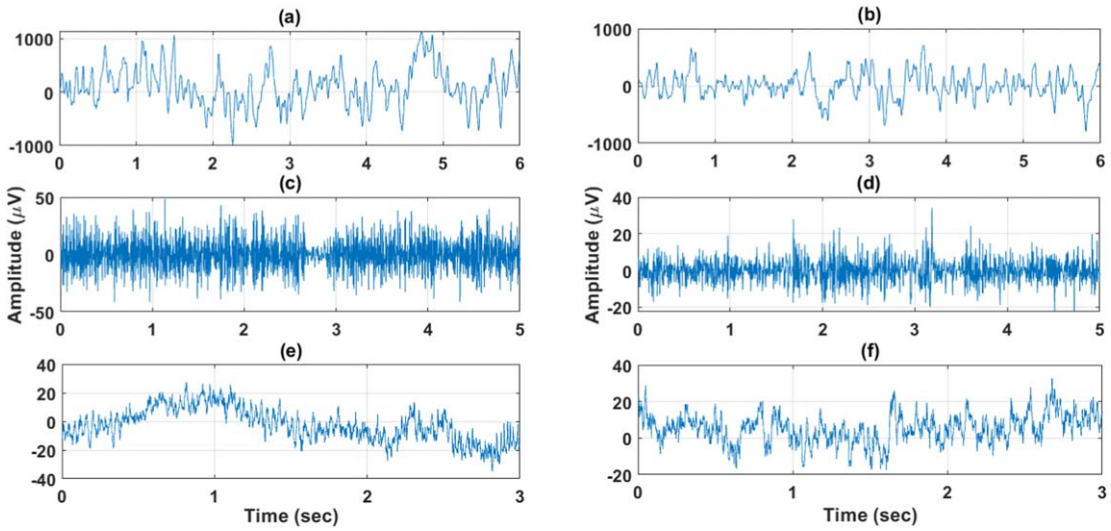

Figure 2. Typical EEG signals used: (a) SZ (dataset 1), (b) HC (dataset 1), (c) SZ (dataset 2), (d) HC (dataset 2), (e) SZ (dataset 3), and (f) HC (dataset 3).

equation (1) (Hatami *et al* 2016)

$$MH_y(t,f) = Real\left(\iiint \exp^{j\pi\tau\nu}\exp^{j2\pi\nu(s-t)}\right. \tag{1}$$

$$\left.y\left(s+\frac{\tau}{2}\right)y^*\left(s-\frac{\tau}{2}\right)\exp^{-j2\pi f\tau}\,dsd\tau d\nu\right) \tag{2}$$

where $y(.)$ denotes the signal to be analyzed, t and f represents time and frequency, $*$ denotes complex conjugate pair, and the kernel function is denoted by $exp^{j\pi\tau\nu}$. The above-mentioned equation can be simplified as (Hatami *et al* 2016)

$$MH_y(t,f) = Real(y(t)\exp^{-j2\pi ft}\,Y^*(f)), \tag{3}$$

where $Y(f)$ is the Fourier transform of $y(t)$. But, MH-TFD produces interference called cross-terms which interrupts the readability of the signal when analyzing it in multi-components. These cross-terms generate non-identical components that severely distort the signals. The cross-term formulation of a signal is denoted as (Hatami *et al* 2016)

$$CT_{y_i,y_k}(t,f) = Real\left(\iiint \exp^{j\pi\tau\nu}\exp^{j2\pi\nu(s-t)}\right. \tag{4}$$

$$\left.\left[y_i\left(s+\frac{\tau}{2}\right)y_k^*\left(s-\frac{\tau}{2}\right)+y_k\left(s+\frac{\tau}{2}\right)y_i^*\left(s-\frac{\tau}{2}\right)\right]\right. \tag{5}$$

$$\left.\exp^{-j2\pi f\tau}\,dsd\tau d\nu\right), \tag{6}$$

where the two-component of a signal is denoted by y_i and y_k with a cross-term CT_{y_i,y_k}. The cross-term in the time and frequency domain can be minimized by using a kernel function. MH-TFD uses time and frequency cross-term reduction kernels with a provision of flexible length to minimize the cross-term of a signal. Due to this reason, MH-TFD is a suitable choice for obtaining the time-frequency representation of a signal. The EEG segments of three datasets with all channels are converted to TFD image. For dataset 1, 60 s of EEG segment (7680 samples) are transformed into TFD image using MHTFD. Similarly, for datasets 2 and 3, we have used 25 s (6250 samples) and 3 s (3072 samples) of EEG segments are converted to TFD. Therefore, for datasets 1, 2, and 3 we have obtained 1344, 21 702, and 49 3824 TFD images, respectively. These TFD images of all the channels are fed to the CNN model for the detection of SZ from HC EEG segments. The typical TFD of an SZ and HC EEG signals obtained by MH-TFD on three datasets are shown in figure 3. The TFD indicates that the energy content of EEG for SZ and HC is dominant in a lower frequency range.

4.3. Convolutional neural network (CNN)

The 2D TFD obtained using MH-TFD is fed to a CNN model. It is an automated tool that enables the extraction and classification of deep features. Convolutional, pooling, dropout, dense, softmax, and classification layers are the main building blocks of CNN. The extraction of deep features is controlled by convolution, pooling, and dropout layers while the classification is done through dense, softmax, and output (classification) layers. The

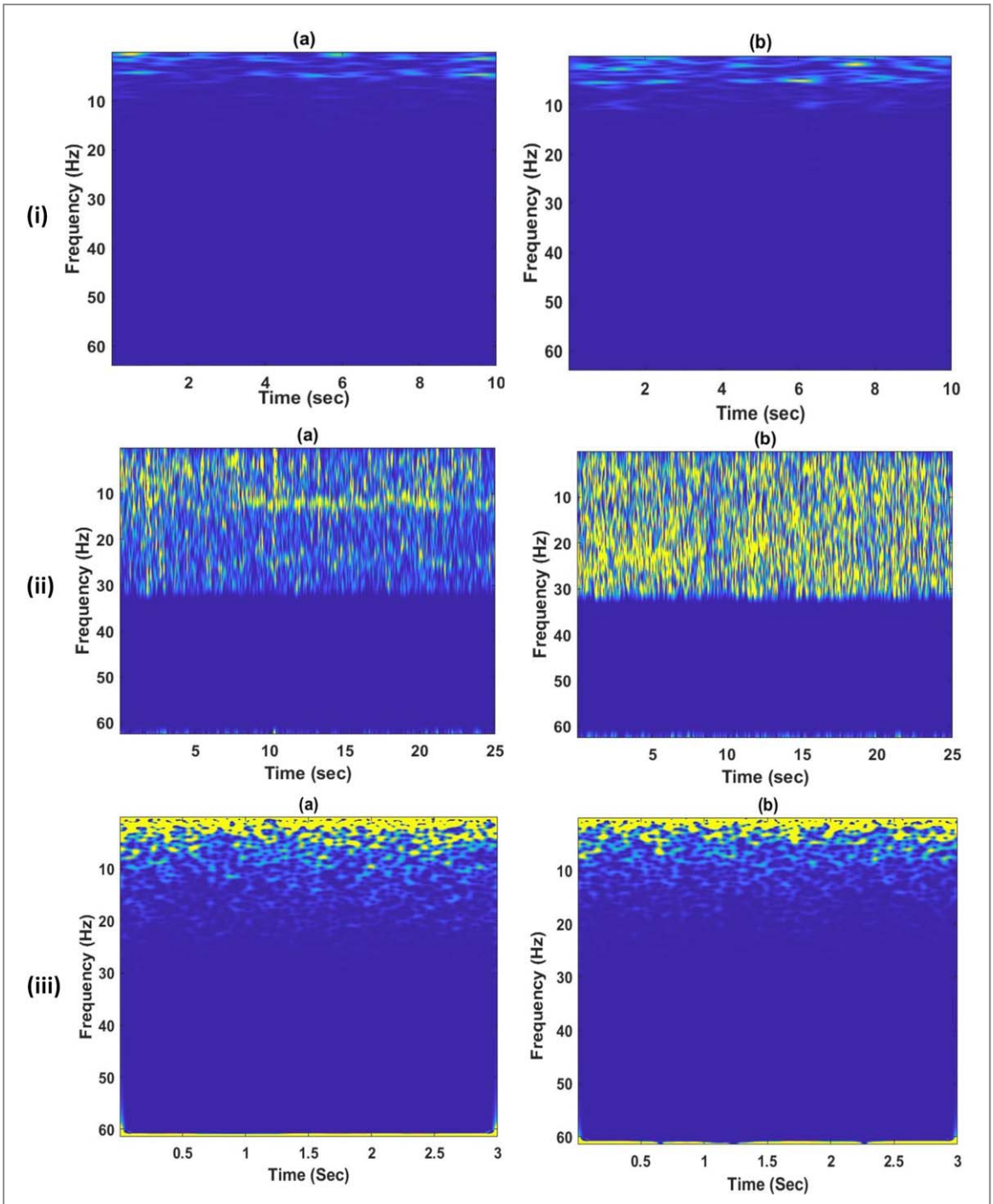

Figure 3. Typical TFD plots obtained: (a) SZ and (b) HZ; (i) Dataset 1, (ii) Dataset 2, and (iii) Dataset 3.

convolutional layer is the heart of the CNN model comprised of filters (kernels) that are moved along the tensor (image) in a fixed length called stride. Convolutions of kernel and tensor are evaluated to obtain output feature maps. Zero paddings are applied to keep image size while non-linearity in the network is added using the activation function. The pooling layer reduces the dimension of the output feature maps by keeping the number of input and output maps unaltered. A dense layer is followed by a pooling layer which transforms a 2D matrix to 1D and assigns some scores to the deep features extracted from the preceded layers. The softmax layer allocates the probability using some algorithms to each feature score. Finally, the classification layer assigns the output class to feature maps. In addition, a CNN model also uses a normalization layer to bring all the feature maps to the same scale which helps regularization, avoiding overfitting. A dropout layer deactivates some of the neurons in the network to lessen generalization error and overfitting.

Various CNN models are developed with different combinations of layers which vary from application to application and user to user. Some use CNN models with fever layers while others use dense CNN models composed of hundreds of layers (Alom *et al* 2018). Even with different configurations of CNN models and

Table 4. Summary of hyperparameters used for the proposed CNN model.

Layer	Type	KS	FS	OS	LP	OP
0	Input	—	—	—	—	—
1	Conv2D (Relu)	9 × 9	96	73 × 73	23 424	Str = 3
2	BatchNorm	—	—	73 × 73	384	—
3	MaxPool	—	—	36 × 36	0	Str = 2 PS = 3
4	Conv2D (Relu)	5 × 5	256	32 × 32	614 656	Str = 1
5	BatchNorm	—	—	32 × 32	1024	—
6	MaxPool	—	—	15 × 15	0	Str = 2 PS = 3
7	Conv2D (Relu)	3 × 3	384	13 × 13	885 120	Str = 1
8	Conv2D (Relu)	3 × 3	384	11 × 11	1327 488	Str = 1
9	Conv2D (Relu)	3 × 3	256	9 × 9	884 992	Str = 1
10	MaxPool	—	—	4 × 4	0	Str = 2 PS = 2
11	Dense (Relu)	—	—	4096	16 781 312	HN = 4096
12	DropOut	—	—	4096	0	Rate = 0.5
13	BatchNorm	—	—	4096	16 384	—
14	Dense (Relu)	—	—	4096	16 781 312	HN = 4096
15	DropOut	—	—	4096	0	Rate = 0.5
16	BatchNorm	—	—	4096	16 384	—
17	Dense (Relu)	—	—	192	786 624	HN = 192
18	DropOut	—	—	192	0	Rate = 0.5
19	BatchNorm	—	—	192	768	—
20	Dense (Softmax)	—	—	2	386	HN = 2

KS-kernel size, FS-filter size, OS-output size, LP-learning parameters, OP-other parameters, Str-stride, PS-pool size.

multiple trials, the desired performance is not achieved. Also, there is no standard model available for this application due to a lack of prior (Wolpert 1996). Therefore, a CNN model is developed with five convolutions, three pooling, three dense, and one output layer. In addition, the developed model uses rectified linear unit (Relu) as an activation function to increase non-linearity, a max-pooling layer to reduce the dimensionality of the feature map, a dropout of 50%, and batch normalization layers to reduce overfitting. The summary of the proposed CNN model is shown in table 4.

5. Results

Traditional machine learning techniques require extensive statistical analysis for selecting handcrafted methods and features. Moreover, a precise selection of the classifier and its parameters is time-consuming and does not guarantee performance success. Thus, the SchizoNET model is developed for the automatic detection of SZ. Three public datasets comprised of push-button tasks and resting-state EEG signals are employed for testing the SchizoNET model. The EEG epochs are transformed into time-frequency representation using MH-TFD. For reducing the cross-term of TFD, the Kaiser time and frequency window of lengths 31 and 63 are used. Obtained TFD is converted into images and fed to the CNN model. The learning rate is 10^{-04}, the bias and weight learning factor are both fixed at 10, the adaptive moment estimation optimizer is used to scale the learning rate of each weight, the batch size is 64, the total number of epochs is 60, and the frequency of validation is 50. All the parameters are selected empirically and maintained uniformly throughout the experimentation.

The DL models often offer very high performance; however, their stability is uncertain. Therefore, to verify the stability of our developed model, we have performed holdout (80% data used for training and 20% data used for testing), five-FCV, and ten-FCV techniques. The accuracy (ACC) obtained for each dataset using the

Table 5. Accuracy (%) obtained using various validation techniques with our proposed SchizoNET.

Validation technique	Dataset 1	Dataset 2	Dataset 3
Holdout	98.14	99.95	97.95
Five-FCV	97.47	99.9	97.44
Ten-FCV	97.4	99.74	96.35

Table 6. Performance measures obtained for SchizoNET using different validation techniques.

Validation	Holdout			Five-FCV			Ten-FCV		
Performance	Dataset: 1	Dataset: 2	Dataset: 3	Dataset: 1	Dataset: 2	Dataset: 3	Dataset: 1	Dataset: 2	Dataset: 3
ACC	98.14	99.95	97.95	97.47	99.89	97.44	97.40	99.74	96.35
SEN	97.93	99.96	97.58	97.25	99.87	97.02	96.85	99.64	95.95
SPE	98.39	99.95	98.53	97.73	99.92	98.10	98.04	99.87	96.97
Kappa	94.80	99.90	95.73	94.90	99.80	94.66	94.72	99.48	92.37
PRC	98.61	99.96	99.03	98.06	99.93	98.75	98.33	99.89	98.02
F-1	98.27	99.96	98.30	97.65	99.90	97.88	97.59	99.77	96.97

aforementioned validation techniques is shown in table 5. The results show that the ACC for holdout validation in all the datasets is highest because it is not averaged, while it is slightly reduced in the case of multi-fold CV techniques.

The SchizoNET model is evaluated by measuring six performance measures: ACC, Cohen's Kappa (Kappa), precision (PRC), sensitivity (SEN), specificity (SPE), and F-1 measure. As our model is tested on balanced and unbalanced datasets, we have chosen the above-mentioned performance measures. Table 6 shows the performance measures obtained using holdout, five-FCV, and ten-FCV with our SchizoNET model for Datasets 1, 2, and 3. The result shows that our developed SchizoNET model provides high performance on all three datasets and validation techniques. Thus, the results of tables 5 and 6 confirm the robustness of our SchizoNET model to obtain high performance in different validation scenarios for all three datasets. This confirms that our model has generated more distinct deep features to accurately categorize SZ and HC EEG segments. Figure 4 depicts plot of accuracy and loss versus iteration obtained for training, testing and validation phases of proposed SchizoNET.

A model with higher accuracy and higher variation from its mean value does not make a significance. As a result, the STD from a mean value of each PM is evaluated to measure the effectiveness of the SchizoNET. Table 7 provides the variation of STD and margin of error (MoE) obtained for each PM for a 95% confidence interval (CI). The table shows that the model got very little STD from the mean for each PM. On dataset 1, F-1 provides the lowest STD of ± 0.63 while the highest is ± 1.43 for Kappa. For dataset 2, SEN achieved the highest STD of ± 0.44 while PRC provided the minimum STD of ± 0.13. Finally, on dataset 3, the least STD of ± 0.68 is provided for F-1, whereas the highest is ± 2.46 for SPE. The MoE obtained using 95% CI on all datasets reveals that PM on each fold during ten-FCV shows no significant variance, indicating that the SchizoNET method is reliable and effective.

To get more insight into the SchizoNET model, the percentage confusion matrix is evaluated for three datasets using ten-FCV as shown in table 8. It can be evident from table 8 that 98.33% of SZ and 96.31% of HC EEG signals are identified correctly in their corresponding classes for Dataset 1. On Dataset 2, 99.89% of SZ and 99.56% of HC EEG signals, and on Dataset 3, 98.02% of SZ while 93.86% of HC EEG signals are correctly classified. The proposed SchizoNET model has the highest and least classification rate for SZ signals is 99.89% and 98.02%.

Further, we have evaluated receiver operating characteristics (ROC) and area under the curve (AUC) for our SchizoNET model, as shown in figure 5. It is evident that our developed model provided the AUC of 97.69%, 99.99%, and 96.52% for dataset 1, 2, and 3, respectively. This shows that our developed model accurately performs binary classification of SZ and HC.

6. Performance comparison

The performance of the SchizoNET is evaluated further by comparing it with current state-of-the-art techniques. Tables 9, 10, and 11 shows the performance comparison of the SchizoNET model on dataset 1, 2, and 3, respectively.

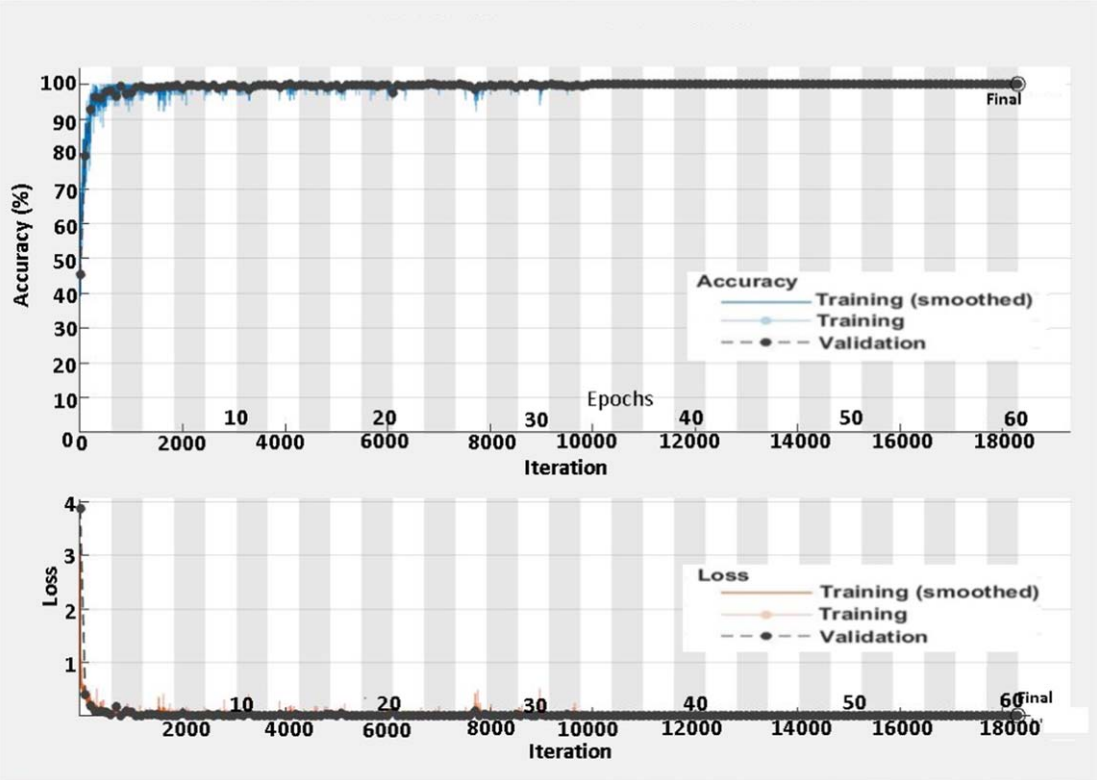

Figure 4. Plot of accuracy and loss versus iteration obtained for training, testing and validation phases of SchizoNET.

Table 7. STD from mean and MoE for 95% CI obtained for the SchizoNET model.

PM	Dataset 1		Dataset 2		Dataset 3	
	$\mu \pm STD$	MoE	$\mu \pm STD$	MoE	$\mu \pm STD$	MoE
ACC	97.40 ± 0.68	± 0.42	99.74 ± 0.22	± 0.14	96.35 ± 0.69	± 0.43
SEN	96.85 ± 1.11	± 0.68	99.64 ± 0.44	± 0.275	95.95 ± 1.55	± 0.96
SPE	98.04 ± 0.80	± 0.49	99.87 ± 0.15	± 0.09	96.97 ± 2.46	± 1.52
Kappa	94.72 ± 1.44	± 0.85	99.48 ± 0.40	± 0.28	92.37 ± 1.38	± 0.86
PRC	98.33 ± 0.70	± 0.43	99.89 ± 0.13	± 0.08	98.02 ± 2.27	± 1.41
F-1	97.59 ± 0.63	± 0.39	99.77 ± 0.20	± 0.12	96.97 ± 0.68	± 0.42

Table 8. Percentage confusion matrix obtained for SchizoNET model.

Dataset	Dataset 1		Dataset 2		Dataset 3	
Class	SZ	HC	SZ	HC	SZ	HC
SZ	708	12	11 882	12	289 076	5836
HC	23	601	43	9765	12 208	186 704

6.1. Dataset: 1

The model developed by Piryatinska *et al* (2017) operated on continuous vector functions obtained from ϵ-complexity to obtain the features classified using RF and SVM classifiers. Their model achieved the highest ACC of 89.3%, an average ACC of 85.3%, SEN of 88.6%, and SPE of 82.6% using the RF classifier. Calhas (2019) used DSTFT-based analysis of EEG signals to extract time-frequency-amplitude features. These features are given to CNN and different classifiers like KNN, SVM, XGBoost, RF, and Naive Bayes. The DSTFT and RF classifier combination has obtained the highest ACC of 84%, SPE of 82%, and SEN of 87%. Singh *et al* (2021) used filtering and FFT-based rhythm separation. Different features combining Hjorth parameters (mobility, complexity, and activity), spectral power, and mean spectral amplitude have been extracted from the delta, beta, gamma, alpha, and theta rhythms. These features are classified using deep learning classifiers like CNN and

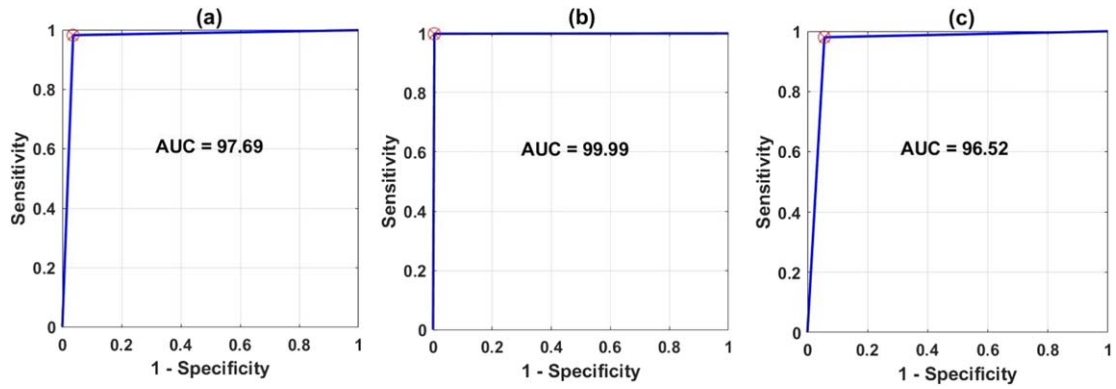

Figure 5. Summary of ROC and AUC curves obtained for SchizoNET: (i) Dataset 1, (ii) Dataset 2, and (iii) Dataset 3.

Table 9. Summary of performance (%) comparison with existing state-of-the-art methods developed using Dataset: 1.

Author	Feature extractor	Classifier (validation)	ACC	SEN	SPE	PRC	F-1
Piryatinska *et al* (2017)	ϵ-complexity	RF (10 FCV)	85.3	88.6	82.6	—	—
Calhas (2019)	DSTFT	RF (5 FCV)	84	87	82	—	—
Singh *et al* (2021)	FFT-based rhythms	CNN (Holdout)	94.08	92.7	95.31	—	93.62
Phang *et al* (2020)	TF-domain VAM	CNN (5 FCV)	91.69	91.11	92.5	94.14	—
Aslan and Akin (2020)	STFT	VGG-16 (Holdout)	95	95	—	95	95
Kutepov *et al* (2020)	Nonlinear features	Statistical analysis	—	—	—	—	—
Dimitriadis (2021)	FI and corrEnv	SVM (5 FCV)	100	100	100	—	—
Kumar *et al* (2023)	CBFS	AdaBoost (10 FCV)	92.85	93.3	92.3	—	—
SchizoNET	**MH-TFD**	**CNN (Holdout)**	**98.14**	**97.93**	**98.39**	**98.61**	**98.27**
		CNN (5 FCV)	**97.47**	**97.25**	**97.73**	**98.06**	**97.65**
		CNN (10 FCV)	**97.4**	**96.85**	**98.04**	**98.33**	**97.59**

Table 10. Summary of performance (%) comparison with existing state-of-the-art methods developed using Dataset: 2.

Author	Feature extractor	Classifier	ACC	SEN	SPE	PRC	F-1	Kappa
Oh *et al* (2019)	CNN	CNN (10 FCV)	98.07	97.32	98.17	98.45	—	—
Jahmunah *et al* (2019)	Nonlinear features	SVM-RBF (10 FCV)	92.91	93.45	92.22	93.6	—	—
Singh *et al* (2021)	FFT-based rhythms	CNN (Holdout)	98.96	99.05	98.88	—	98.87	—
Krishnan *et al* (2020)	Multivariate EMD	SVM-RBF (10 FCV)	93	94	92	92.71	93.04	—
Aslan and Akin (2020)	STFT	VGG-16 (Holdout)	97	97	—	97	97	—
Shalbaf *et al* (2020)	CWT	ResNet-SVM (10 FCV)	95.3	96.45	94.5	—	—	—
Buettner *et al* (2020)	FFT	RF (10 FCV)	96.77	96.77	96.77	96.77	96.77	93.55
Prabhakar *et al* (2020b)	EM-PCA, PLS, and Isomap	RF (Holdout)	98.77	—	—	—	—	—
Prabhakar *et al* (2020a)	Nonlinear features	SVM-RBF (10 FCV)	89.85	—	—	—	—	—
Sharma and Acharya (2020)	Optimal wavelet	KNN (10 FCV)	99.21	98.84	99.42	99.05	—	—
Nikhil *et al* (2021)	Nonlinear features	LSTM (Holdout)	99	98.9	—	99.2	99	—
Kumar *et al* (2023)	CBFS	AdaBoost (10 FCV)	99.36	99.2	99.4	—	—	—
Racz *et al* (2020)	DFC	RF (LOSO)	89.29	78.57	100	78.57	—	—
Aydemir *et al* (2022)	Complexity and HFD	KNN (10 FCV)	99.82	99.84	99.81	—	—	—
Goshvarpour and Goshvarpour (2020)	Nonlinear features	PNN (Holdout)	100	100	100	—	—	—
SchizoNET	**MH-TFD**	**CNN (Holdout)**	**99.95**	**99.96**	**99.95**	**99.96**	**99.96**	**99.9**
		CNN (5 FCV)	**99.89**	**99.87**	**99.92**	**99.93**	**99.90**	**99.8**
		CNN (10 FCV)	**99.74**	**99.64**	**99.87**	**99.89**	**99.77**	**99.48**

Table 11. Summary of performance (%) comparison with existing state-of-the-art methods developed using Dataset: 3.

Author	Feature extractor	Classifier (Validation)	ACC	SEN	SPE	PRC	F-1	Kappa
Zhang (2019)	Statistical features (2)	RF (10 FCV)	80.15	—	—	—	—	56.44
	Statistical features (5)	RF (10 FCV)	80.67	—	—	—	—	58.25
	Statistical features (8)	RF (10 FCV)	81.1	—	—	—	—	59.3
Siuly *et al* (2020)	EMD	SVM (10 FCV)	89.59	89.76	89.32	93.21	91.45	78.17
Khare *et al* (2020)	EWT	EBT (10 FCV)	88.7	91.13	89.29	83.78	—	—
Khare and Bajaj (2021b)	FTQWT	FLSSVM (10 FCV)	91.39	92.65	93.22	95.57	93.06	—-
Khare and Bajaj (2021a)	OVMD	OELM (10 FCV)	92.93	97.15	91.06	93.94	94.07	85.32
Siuly *et al* (2022)	Average filtering	GoogLeNet (Holdout)	95.09	93.81	97.02	97.95	95.83	—
Baygin *et al* (2023)	TQWT	KNN	95.84	97.01	94.06	96.11	—-	—-
Smith *et al* (2021)	SPWVD	CNN (10 FCV)	93.36	94.25	92.03	94.66	94.5	—
	STFT	CNN (10 FCV)	79.17	—	—	—	—	—
	CWT	CNN (10 FCV)	90.64	—	—	—	—	—
SchizoNET	**MH-TFD**	**CNN (Holdout)**	**97.95**	**97.58**	**98.53**	**99.03**	**98.3**	**95.73**
		CNN (5 FCV)	**97.44**	**97.02**	**98.1**	**98.75**	**97.88**	**94.66**
		CNN (10 FCV)	**96.35**	**95.95**	**96.97**	**98.02**	**96.97**	**92.37**

LSTM models. The spectral features classified using the CNN-based deep learning model have obtained the highest ACC of 94.08%, SEN, F-1 measure, and SPE of 92.7%, 93.62%, and 95.31%, respectively. Phang *et al* (2020) examined direct connectivity estimated from EEG to capture brain network during SZ. VAR model-based time-domain, PDC-based frequency-domain, and the network topology-based complex network measures for spatial features and their combination have been using the connectome CNN model. Their model achieved the highest ACC of 91.69%, SEN of 91.11%, SPE of 92.5%, and PRC of 94.14%. Aslan and Akin (2020) developed an automated model based on the spectrograms of EEG signals obtained using STFT and the visual geometry group (VGG-16) to classify deep features has obtained an ACC, SEN, PRC, and F-1 of 95%. Kutepov *et al* (2020) extracted nonlinear features and performed its statistical analysis to determine significant differences during SZ from HC subjects. Dimitriadis (2021) developed a hybrid model that analyzed relative power spectrum based on Welch's algorithm along with probability distribution of flexibility index (FI) and SVM classifier has achieved an ACC, SEN, and SPE of 100% using corrEnv features. Kumar *et al* (2023) used a histogram of local variance, symmetrically weighted local binary patterns, and correlation-based feature selection (CBFS) with an Adaboost classifier. Their model obtained 92.85%, 93.3%, and 92.3% of ACC, SEN, and SPE, respctively with thirteen features. The SchizoNET model developed using MH-TFD-based time-frequency representation with the CNN model has obtained the highest ACC, SEN, SPE, PRC, and F-1 using different validation techniques which is higher than all state-of-the-art models implemented on dataset 1.

6.2. Dataset: 2

Oh *et al* (2019) used a deep CNN model to automatically extract and classify has obtained an ACC, SEN, SPE, and PRC of 98.07%, 97.32%, 98.17%, and 98.45%. Jahmunah *et al* (2019) developed a nonlinear-based feature extractor combining entropy (Renyi, permutation, Tsallis, Kolmogorov-Sinai, and Shannon), activity, Hjorth and Kolmogorov complexity, mobility, largest Lyapunov exponent, and bispectrum to obtain recurrence plots. These features classified using an SVM classifier with radial basis function (RBF) kernel have achieved an ACC, SEN, SPE, and PRC of 92.91%, 93.45%, 92.22%, and 93.6%. The temporal, frequency and spectral feature-based model developed by Singh *et al* (2021) using FFT-based rhythms combined with the CNN model has achieved an ACC of 98.96%, SPE, F-1, and SEN of 98.88%, 98.87%, and 99.05%. Krishnan *et al* (2020) explored the utility of MEMD to obtain instantaneous amplitude and frequency-based mode functions. Several entropy-based features are extracted from these mode functions and selected using recursive feature selection. Their method has obtained the highest ACC of 93%, SEN of 94%, SPE of 92%, PRC and F-1 of 92.71%, and 93.04%C with all feature-set classified using radial basis function kernel of SVM classifier. Aslan and Akin (2020) developed an automated model based on the spectrograms of EEG signals obtained using STFT and the visual geometry group (VGG-16) to classify deep features has produced an ACC, SEN, PRC, and F-1 of 97%. Shalbaf *et al* (2020) explored the scalogram analysis using CWT to extract simultaneous time-frequency information from EEG signals. The scalogram is fed to a deep ResNet model to extract deep features. Using an SVM classifier, their hybrid model achieved an ACC, SEN, and SPE of 95.30%, 96.45%, and 94.50%. Buettner *et al* (2020) used analysis of spectra from different EEG rhythms extracted using FFT. These rhythms have been given to RF classifier and obtained an ACC, SEN, SPE, PRC, and F-1 of 96.77%, and Kappa of 93.55%. Prabhakar *et al* (2020b) extracted EM-PCA, PLS, and Isomap-based features from EEG signals. The statistically significant features have been selected using optimization and classified with different machine learning techniques. Their

model obtained an average ACC of 98.77%. In another method, Prabhakar *et al* (2020a) extracted Hurst exponent, largest Lyapunov exponent, Hjorth exponents, detrended fluctuation analysis, sample entropy, recurrence quantification analysis, fractal dimension, Kolmogorov, and Lampel Ziv complexity features. The features selected using black hole optimization have obtained a maximum ACC of 89.85% when classified with the SVM-RBF classifier. Sharma and Acharya (2020) developed an automated model using an optimal two-band orthogonal wavelet filter bank called optimal root-mean-squared frequency spread. Multiple L1-norm-based features have been extracted and classified using a KNN classifier to obtain 99.21%, 98.84%, 99.42%, and 99.05% of ACC, SEN, SPE, and PRC. Nikhil *et al* (2021) combined handcrafted Katz FD (KFD) and approximate entropy, along with the time-domain measure of variance values features an LSTM network. Their method has obtained an ACC and F-1 of 99%, PRC of 99.2%, and SEN of 98.9%, respectively. Kumar *et al* (2023) obtained the ACC, SEN, and SPE of 99.36%, 99.2%, and 99.4% using CBFS and AdaBoost classifier. Racz *et al* (2020) used DFC and RF classifier to extract and classify dynamic connectivity features. They obtained an ACC, SEN, and SPE of 89.29%, 78.57%, and 100%, respectively with leave one subject out (LOSO) validation. Aydemir *et al* (2022) used the analysis of the complexity and HFD features to classify it with the KNN classifier and reported an ACC, SEN, and SPE of 99.82%, 99.84%, and 99.81%, respectively. Goshvarpour and Goshvarpour (2020) used analysis of nonlinear features with PNN classifier to obtain a performance (ACC, SEN, and SPE) of 100%. Our developed SchizoNET model has obtained an ACC, SEN, SPE, PRC, F-1, and Kappa of 99.74%, 99.64%, 99.87%, 99.89%, 99.77%, and 99.48%, respectively which is higher than all using ten-FCV state-of-the-art techniques implemented on dataset 2.

6.3. Dataset: 3

Zhang (2019) used a combination of different statistical features to classify SZ and HC subjects using an RF classifier. The model developed has obtained the highest ACC and kappa of 81.81% and 59.3% using eight features. Khare *et al* (2020) used EWT-based analysis to extract various statistical features from the decomposed components. Among different classification techniques, the highest ACC, SEN, SPE, and PRC of 88.7%, 91.13%, 89.29%, and 83.78% have been obtained with the ensemble Bagged tree (EBT) classifier. Siuly *et al* (2020) developed an EMD-based automatic identification of SZ which provides instantaneous amplitude and frequency information about EEG signals. Different statistical measures obtained from the modes have been classified using an SVM classifier with an ACC of 89.59%, PRC of 83.78%, SEN of 91.13%, and SPE of 89.29%. Khare and Bajaj (2021b) used the Fisher score-based channel selection technique and FTQWT to extract the subbands of EEG signals. Five statistically significant features selected by the Kruskal Wallis test have been classified using the FLSSVM classifier. Their self-learned model achieved the performance values of ACC, SEN, SPE, PRC, and F-1 of 91.39%, 92.65%, 93.22%, 95.57%, and 93.06% respectively for the fourth subband. In another method, Khare and Bajaj (2021a) developed an optimized model combining OVMD and OELM classifier. The features of decomposed modes are selected using statistical analysis and classified using the OELM classifier. Their model has obtained the highest performance for SEN, SPE, Kappa, ACC, F-1, and PRC values of 97.15%, 91.06%, 85.32%, 92.93%, 94.07%, and 93.94% in the third mode due to chaotic behaviour. Smith *et al* (2021) extracted simultaneous information about time-frequency representations (TFR) from EEG signals using SPWVD, STFT, and CWT. The TFR has been transformed into images and fed to the CNN model. Their method has obtained an ACC of 93.36%, SEN of 94.25%, SPE of 92.03%, PRC of 94.66%, and F-1 of 94.5% using SPWVD-based TFR while an ACC of 79.17% and 90.64% has been achieved with spectrograms and scalograms techniques. Siuly *et al* (2022) used filtering and deep learning feature extraction and classification using GoogLeNet. They obtained an ACC, SEN, and SPE of 95.09%, 93.81%, and 97.02%, respectively with holdout validation using GoogLeNet classifier and 98.84%, **99.02%, and 98.58% using SVM classifier**. Baygin *et al* (2023) used iterative TQWT-based feature extractor combined with KNN classifier to detect SZ. They obtained an ACC, SEN, SPE, and PRC of 95.84%, 97.01%, 94.06%, and 96.11%, respectively with Cz channel. The proposed method combines MH-TFD and the CNN model has obtained an ACC, SEN, SPE, PRC, F-1, and Kappa of 96.35%, 95.95%, 96.97%, 98.02%, 96.97%, and 92.37% using the ten-FCV technique. Our developed model has obtained 97.95%, 97.58%, and 98.53% of ACC, SEN, and SPE, respectively with holdout validation. The performance of the SchizoNET model is higher than all state-of-the-art with different validation techniques developed on dataset 3 confirming that MH-TFD can capture the subtle details from the EEG signals.

7. Discussion

The methods developed using nonlinear features require tuning of multiple parameters and showing degraded performance due to different noise and artefacts. FFT-based techniques analyze the EEG signals in the time-frequency domain. Still, time-based representations of a signal use the entire frequency span over which it is defined and may ignore some hidden characteristics along with frequency. Similar drawbacks are also true for

frequency-based representations resulting in poor time-frequency localization. The analysis of EEG signals using STFT assumes the signal to be stationary over a duration, and also requires the selection of length and type of window. The discrete wavelet transform-based techniques decompose a signal into subbands by selecting a mother wavelet and decomposition level that are difficult to find. The EMD-based decomposition extracts instantaneous information regarding amplitude and frequency but lacks mathematical modelling and suffers from mode mixing. The TFD provided by STFT and CWT requires to satisfy Heisenberg-Gabor inequality, due to which the resolution in time-frequency of this transformation is limited by localizing window parameters like duration and bandwidth. The TFD obtained using SPWVD requires kernel function and its length as parameters. Improper choice of these parameters may produce severe distortion in the TFR. These shortcomings result in overlapping information about SZ and HC EEG signals that degrade the system performance. Some CNN-based classification models use single-fold fixed-length training and testing sets that might produce an overfitted model. Moreover, many studies are limited to a single EEG data evaluation which does not guarantee similar performance on other or different EEG datasets of the same problem. Our proposed SchizoNET model combines MH-TFD and CNN for automatic time-frequency-amplitude feature extraction and classification. The MH-TFD does not require any choice of a window but uses the autocorrelation of signals to be analyzed. In addition, the problem of cross-term is overcome due to the use of cross-term reduction window in frequency and time domain. This enables the extraction of more hidden information from EEG signals which reflect representative and distinguishable characteristics of it. The CNN model enables automatic classification to detect SZ and HC EEG signals. The evaluation of SchizoNET on three different EEG datasets of SZ developed using holdout, five-FCV, and ten-FCV has provided the highest performance over existing state-of-the-art techniques on most of the datasets. It is evident from tables 9, 10, and 11 that this is the first study to develop a novel DL model which can be used for all three public datasets and yield the highest performance. In addition, our developed model is simple as it has only five convolutional layers compared to benchmark CNN models like AlexNet, VGG-16, and ResNet-50 (Smith *et al* 2021). Also, our SchizoNET model requires fewer learning parameters i.e. about 42.8 million compared to existing AlexNet (approx. 61 million) and VGG-16 (approx. 138 million) parameters (Smith *et al* 2021). The merits of our developed SchizoNET model are as follows:

- **Robust**: the SchizoNET model is robust because it is developed using three different EEG datasets.
- **Accurate and stable**: the developed model reported the highest and most consistent performance with holdout and cross-validation techniques.
- **Simple and effective**: our model is simple (only five convolutional layers) and has fewer learning parameters than benchmark CNN models.

The limitations of our proposed model are given below:

- In our used three datasets, the number of subjects used in each dataset is too few to explore the LOSO validation technique.
- Our work does not localize the region of SZ.

8. Conclusion

The proposed SchizoNET model combines MH-TFD and CNN to automatically detect SZ patients using EEG signals. The TFD generated by MH-TFD has provided excellent resolution, hidden information, and detailed insight into EEG signals due to a reduction in cross-term. The TFD has facilitated the CNN model to extract deep features that drastically reduced manual efforts. The simple architecture of the proposed CNN model has drastically improved the system performance with fewer learnable parameters. Our developed model correctly identified 99.74% of SZ signals, the highest among current state-of-the-art techniques. Thus, the proposed SchizoNET model is robust, effective, accurate, and versatile as it obtained the highest performance matrices on three EEG datasets. Also, the designed model can detect SZ in a resting state, evoked potential, and tasks related to EEG acquisition. Our SZ model is more generalized as it does not require any feature engineering and can automatically extract and classify features. The limitation of our model is that it has not been developed using LOSO cross-validation due to the fewer subjects in each of the three datasets. In the future, we will develop a subject-based and channel-wise SZ detection model by using more subjects in each class.

Data availability statement

No new data were created or analysed in this study. Data will be available from 31 January 2028.

ORCID iDs

Smith K Khare https://orcid.org/0000-0001-8365-1092
Varun Bajaj https://orcid.org/0000-0002-8721-1219
U Rajendra Acharya https://orcid.org/0000-0003-2689-8552

References

Advanced Time-Frequency Signal and System Analysis 2016 Advanced time-frequency signal and system analysis *Time-Frequency Signal Analysis and Processing* ed B Boashash 2nd edn (Oxford: Academic Press) ch 4 pp 141–236
Akar S A *et al* 2016 Analysis of the complexity measures in the EEG of schizophrenia patients *Int. J. Neural Syst.* **26** 1650008
Alimardani F *et al* 2018 Classification of bipolar disorder and schizophrenia using steady-state visual evoked potential based features *IEEE Access* **6** 40379–88
Alom M Z *et al* 2018 The history began from AlexNet: a comprehensive survey on deep learning approaches arXiv:1803.01164
Aslan Z and Akin M 2020 Automatic detection of schizophrenia by applying deep learning over spectrogram images of EEG signals *Trait. Signal* **37** 235–44
Aydemir E *et al* 2022 CGP17Pat: automated schizophrenia detection based on a cyclic group of prime order patterns using EEG signals *Healthcare* **10** 643
Baygin M 2021 An accurate automated schizophrenia detection using TQWT and statistical moment based feature extraction *Biomed. Signal Process. Control* **68** 102777
Baygin M *et al* 2023 CCPNet136: automated detection of schizophrenia using carbon chain pattern and iterative TQWT technique with EEG signals *Physiol. Meas.* (https://doi.org/10.1088/1361-6579/acb03c)
Begić D, Hotujac L and Jokić-Begić N 2000 Quantitative EEG in positive and negative schizophrenia *Acta Psychiatrica Scandinavica* **101** 307–11
Sergey B and Gorbachevskaya N *EEG of Healthy Adolescents and Adolescents with Symptoms of Schizophrenia* (http://brain.bio.msu.ru/eeg_schizophrenia.htm)
Borisov S *et al* 2005 Analysis of EEG structural synchrony in adolescents with schizophrenic disorders *Hum. Physiol.* **31** 255–61
Bromet E J and Fennig S 1999 Epidemiology and natural history of schizophrenia *Biol. Psychiatry* **46** 871–81
Buettner R *et al* 2020 *Development of a Machine Learning Based Algorithm To Accurately Detect Schizophrenia based on One-minute EEG Recordings* pp 3216–25
Caldwell C and Gottesman I 1990 Schizophrenics kill themselves too: a review of risk factors for suicide *Schizophrenia Bull.* **16** 571–89
Calhas D 2019 Schizophrenia diagnosis using electroencephalographic data *PhD Thesis* Spain: information and Computer Engineering, Universitat Politécnica de Catalunya
Cho K-O and Jang H-J 2020 Comparison of different input modalities and network structures for deep learning-based seizure detection *Sci. Rep.* **10** 122
Ciprian C *et al* 2021 Diagnosing schizophrenia using effective connectivity of resting-state EEG data *Algorithms* **14** 139
Dimitriadis S I 2021 Reconfiguration of amplitude driven dominant coupling modes (DoCM) mediated by alpha-band in adolescents with schizophrenia spectrum disorders *Progress Neuro-Psychopharmacol. Biol. Psychiatry* **108** 110073
Dvey-Aharon Z *et al* 2015 Schizophrenia detection and classification by advanced analysis of eeg recordings using a single electrode approach *PLoS One* **10** 1–12
Ford J M *et al* 2013 Did i do that? abnormal predictive processes in schizophrenia when button pressing to deliver a tone *Schizophrenia Bull.* **40** 804–12
Goshvarpour A and Goshvarpour A 2020 Schizophrenia diagnosis using innovative EEG feature-level fusion schemes *Phys. Eng. Sci. Med.* **43** 227–38
Green M F, Horan W P and Lee J 2015 Social cognition in schizophrenia *Nat. Rev. Neurosci.* **16** 620–31
Hatami M, Farrokhifard M and Parniani M 2016 A non-stationary analysis of low-frequency electromechanical oscillations based on a refined Margenau–Hill distribution *IEEE Trans. Power Syst.* **31** 1567–78
Hettige N *et al* 2017 A biopsychosocial evaluation of the risk for suicide in schizophrenia *CNS Spectrums* **23** 1–11
Hiesh M *et al* 2013 Classification of schizophrenia using genetic algorithm- support vector machine (GA-SVM) *2013 35th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)* pp 6047–50
https://kaggle.com/broach/button-tone-sz (n.d.). Access 22 Jun 2020
Jahmunah V *et al* 2019 Automated detection of schizophrenia using nonlinear signal processing methods *Artif. Intell. Med.* **100** 101698
Jin H and Mosweu I 2017 The societal cost of schizophrenia: a systematic review *Pharmacoeconomics* **35** 25–42
Khare S K and Bajaj V 2021a A hybrid decision support system for automatic detection of Schizophrenia using EEG signals *Comput. Biol. Med.* **141** 105028
Khare S K and Bajaj V 2021b A self-learned decomposition and classification model for schizophrenia diagnosis *Comput. Methods Programs Biomed.* **211** 106450
Khare S K and Bajaj V 2021c Time-frequency representation and convolutional neural network- based emotion recognition *IEEE Trans Neural Netw. Learn. Syst.* **32** 2901–9
Khare S K, Gaikwad N and Bokde N D 2022 An intelligent motor imagery detection system using electroencephalography with adaptive wavelets *Sensors* **22** 8128
Khare S K, Gaikwad N B and Bajaj V 2022 VHERS: a novel variational mode decomposition and hilbert transform-based EEG rhythm separation for automatic ADHD detection *IEEE Trans. Instrum. Meas.* **71** 1–10
Khare S K *et al* 2020 Classification of schizophrenia patients through empirical wavelet transformation using electroencephalogram signals *Modell. Anal. Active Biopotential Signals Healthcare* **1** 2053–563

Krishnan P T et al 2020 Schizophrenia detection using multivariate empirical mode decomposition and entropy measures from multichannel EEG signal *Biocybernetics Biomed. Eng.* **40** 1124–39
Kumar J S and Bhuvaneswari P 2012 Analysis of electroencephalography (EEG) signals and its categorization—a study. Procedia Engineering 38 *Int. Conf. on Modelling Optimization and Computing* pp 2525–36
Kumar T S et al 2023 Automated Schizophrenia detection using local descriptors with EEG signals *Eng. Appl. Artif. Intell.* **117** 105602
Kutepov I E et al 2020 EEG analysis in patients with schizophrenia based on Lyapunov exponents *Inf. Med. Unlocked* **18** 100289
Lai J W et al 2021 Schizophrenia: a survey of artificial intelligence techniques applied to detection and classification *Int. J. Environ. Res. Public Health* **18** 1660–4601
Li F et al 2019 Differentiation of schizophrenia by combining the spatial eeg brain network patterns of rest and task P300 *IEEE Trans. Neural Syst. Rehabil. Eng.* **27** 594–602
Liu H et al 2017 A data driven approach for resting-state EEG signal *Classification of Schizophrenia with Control Participants using Random Matrix Theory* arXiv:1712.05289
Lloyd J et al 2017 Treatment outcomes in schizophrenia: qualitative study of the views of family carers *BMC Psychiatry* **17** 266
Masychev K et al 2021 Advanced signal processing methods for characterization of schizophrenia *IEEE Trans. Biomed. Eng.* **68** 1123–30
Namazi H, Aghasian E and Ala T S 2019 Fractal-based classification of electroencephalography (EEG) signals in healthy adolescents and adolescents with symptoms of schizophrenia *Technol. Health Care* **27** 233–41
Nikhil C, Sreekumar K and Subha D P 2021 EEG-based automated detection of schizophrenia using long short-term memory (LSTM) network *Advances in Machine Learning and Computational Intelligence* ed S Patnaik, X-S Yang and I K Sethi (Singapore: Springer) pp 229–36
Oh S L et al 2019 Deep convolutional neural network model for automated diagnosis of schizophrenia using EEG signals *Appl. Sci.* **9** 2870
Olejarczyk E and Jernajczyk W 2017 Graph-based analysis of brain connectivity in schizophrenia *PLoS One* **12** 1–28
Parvinnia E et al 2014 Classification of EEG Signals using adaptive weighted distance nearest neighbor algorithm *J. King Saud Univ.—Comput. Inf. Sci.* **26** 1–6
Phang C-R et al 2020 A multi-domain connectome convolutional neural network for identifying schizophrenia from eeg connectivity patterns *IEEE J. Biomed. Health Inform.* **24** 1333–43
Phang C et al 2019 Classification of EEG-based effective brain connectivity in schizophrenia using deep neural networks *2019 9th Int. IEEE/EMBS Conf. on Neural Engineering (NER)* pp 401–6
Piryatinska A, Darkhovsky B and Kaplan A 2017 Binary classification of multichannel-EEG records based on the ϵ-complexity of continuous vector functions *Comput. Methods Programs Biomed.* **152** 131–9
Prabhakar S K, Rajaguru H and Lee S-H 2020b A framework for schizophrenia eeg signal classification with nature inspired optimization algorithms *IEEE Access* **8** 39875–97
Prabhakar S K, Rajaguru H and Lee S-W 2020a Schizophrenia EEG signal classification based on swarm intelligence computing *Comput. Intell. Neurosci.* **2020** 1–14
Racz F S et al 2020 Multifractal and entropy-based analysis of delta band neural activity reveals altered functional connectivity dynamics in schizophrenia *Front. Syst. Neurosci* **14**
Ravan M et al 2015 A machine learning approach using auditory oddball responses to investigate the effect of Clozapine therapy *Clin. Neurophysiol.* **126** 721–30
Sabeti M, Katebi S and Boostani R 2009 Entropy and complexity measures for EEG signal classification of schizophrenic and control participants *Artif. Intell. Med.* **47** 263–74
Sadeghi D et al 2022 An overview of artificial intelligence techniques for diagnosis of Schizophrenia based on magnetic resonance imaging modalities: Methods, challenges, and future works *Comput. Biol. Med.* **146** 105554
Shalbaf A, Bagherzadeh S and Maghsoudi A 2020 Transfer learning with deep convolutional neural network for automated detection of schizophrenia from EEG signals *Australas. Phys. Eng. Sci. Med.* **43** 1229–39
Sharma M and Acharya U R 2020 Automated detection of schizophrenia using optimal wavelet-based l1 norm features extracted from singlechannel EEG *Cognit. Neurodynamics* **15** 661–74
Singh K, Singh S and Malhotra J 2021 Spectral features based convolutional neural network for accurate and prompt identification of schizophrenic patients *Proc. Inst. Mech. Eng.* H **235** 167–84
Siuly S et al 2020 A computerized method for automatic detection of schizophrenia using eeg signals *IEEE Trans. Neural Syst. Rehabil. Eng.* **28** 2390–400
Siuly S et al 2022 SchizoGoogLeNet: the googlenet-based deep feature extraction design for automatic detection of schizophrenia *Comput. Intell. Neurosci.* **2022** 1–13
Smith K K, Bajaj V and Acharya U R 2021 SPWVD-CNN for automated detection of schizophrenia patients using eeg signals *IEEE Trans. Instrum. Meas.* **70** 1–9
Sui J et al 2014 Combination of FMRI-SMRI-EEG data improves discrimination of schizophrenia patients by ensemble feature selection *2014 36th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society* pp 3889–92
Sun J et al 2021 A hybrid deep neural network for classification of schizophrenia using EEG data *Sci. Rep.* **11** 4706
Talo M et al 2019 Application of deep transfer learning for automated brain abnormality classification using MR images *Cognit. Syst. Res.* **54** 176–88
Vareka L 2021 Comparison of convolutional and recurrent neural networks for the P300 detection *BIOSIGNALS* 186–91
Wolpert D H 1996 The lack of a priori distinctions between learning algorithms *Neural Comput.* **8** 1341–90
WHO (2022). In.
Yin Z et al 2017 Functional brain network analysis of schizophrenic patients with positive and negative syndrome based on mutual information of EEG time series *Biomed. Signal Process. Control* **31** 331–8
Zhang L 2019 EEG signals classification using machine learning for the identification and diagnosis of schizophrenia *2019 41st Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)* pp 4521–4