



Detection of schizophrenia using hybrid of deep learning and brain effective connectivity image from electroencephalogram signal

Sara Bagherzadeh^a, Mohsen Sadat Shahabi^b, Ahmad Shalbaf^{b,*}

^a Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

^b Department of Biomedical Engineering and Medical Physics, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran



ARTICLE INFO

Keywords:

Electroencephalogram (EEG)
Schizophrenia (SZ)
Effective connectivity
Transfer entropy (TE)
Convolutional neural network (CNN)
Long short-term memory (LSTM)

ABSTRACT

Detection of mental disorders such as schizophrenia (SZ) through investigating brain activities recorded via Electroencephalogram (EEG) signals is a promising field in neuroscience. This study presents a hybrid brain effective connectivity and deep learning framework for SZ detection on multichannel EEG signals. First, the effective connectivity matrix is measured based on the Transfer Entropy (TE) method that estimates directed causalities in terms of brain information flow from 19 EEG channels for each subject. Then, TE effective connectivity elements were represented by colors and formed a 19×19 connectivity image which, simultaneously, represents the time and spatial information of EEG signals. Created images are used to be fed into the five pre-trained Convolutional Neural Networks (CNN) models named VGG-16, ResNet50V2, InceptionV3, EfficientNetB0, and DenseNet121 as Transfer Learning (TL) models. Finally, deep features from these TL models equipped with the Long Short-Term Memory (LSTM) model for the extraction of most discriminative spatio-temporal features are used to classify 14 SZ patients from 14 healthy controls. Results show that the hybrid framework of pre-trained CNN-LSTM models achieved higher accuracy than pre-trained CNN models. The highest average accuracy and F1-score were achieved using the EfficientNetB0-LSTM model through the 10-fold cross-validation method equal to 99.90% and 99.93%, respectively. Therefore, the superior performance of the hybrid framework of brain effective connectivity images from EEG signals and pre-trained CNN-LSTM models show that the proposed method is highly capable of detecting SZ patients from healthy controls.

1. Introduction

Schizophrenia (SZ) is a psychiatric disorder that severely influences perception and individual relations [1,2]. These patients lose trust in others and this causes damage to relationships with family or other people in ordinary situations or workplaces [3,4]. Schizophrenia is related to the changes in structure and function of the cortex, and subsequently connections of different cortical regions [5,6]. There is a reduction in the gray matter volume of SZ patients or disruption in the integrity of white matter [5–7]. Moreover, there are abnormalities in the neural activity of SZ patients that affects cognition [8], emotion, and memory tasks [6]. Finally, researchers had identified different gene expressions of SZ and healthy individuals [9,10]. About 20 million people throughout the world suffer from SZ disease [11]. These patients can be treated through an accurate prognosis. Therefore, designing a semi- or fully automated framework to diagnose SZ patients from brain imaging modalities like Electroencephalogram (EEG) signals using

advanced machine learning algorithms can be very useful in clinical procedures. EEG is a common tool to evaluate or diagnose psychiatric disorders [12,13] and has several advantages compared to other brain mapping techniques because of its high temporal resolution, non-invasiveness, portability, and availability in most clinical centers.

Multiple studies tried to extract features from EEG signals using various methods to detect SZ patients such as nonlinear analysis including different types of entropies [14–18], largest Lyapunov exponent [14,19], Hjorth complexity [20,21], spectral analysis like Fourier transform [20,22,23], time-frequency analysis like Short-Time Fourier Transform (STFT) [24], Continuous Wavelet Transform (CWT) [25–27] or smoothed pseudo-Wigner–Ville distribution [25], structural features like collatz pattern [28] and graphical features [29]. Other studies applied deep learning techniques i.e., Convolutional Neural Networks (CNNs) to extract high-level features and detect SZ from EEG signals [30–32]. Oh et al. [30] developed an end-to-end approach using CNN models, and Phang et al. [31] utilized CNN with a multi-domain

* Corresponding author.

E-mail address: shalbaf@sbmu.ac.ir (A. Shalbaf).

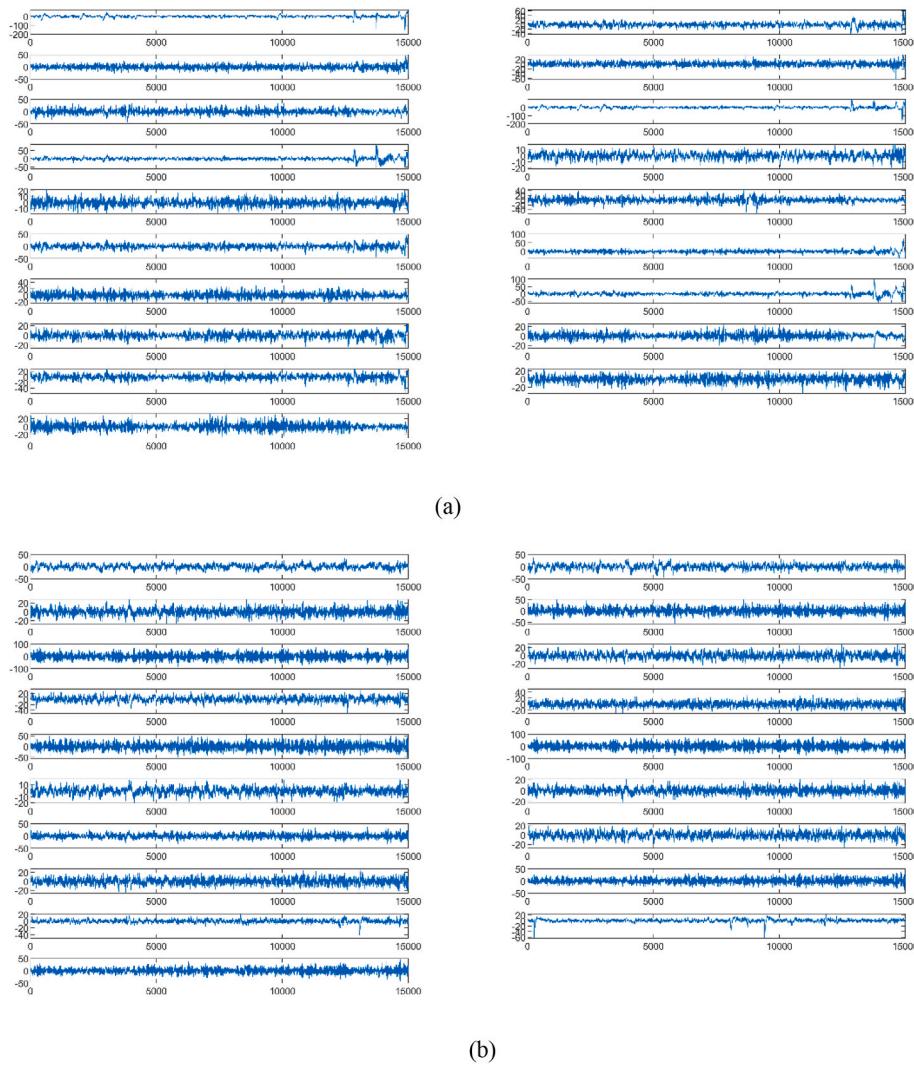


Fig. 1. A sample from 19 EEG channels for (a) SZ patient and (b) healthy participant during 1 min of recording.

connectome approach to diagnose SZ patients. CNN models are powerful image processing tools and the deep structure of these networks and their ability to learn abstracted patterns results in high performance [34, 35]. Due to these advantages, they have become a widespread tool in several fields such as computer-assisted medical diagnoses [25–32, 34–40]. Another popular method among deep learning algorithms is Long Short-Term Memory (LSTM) method which takes the time dependency into account. On the other hand, this model explores and learns the sequential order of input samples. LSTM models have been used in many EEG processing studies due to their effectiveness, high performance, and innovative architecture [41–43].

The main contribution of this paper is divided into two issues. First, represent EEG signals of 19 channels with a three-dimensional spatial-time color image using the popular brain effective connectivity measure based on information theory named Transfer Entropy (TE) to be fed into the deep learning method. Brain effective connectivity investigates a specific relationship between pairs of multiple EEG channels. Then, design a hybrid framework based on created effective connectivity images, pre-trained CNN models, and LSTM model to improve the detection of SZ from healthy controls.

2. Material and methods

2.1. Participant and EEG recording

EEG signals from the publicly available database were used in this study [44]. We downloaded and used this database “EEG in schizophrenia” from the link: <https://repod.icm.edu.pl/dataset.xhtml?persistentId=doi:10.18150/repod.0107441>. This database was collected from 14 patients suffering from SZ and 14 healthy subjects with the same range of age and equal sex ratio. EEG signals are recorded for 12 min in which subjects were relaxed and their eyes were closed. There are 19 EEG channels recorded according to the international 10–20 electrode placement system from each participant (Fp1, Fp2, F7, F3, Fz, F4, F8, C3, Cz, C4, P3, Pz, P4, T3, T4, T5, T6, O1, O2) with the sampling frequency of 250 Hz. All subjects meet the International Classification of Diseases (ICD)-10 criteria for paranoid SZ and the recording protocol was approved by the Ethics Committee of the Institute of Psychiatry and Neurology in Warsaw, Poland. Individuals from both groups of SZ and healthy must be higher than 18 years old. SZ patients were diagnosed based on the ICD-10 with F20.0 (paranoid schizophrenia) and patients at a very early stage of SZ were not permitted to participate. In other words, these patients suffer from one specific type of SZ named paranoid schizophrenia. Also, they must not take any type of medicine till seven days before the experiment. Pregnant women, patients with brain pathology and neurological disorders such as epilepsy, Parkinson’s, or

Alzheimer's disease were not permitted to participate in the experiment. EEG signals were passed through a low pass and high pass Finite Impulse Response (FIR) Butterworth filters with the cutting of frequencies of 0.5 and 45 Hz, respectively, in the EEGLAB toolbox [45] on MATLAB software (version 2020a). Fig. 1 shows an example of 1 min EEG signal from an SZ patient (a) and a normal participant (b).

2.2. TE connectivity measure

TE is a model-free, non-parametric and nonlinear effective connectivity measure which denotes causality between two time-series based on the conditional entropy [46]. TE is a popular estimator of brain effective connectivity in the neuroscience area that has been used in cognition problems [47–49], assessing the drowsiness [50], anesthesia [51], and recognition of emotional states [52]. Given the time series of x (t) and $y(t)$ from a Markov process, Niso et al. [53] proposed a measure based on causality which could calculate the generalized Markov condition from Eq. (1):

$$p(y_{t+1}|y_t^n, x_t^m) = p(y_{t+1}|y_t^n) \quad (1)$$

Where $x_t^m = (x_t, x_{t+1}, \dots, x_{t-m+1})$ and $y_t^n = (y_t, y_{t+1}, \dots, y_{t-n+1})$ and m and n are memory of Markov processes in x and y , respectively. The right side of this equation denotes the probability of observing a value of y_{t+1} supposing the previous n steps are available and the left side calculates this probability given both histories of $x(t)$ and $y(t)$. Eq. (1) is fully satisfied when the dynamics or transition probabilities of y are independent of the past of x (in the absence of causality from x to y). TE from x to y using the Kullback-Leibler divergence between two probability distributions is determined as $E(q)$: 2:

$$T_{X \rightarrow Y} = \sum_{y_{t+1}, y_t^n, x_t^m} p(y_{t+1}|y_t^n, x_t^m) \log \left(\frac{p(y_{t+1}|y_t^n, x_t^m)}{p(y_{t+1}|y_t^n)} \right) \quad (2)$$

This equation calculates directed information flow from x to y . Therefore, we can write the TE relation according to Eq. (2) from time series x_t to y_t as Eq. (3):

$$T_{X \rightarrow Y} = \sum_{y_{t+1}, y_t^{d_y}, x_t^{d_x}} p(y_{t+1}|y_t^{d_y}, x_t^{d_x}) \log \left(\frac{p(y_{t+1}|y_t^{d_y}, x_t^{d_x})}{p(y_{t+1}|y_t^{d_y})} \right) \quad (3)$$

Where t and u are time-index and prediction time, respectively, and both are discrete values. $y_t^{d_y}$ and $x_t^{d_x}$ denote d_y - and d_x -dimensional delay vectors and are defined as Eqs. (4) and (5):

$$x_t^{d_x} = (x(t), x(t-\tau), \dots, x(t-d_x-1)\tau) \quad (4)$$

$$y_t^{d_y} = (y(t), y(t-\tau), \dots, y(t-d_y-1)\tau) \quad (5)$$

Where τ is time delay. TE brain effective connectivity was estimated via the HERMES connectivity toolbox (version 2020) [53] in MATLAB software.

2.3. Pre-trained CNN models

CNNs are popular deep learning tools that have an integrated structure including features extraction, feature reduction, and finally classification with high performance [34,35].

In a typical CNN, first, several convolution operations provide local patterns or features by sliding a mask over the sample in the convolutional layers [54–56]. This operation for the X (pixel of the input image) toward the W , weight array, and the b_l , bias of l th neuron from the feature vector, is described in Eq (6):

$$a_l^c = f \left(\sum_{i=1}^p \sum_{k=1}^p X_{(i+l-1)k} * W_{ik} \right) + b_l \quad (6)$$

where, a_l^c and p denote the output of l th neuron from feature vector and length of the convolutional kernel [55]. f is the activation function that can be a Rectified Linear Unit (ReLU). Then, the maximum or average operations reduce the size of feature maps in the pooling layer. Max pool is more common than average since it leads to the convergence of the network quicker [34]. The output of the l th neuron in this layer (a_l^{max}) is computed by Eq (7).

$$a_l^{max} = \max(a_l^c), (i=1, 2, \dots, L-p+1) \quad (7)$$

Finally, the fully connected layer performs the classification of the optimized feature maps via the softmax function (Eq (8)) which estimates the probability (Y) of each sample to each class.

$$Y = \text{softmax}(a' w' + b') \quad (8)$$

where, a' , w' and b' are output of the fully connected layer, weigh and bias, respectively.

There are multiple CNN models which have been trained on very large databases like ImageNet [57]. ImageNet database uses over one million images from one thousand classes of animals, objects and etc. Five robust pre-trained CNNs from different architectures named VGG16 [58] from straightforward CNNs, ResNet50V2 from residual network family [59], InceptionV3 [60] with multiple parallel convolutional layers, the EfficientNetB0 [61] with scaling CNNs and the DenseNet121 [62] with the densely connected convolutional network were used in this study to diagnose SZ patients. All of these models had been implemented in the Keras package with the TensorFlow backend and Python programming language.

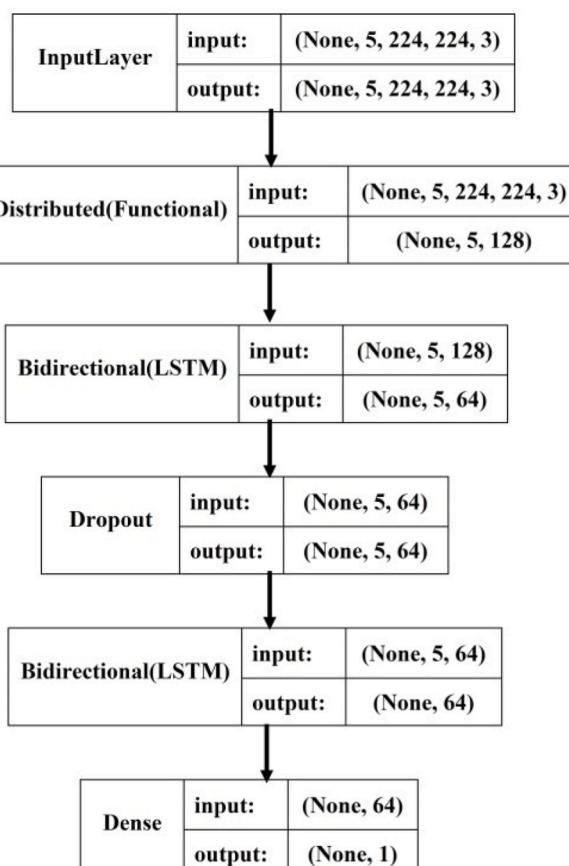


Fig. 2. Block diagram of the hybrid pre-trained CNN-LSTM model.

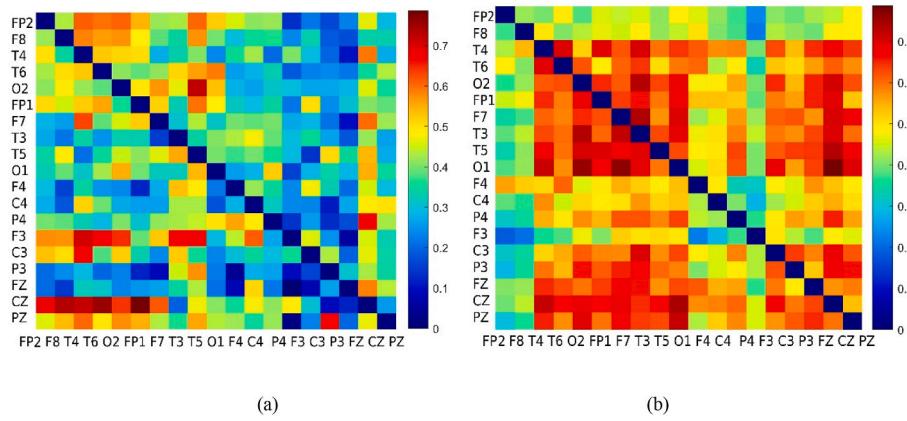


Fig. 3. TE representation for (a) schizophrenia and (b) healthy subject on 5 s time window. Each element represents the TE value between two channels, in which the values of diagonal elements are zero.

2.4. LSTM models

LSTM models are some types of deep learning algorithms that could learn information of dependency in sequential inputs to classify categories [34,35]. LSTM models can store information in memory blocks for a specific duration that provides valuable information to solve the desired problem efficiently. These memory blocks consist of multiple particular gates, named input, output and forget gates. These models have been used in the biomedical signal processing area due to their particular abilities and high performance [41–43].

2.5. Hybrid pre-trained CNN-LSTM models

The hybrid model of pre-trained CNN-LSTM provides both abilities to extract deep particular features from pre-trained CNN and explore time dependency from LSTM and eventually improves solving of the desired problem. First, every five sequential TE images with the size of $224 \times 224 \times 3$ form an input set to be fed to the CNN-LSTM model. Then, the TimeDistributed layer is utilized upon CNN models to handle all five images across the temporal dimension. 128 features will be extracted from each sample input. Bidirectional LSTM layers used in this study, exploit the most advantageous features from each input sample by analyzing it in both directions of the temporal dimension. The first bidirectional LSTM layer produces 64 features because of the 50% dropout layer. Then, the output of the second bidirectional layer with 64 neurons goes to the dense layer with the sigmoid activation function to classify SZ patients and healthy controls. The structure of the proposed hybrid CNN-LSTM is shown in Fig. 2.

2.6. Fine-tune procedure

All weighable layers of pre-trained CNN models were fine-tuned in both single and hybrid scenarios. This work adapts the weight parameters toward the SZ detection problem using TE images. Before starting the fine-tuning procedure, the original fully connected layer and classification layers are replaced by a new one for the classification of SZ and normal classes (the original fully connected layer of pre-trained CNNs was designed to classify 1000 categories) and all parameters of the network are re-trained. The weights of all network models are fine-tuned based on new constructed images using the adaptive moment estimation optimizer (ADAM) algorithm in the optimization phase and cross-entropy was chosen as the loss function. The training phase had been run on Nvidia K80 GPU with 12 GB RAM from Google Colaboratory. Python programming language has been used to write all the functions.

2.7. Statistical analysis

The 10-fold Cross-validation (10-fold CV) method was used to assess the performance of the proposed method. Consequently, five evaluation metrics of accuracy, specificity, sensitivity, F1-score and Area Under the Curve (AUC) were computed via Eq (9) to Eq (13), respectively [63]:

$$\text{Accuracy} = \frac{tp + tn}{tp + tn + fp + fn} \quad (9)$$

$$\text{specificity} = \frac{tn}{fp + tn} \quad (10)$$

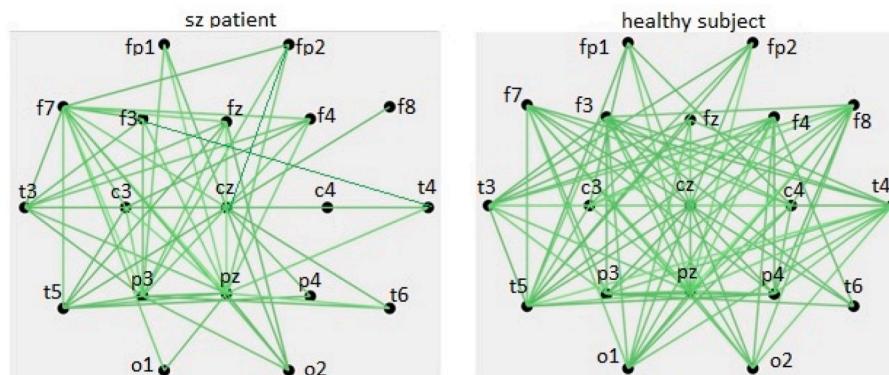


Fig. 4. Topographic map of the electrical activity of the brains of an SZ patient (left) and a healthy individual (right) and using the TE effective connectivity measure. The TE effective connectivity measure with a value of higher than 0.2 has been kept.

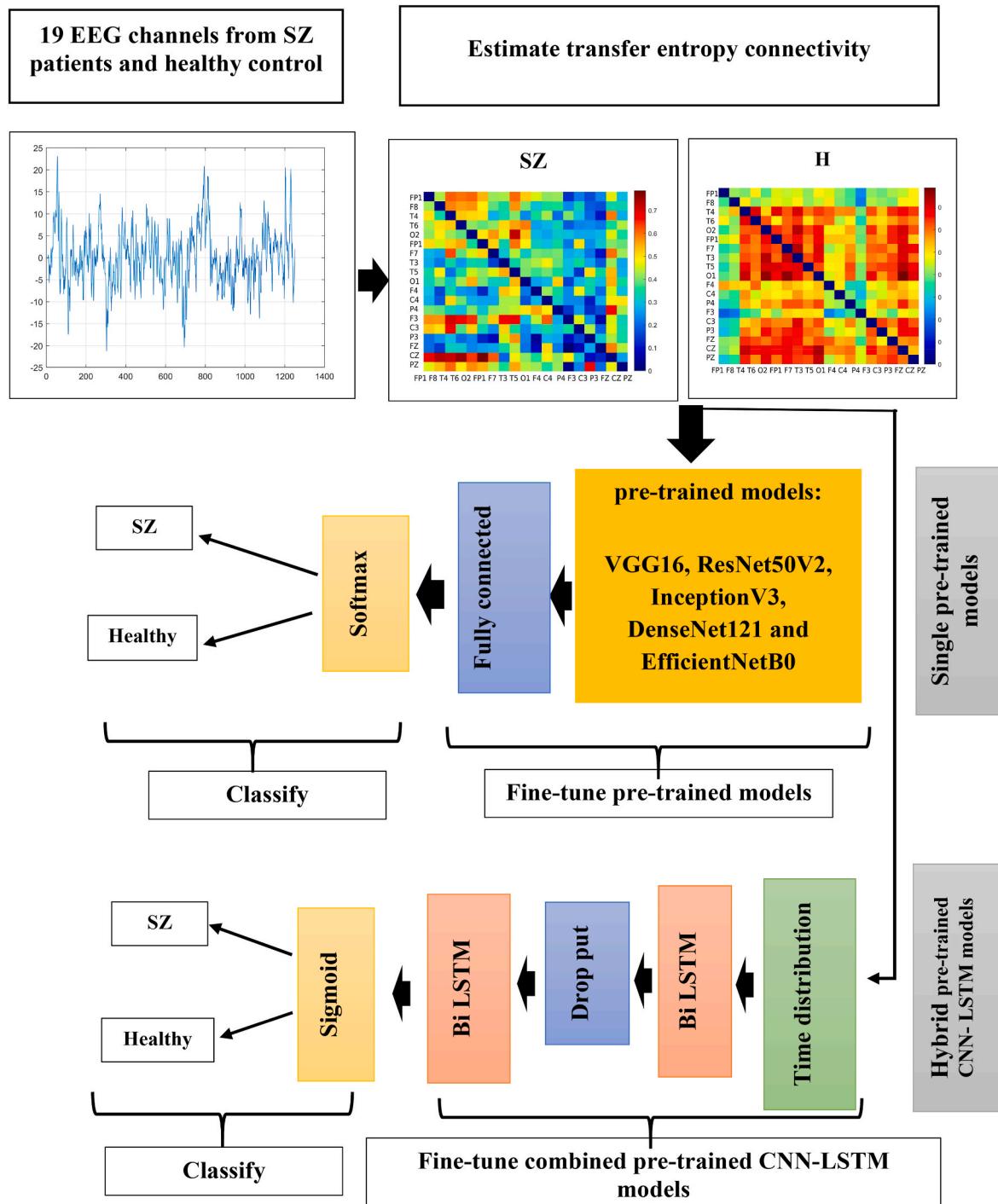


Fig. 5. Block diagram of proposed SZ detection method from multichannel EEG based on single pre-trained CNN models and hybrid CNN-LSTM models.

$$\text{sensitivity} = \frac{tp}{tp + fn} \quad (11)$$

$$F1 - \text{score} = \frac{tp}{tp + \frac{1}{2}(fn + fp)} \quad (12)$$

$$\text{AUC} = \frac{1}{2} \left(\frac{tp}{tp + fn} + \frac{tn}{tn + fp} \right) \quad (13)$$

where, tp , tn , fp and fn denote, true positive, true negative, false positive and false negative elements for two classes of SZ and healthy from the confusion matrix obtained from each deep network, respectively.

3. Results

The brain effective connectivity by TE method was estimated from each 5 s window of 19 EEG channels from 14 SZ patients and 14 healthy controls and resulted in a 19×19 asymmetric connectivity matrix. TE computation is optimized with the embedding dimension, time delay, and the number of neighbors of 3, 10, and 4, respectively. Then each 19×19 connectivity matrix is represented by a color map as a 19×19 image. Considering 12 min or 720 s of EEG signals and computation of TE on 5 s, we have 144 images or datasets for each subject. By this work, the number of datasets for independent testing in the pre-trained CNN models is increased. So, the number of images for all subjects in the two

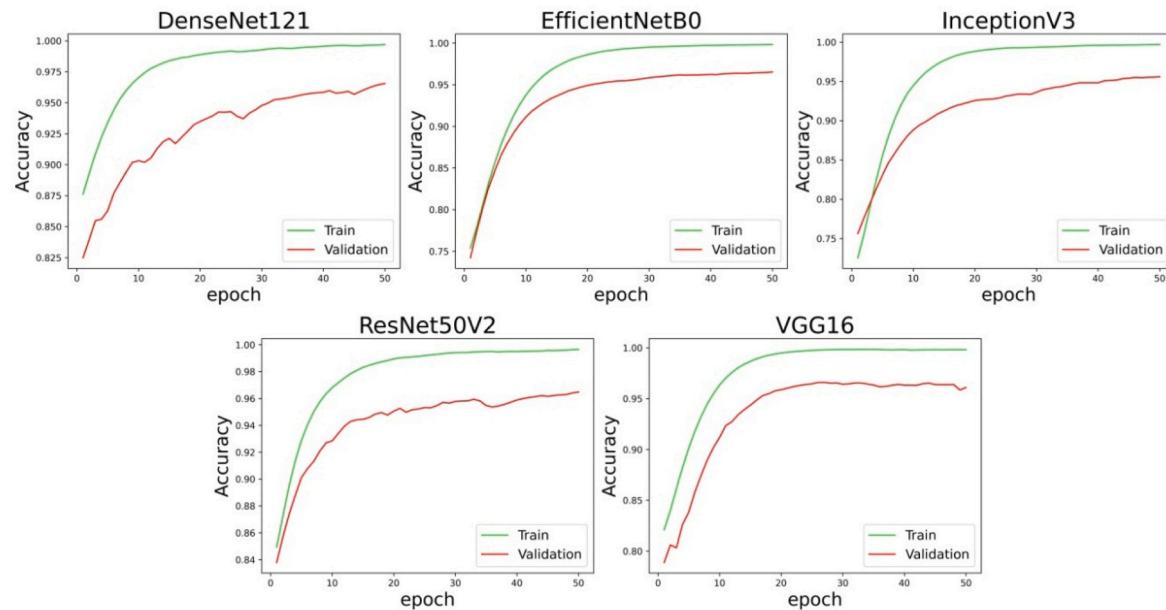


Fig. 6. The accuracy curve obtained for pre-trained CNN models on training (green color) and validating (red color) TE images at one-fold from 10-folds.

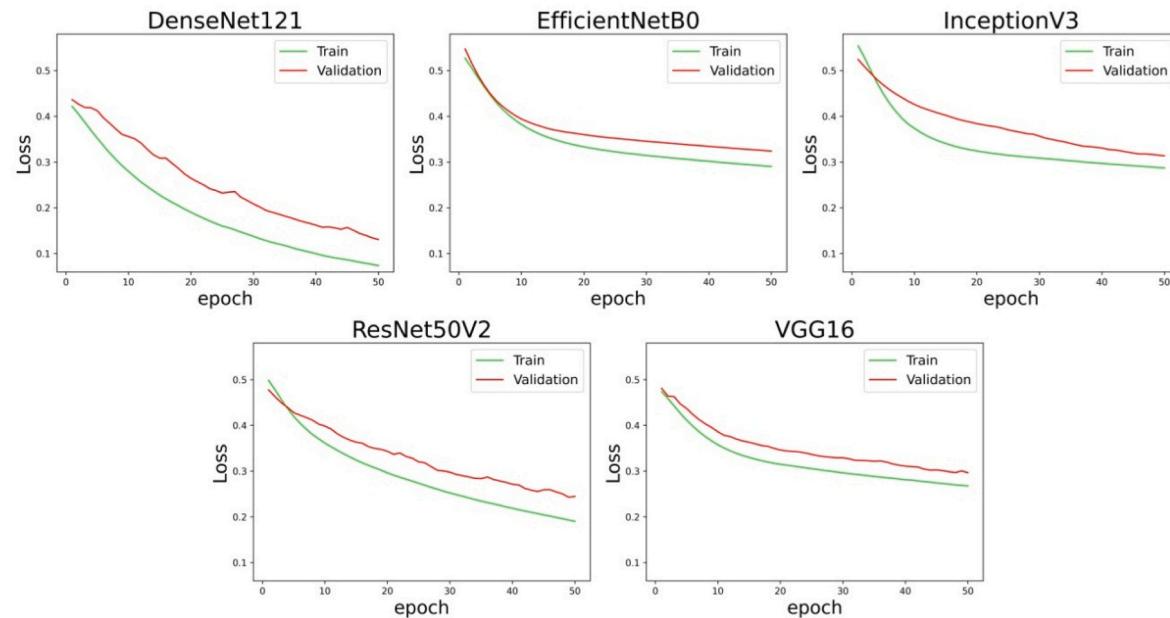


Fig. 7. The loss curve obtained for pre-trained CNN models on training (green color) and validating (red color) TE images at one-fold from 10-folds.

classes was $4032 = 28. Fig. 3 shows the average TE image of SZ patients (a) and normal participants (b). High TE values are represented by red spectrum and low values by the blue spectrum. As it can be understood from comparing Fig. 3 (a) and (b), the normal individual had higher values of TE and subsequently higher values of information flow than SZ patients in almost all brain regions. Fig. 4 shows the topography of the head using the TE measure as the effective connectivity method for one healthy individual (right) and one SZ patient (left). In this figure, the TE effective connectivity measure with a value of higher than 0.2 has been kept. As can be observed, brain regions in this healthy individual have more connections than in SZ patient in terms of TE measure. In the other words, there is lower information flow between brain regions of SZ patients than in healthy individuals.$

As mentioned in the previous section, two different strategies of deep

learning models have been investigated in this study from EEG derived brain effective connectivity images. In the first model, only pre-trained CNNs (ResNet50V2, VGG16, DenseNet121, InceptionV3, and EfficientNetB0), and in the second model, hybrid-based pre-trained CNN-LSTM models are used to diagnose SZ patients from healthy participants. Pre-trained CNN models require input images with specific sizes, therefore, each TE image is resized to be fed to each pre-trained CNN. For example, ResNet50V2 requires 224×224 images as input, so all TE images were resized according to each acceptable input size. It should be noted that TE images were fed into the LSTM model in five sets, i.e., each five TE images form a set of samples as input of the LSTM model. Therefore, the TE images of each participant include 28 sets of five images ($28 \times 5 = 144$ TE images). Therefore, we create spatial-time TE images from EEG signals and fed them into a hybrid of deep transfer learning models including pre-trained CNNs and the LSTM models.

Table 1

Main hyper-parameters of pre-trained CNN models.

Model Name	Batch size	Learning Rate
DenseNet121	20	3e-04
ResNet50V2	20	1e-04
VGG16	32	9e-05
EfficientNetB0	20	1e-05
InceptionV3	16	2e-05

Table 2Average accuracy, specificity, sensitivity, F1-score and AUC values (mean \pm std) of SZ detection from healthy controls for TE images using various pre-trained CNN models through the 10-fold CV criterion.

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)	AUC
DenseNet121	96.26 \pm 1.07	95.48 \pm 2.90	97.02 \pm 1.76	96.16 \pm 1.18	0.963 \pm 0.011
ResNet50V2	95.47 \pm 0.77	93.32 \pm 1.40	97.56 \pm 0.82	95.30 \pm 0.83	0.954 \pm 0.008
VGG16	95.27 \pm 0.97	93.52 \pm 3.00	96.98 \pm 1.63	95.10 \pm 1.11	0.953 \pm 0.010
EfficientNetB0	94.33 \pm 0.46	92.66 \pm 0.75	95.95 \pm 0.73	94.15 \pm 0.48	0.943 \pm 0.004
InceptionV3	93.94 \pm 0.82	92.06 \pm 2.87	95.76 \pm 1.72	93.72 \pm 0.96	0.939 \pm 0.085

Fig. 5 summarizes the proposed block diagram.

Figs. 6 and 7 show the accuracy and loss function curves at the fine-tuning (green color) and validation (red color) process on TE images in the first deep learning model for the five pre-trained CNN models at one-fold of 10-folds, respectively. The initial learning rate and mini-batch size of pre-trained CNN models were reported in **Table 1**. Also, the maximum epochs to train all the networks were considered 50. **Table 2** reports the results of classifying TE images of two classes of SZ patients and healthy controls using the first deep learning structure (only pre-

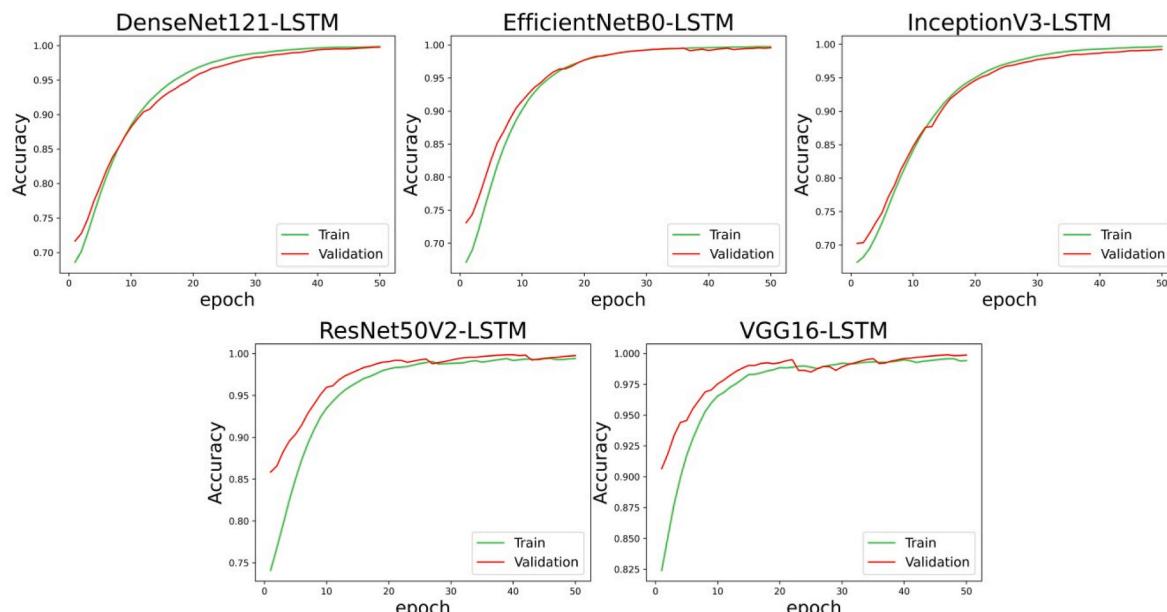
trained CNN models) through the 10-fold CV method. Comparing the five pre-trained CNN models in **Table 2**, the highest average accuracy, specificity, sensitivity, F1-score and AUC are obtained using the DenseNet121 model equal to 96.26%, 97.02%, 95.48%, 96.16% and 0.963, respectively. But, as it can be observed that the efficientNetB0 is the most stable model because it had the lowest accuracy variance (0.46 (EfficientNetB0) vs. 0.82 (InceptionV3), 0.97 (VGG16), 0.77 (ResNet50V2), and 1.07 (DenseNet121)). Also, this model had the lowest fluctuations for the validation process and lower differences between the loss and accuracy values of train and validation processes (**Figs. 6 and 7**).

Figs. 8 and 9 show the accuracy and loss function curves at the fine-tuning (green color) and validation (red color) process on TE images in the second deep learning model for the five hybrid pre-trained CNN-LSTM models at one-fold of 10-folds, respectively. The batch sizes and learning rates of each hybrid model are shown in **Table 3**. Also, the number of epochs was 50 to train all models. **Table 4** shows the results of these hybrid CNN-LSTM models on TE effective connectivity images. The highest average accuracy, sensitivity, specificity, F1-score and AUC are obtained using the EfficientNetB0-LSTM equal to 99.90%, 99.54%, 100%, 99.93% and 0.998, respectively. Also, it had the lowest accuracy variance (0.06 vs. 0.24 (ResNet50V2), 0.12 (VGG16), 0.09 (InceptionV3) and 0.08 (DenseNet121)) and therefore is the most stable hybrid model. Therefore, the second proposed model with the hybrid CNN-LSTM models had learned patterns of SZ patients and healthy controls better and had higher average accuracy and lower variance.

4. Discussion

In this paper, we proposed an SZ detection framework based on images of a valuable effective connectivity measure based on TE and the hybrid deep learning models with regard to TL techniques. The highest accuracy of 99.90% is obtained for applying the hybrid fine-tuned EfficientNetB0-LSTM model on 14 SZ patients and 14 healthy controls.

According to **Fig. 3**, there are obvious differences between the information flow of brain channels between SZ patients and normal subjects. SZ patients had lower values of TE (information flows) in most brain channels than normal subjects. According to functional Magnetic Resonance Imaging (fMRI) studies, SZ patients have impaired functionality in frontal lobes and sensory areas [64]. Also, frontal, central,

**Fig. 8.** The accuracy curve obtained for hybrid pre-trained CNN-LSTM models on training (green color) and validating (red color) TE images at one-fold from 10-folds.

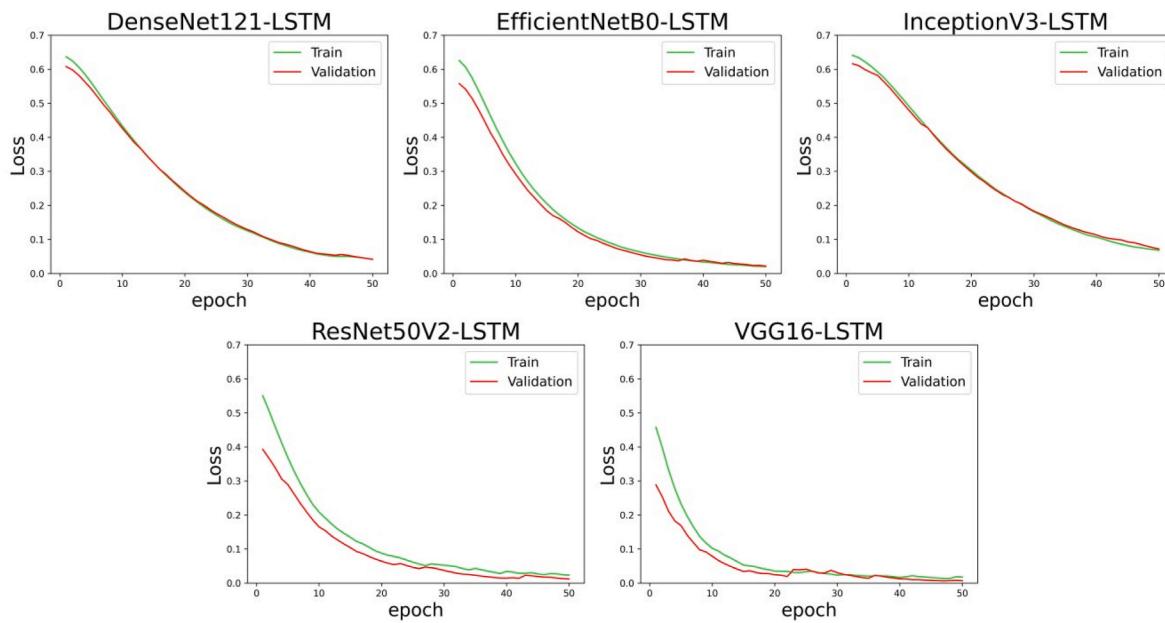


Fig. 9. The loss curve obtained for hybrid pre-trained CNN-LSTM models on training (green color) and validating (red color) TE images at one-fold from 10-folds.

Table 3

Main hyper-parameters of TL-LSTM models.

Model Name	Batch size	Learning Rate
DenseNet121-LSTM	10	2e-06
ResNet50V2-LSTM	8	1e-06
VGG16-LSTM	10	7e-06
EfficientNetB0-LSTM	12	2e-06
InceptionV3-LSTM	12	2e-06

Table 4

Average accuracy, specificity, sensitivity, F1-score and AUC values (mean \pm std) of SZ detection from healthy controls for TE images using various hybrid pre-trained CNN-LSTM models through the 10-fold CV evaluation criterion.

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)	AUC
DenseNet121-LSTM	99.87 \pm 0.08	99.74 \pm 0.26	100.0 \pm 0.00	99.87 \pm 0.03	0.998 \pm 0.03
EfficientNetB0-LSTM	99.90 \pm 0.06	99.54 \pm 0.16	100.00 \pm 0.00	99.93 \pm 0.13	0.998 \pm 0.01
InceptionV3-LSTM	99.49 \pm 0.09	99.36 \pm 0.25	99.21 \pm 0.18	99.49 \pm 0.51	0.995 \pm 0.05
VGG16-LSTM	99.87 \pm 0.12	99.32 \pm 0.38	100.00 \pm 0.00	99.82 \pm 0.44	0.998 \pm 0.03
ResNet50V2-LSTM	99.74 \pm 0.24	99.29 \pm 0.41	99.75 \pm 0.05	99.80 \pm 0.50	0.997 \pm 0.04

parietal, and occipital regions were the most discriminative regions in SZ patients [22,23,33,65,66] which are consistent with our findings.

Comparing the results of the first and second proposed models in Tables 2 and 4, it is shown that combining the CNN and LSTM methods as two deep learning models increased the average accuracy of detection of SZ patients. Each one of the hybrid pre-trained CNN-LSTM models had higher accuracy results than corresponding only pre-trained CNN models. For example, the accuracy value of DenseNet121-LSTM and EfficientNetB0-LSTM increased by nearly 3.7% (from 96.26% in only the pre-trained CNN model to 99.87% in the hybrid model), and 5.9% (from 94.33% in only pre-trained CNN model to 99.90% in the hybrid model) than corresponding only pre-trained CNN models, respectively. Also, the hybrid pre-trained CNN-LSTM models had more stability and lower

Table 5

Comparison of schizophrenia recognition studies from EEG signals on same databases.

Ref	Processing Method	Accuracy (%)
[22]	independent component analysis, FFT, random forest	80
[14]	non-linear features, decision-tree, probabilistic neural network, support vector machine	92.90
[23]	independent component analysis, calculating mean values over all sensors, spectral analysis, random forest	96.77
[15]	multivariate empirical mode decomposition, approximate entropy, sample entropy, permutation entropy, spectral entropy, and singular value decomposition entropy, support vector machine	93
[30]	1-D CNN	98.07
[31]	multi-domain connectome CNN	91.69
[26]	CWT, ResNet18-support vector machine	98.60
[20]	spectral features, Hjorth parameters, CNN, LSTM	98.56
[25]	CWT, STFT, SPWVD, CNN	93.36
[17]	microstates analysis, Nonlinear features, multivariate recursive feature elimination, support vector machine	76.85
[18]	Nonlinear features, LSTM	99
[27]	wavelet transform, K-nearest neighbor	99.21
[28]	Collatz pattern, iterative neighborhood component analysis, K-nearest neighbor	99.47
[29]	Phase space dynamic, graphical features, K-nearest neighbor, generalized regression neural Network	94.80
[32]	CNN-LSTM	99.25
Proposed method	TE, ensemble of hybrid pre-trained CNNs-LSTM models	99.90

accuracy variance than the first method with only pre-trained CNN models. For example, the accuracy variance value of DenseNet121-LSTM and EfficientNetB0-LSTM decreased to 0.08 and 0.06 from 1.07 to 0.46 in the corresponding only pre-trained CNN model, respectively. Finally, the hybrid pre-trained CNN-LSTM models had lower fluctuations for the validation process and lower differences between the loss and accuracy values of train and validation processes than those corresponding only pre-trained CNN models (Figs. 6–9). This issue shows the importance of time dependency information added by LSTM models to the proposed SZ detection framework. On the other hand, observing

sequences of TE images improves the discrimination ability of the proposed framework. The proposed LSTM model via its particular structure i.e., forget gate, input gate and output gate provided more clear information. On the other hand, extracted deep features from CNN models were investigated in forget gates and non-appropriate ones were not forgotten and fed into the input gate. Then, input gates investigate them and learned and updated relevant and zeros the non-relevant ones. Then, output gates decided which element to be transferred out and which one be zero.

The results of this study are compared with new best-related studies to classify SZ patients and normal participants from the same EEG signals in Table 5. As it is observed, the accuracy achieved in this study (99.90%) is higher than those studies with the traditional machine learning methods [14,15,17,22,23,27,28] and other deep learning methods [18,20,25,26,29,31,32] and proves the preference of the proposed method. Our method has some advantages compared to other similar studies. First, our image-making method from one-dimensional EEG signals of 19 channels i.e., brain effective connectivity measure based on information theory named TE, could better represent differences between SZ patients and healthy subjects. In other words, spatial color images based on the brain effective connectivity is created to better represent discriminative characteristics between two groups. STFT [25] and CWT [25–27] are classical time-frequency methods that have been used to convert the one-dimensional signal into a two-dimensional image. Comparing our results with [25–27] according to Table 5, proves the superiority of TE images over these time-frequency images (99.90% vs. 93.36%, 98.60% and 99.21%). Second, combining two powerful categories of deep learning techniques i.e., the pre-trained CNNs and LSTM models improved the diagnostic ability of SZ patients. Finally, this work has the advantage of the combination of the TE effective connectivity image converted from EEG signals and hybrid CNN-LSTM models.

The small sample size is the main limitation of developing deep learning models and can make models vulnerable to bias and lack of generalizability. In this study, we tried to minimize this problem using TL models. Other limitations of this study were the high computational cost and the long time needed to do the fine-tuning and evaluation steps of pre-trained CNNs, especially DenseNet121 or InceptionV3 with very layers and complicated structures through the LOSO. However, we had run them on the Google Colaboratory with no hardware limitations. In the future, we try to develop and verify the performance of other EEG signals to image conversion methods in conjunction with other deep learning schemes.

5. Conclusion

A comprehensive study using the concept of information flow from multichannel EEG signals via effective connectivity measure of TE and the hybrid well-known pre-trained CNN-LSTM models were exploited for recognition of SZ from normal subjects using a 10-fold cross-validation criterion. The highest average accuracy value of 99.90% is achieved for applying the hybrid pre-trained EfficientNetB0-LSTM model on images of the TE method. Relying on the results, the newly proposed hybrid-based model is capable of effectively analyzing brain function and produces the best results compared to all studies in recent years for the detection of SZ patients from healthy controls.

Fund

The authors did not receive support from any organization for the submitted work.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

the work reported in this paper.

References

- [1] A. Savio, J. Charpentier, M. Termenón, A.K. Shinn, M. Grana, Neural classifiers for schizophrenia diagnostic support on diffusion imaging data, *Neural Netw. World* 20 (7) (2010) 935.
- [2] I. Chatterjee, M. Agarwal, B. Rana, N. Lakhani, N. Kumar, Bi-objective approach for computer-aided diagnosis of schizophrenia patients using fMRI data, *Multimed. Tool. Appl.* 77 (2018) 26991–270153.
- [3] C.R. Bowie, P.D. Harvey, Cognitive deficits and functional outcome in schizophrenia, *Neuropsychiatric Dis. Treat.* 2 (4) (2006) 531.
- [4] E.M. Joyce, J.P. Roiser, Cognitive heterogeneity in schizophrenia, *Curr. Opin. Psychiatr.* 20 (3) (2007) 268.
- [5] L.E. DeLisi, K.U. Szulc, H.C. Bertisch, M. Majcher, K. Brown, Understanding structural brain changes in schizophrenia, *Dialogues Clin. Neurosci.* 8 (1) (2006) 71.
- [6] K.H. Karlsgodt, D. Sun, T.D. Cannon, Structural and functional brain abnormalities in schizophrenia, *Curr. Dir. Psychol. Sci.* 19 (4) (2010) 226–231.
- [7] M. Kubicki, R.W. McCarley, M.E. Shenton, Evidence for white matter abnormalities in schizophrenia, *Curr. Opin. Psychiatr.* 18 (2) (2005) 121.
- [8] M. Giraldo-Chica, B.P. Rogers Bp, S.M. Damon, B.A. Landman, N.D. Woodward, Prefrontal-thalamic anatomical connectivity and executive cognitive function in schizophrenia, *Biol. Psychiatr.* 83 (6) (2018) 509–517.
- [9] Q.X. Yang, Y.X. Wang, F.C. Li, S. Zhang, Y.C. Luo, Y. Li, J. Tang, B. Li, Y.Z. Chen, W. W. Xue, F. Zhu F, Identification of the gene signature reflecting schizophrenia's etiology by constructing artificial intelligence-based method of enhanced reproducibility, *CNS Neurosci. Ther.* 25 (9) (2019) 1054–1063.
- [10] Q. Yang, B. Li, J. Tang, X. Cui, Y. Wang, X. Li, J. Hu, Y. Chen, W. Xue, Y. Lou, Y. Qiu, Consistent gene signature of schizophrenia identified by a novel feature selection strategy from comprehensive sets of transcriptomic data, *Briefings Bioinf.* 21 (3) (2020) 1058–1068.
- [11] World Health Organization, Schizophrenia. https://www.who.int/mental_health/management/schizophrenia/en/, 2020. (Accessed 13 June 2021).
- [12] M. Sadat Shahabi, A. Shalbaf, A. Maghsoudi, Prediction of drug response in major depressive disorder using ensemble of transfer learning with convolutional neural network based on EEG, *Biocybern. Biomed. Eng.* 41 (3) (2021) 946–959.
- [13] A. Saeedi, M. Saeedi, A. Maghsoudi, A. Shalbaf, Major depressive disorder diagnosis based on effective connectivity in EEG signals: a convolutional neural network and long short-term memory approach, *Cognitive Neurodynamics* 15 (2) (2021) 239–252.
- [14] V. Jahmunah, S.L. Oh, V. Rajnikanth, E.J. Ciaccio, K.H. Cheong, N. Arunkumar, U. R. Acharya, Automated detection of schizophrenia using nonlinear signal processing methods, *Artif. Intell. Med.* 100 (2019), 101698.
- [15] P.T. Krishnan, A.N.J. Raj, P. Balasubramanian, Y. Chen, Schizophrenia detection using Multivariate Empirical Mode Decomposition and entropy measures from multichannel EEG signal, *Biocybern. Biomed. Eng.* 40 (3) (2020) 1124–1139.
- [16] J.R. Sun ., Cao, M. Zhou, W. Hussain, B. Wang, J. Xue, J. Xiang, A hybrid deep neural network for classification of schizophrenia using EEG Data, *Sci. Rep.* 11 (1) (2021) 1–16.
- [17] K. Kim, N.T. Duc, M. Choi, B. Lee, EEG microstate features for schizophrenia classification, *PLoS One* 16 (5) (2021), e0251842.
- [18] A.N. Chandran, K. Sreekumar, D.P. Subha, EEG-based automated detection of schizophrenia using long short-term memory (LSTM) network, in: *Advances in Machine Learning and Computational Intelligence*, Springer, Singapore, 2021, pp. 229–236.
- [19] I.E. Kutepov I, V.V. Dobriyan, M.V. Zhigalov, M.F. Stepanov, A.V. Krysko, T. V. Yakovleva, V.A. Krysko, EEG analysis in patients with schizophrenia based on Lyapunov exponents, *Inform. Med. Unlocked* 18 (2020), 100289 .E.
- [20] K. Singh, S. Singh, J. Malhotra, Spectral features based convolutional neural network for accurate and prompt identification of schizophrenic patients, *Proc. IME H J. Eng. Med.* 235 (2) (2021) 167–184.
- [21] K. Das, R.B. Pachori, Schizophrenia detection technique using multivariate iterative filtering and multichannel EEG signals, *Biomed. Signal Process Control* 67 (2021), 102525.
- [22] R. Buettner, M. Hirschmiller, K. Schlosser, M. Rössle, M. Fernandes, I.J. Timm, High-performance exclusion of schizophrenia using a novel machine learning method on EEG data, *Health* (2019) 1–6.
- [23] R. Buettner, D. Beil, S. Scholtz, A. Djemai, Development of a machine learning based algorithm to accurately detect schizophrenia based on one-minute EEG recordings, in: *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 2020.
- [24] Z. Dvye-Aharon, N. Fogelson, A. Peled, N. Intrator, Schizophrenia detection and classification by advanced analysis of EEG recordings using a single electrode approach, *PLoS One* 10 (4) (2015), e0123033.
- [25] S.K. Khare, V. Bajaj, U.R. Acharya, SPWVD-CNN for automated detection of schizophrenia patients using EEG signals, *IEEE Trans. Instrum. Meas.* 70 (2021) 1–9.
- [26] A. Shalbaf, S. Bagherzadeh, A. Maghsoudi, Transfer learning with deep convolutional neural network for automated detection of schizophrenia from EEG signals, *Phys. Eng. Sci. Med.* 43 (4) (2020) 1229–1239.
- [27] M. Sharma, U.R. Acharya, Automated Detection of Schizophrenia Using Optimal Wavelet-Based L_1 -Norm Features Extracted from Single-Channel EEG, *Cognitive Neurodynamics*, 2021, pp. 1–14.

- [28] M. Baygin, O. Yaman, T. Tuncer, P.D. Dogan, S. Barua, U.R. Acharya, Automated accurate schizophrenia detection system using Collatz pattern technique with EEG signals, *Biomed. Signal Process Control* 70 (2021), 102936.
- [29] A. Akbari, S. Ghofrani, P. Zakalvand, M.T. Sadiq, Schizophrenia recognition based on the phase space dynamic of EEG signals and graphical features, *Biomed. Signal Process Control* 69 (2021), 102917.
- [30] S.L. Oh, J. Vicnesh, E.J. Ciaccio, R. Yuvaraj, U.R. Acharya, Deep convolutional neural network model for automated diagnosis of schizophrenia using EEG signals, *Appl. Sci.* 9 (14) (2019) 2870.
- [31] C.R. Phang, F. Noman, H. Hussain, C.M. Ting, H. Ombao, A multi-domain connectome convolutional neural network for identifying schizophrenia from EEG connectivity patterns, *IEEE J.Biomed. Health Info.* 24 (5) (2019) 1333–1343.
- [32] A. Shoeibi, D. Sadeghi, P. Moridian, N. Ghassemi, J. Heras, R. Alizadehsani, A. Khadem, Y. Kong, S. Nahavandi, J.M. Gorritz, Automatic diagnosis of schizophrenia using EEG signals and CNN-LSTM models, *arXiv preprint arXiv.* 2109 (2021), 01120.
- [33] E. Olejarczyk, W. Jernajczyk, Graph-based analysis of brain connectivity in schizophrenia, *PLoS One* 12 (11) (2017), e0188629.
- [34] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, M.S. Lew, Deep learning for visual understanding: a review, *Neurocomputing* 187 (2016) 27–48.
- [35] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*, 2016. Cambridge.
- [36] H. Greenspan, B. Van Ginneken, R.M. Summers, Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique, R.M., *IEEE Trans. Med. Imag.* 35 (5) (2016) 1153–1159.
- [37] W. Sun, T.L.B. Tseng, J. Zhang, W. Qian, Enhancing deep convolutional neural network scheme for breast cancer diagnosis with unlabeled data, *Comput. Med. Imag. Graph.* 57 (2017) 4–9.
- [38] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, et al., A survey on deep learning in medical image analysis, *Med. Image Anal.* 42 (2017) 60–88.
- [39] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, S. Thrun, Dermatologist-level classification of skin cancer with deep neural networks, *Nature* 542 (7639) (2017) 115–118.
- [40] R. Poplin, A.V. Varadarajan, K. Blumer, Y. Liu, M.V. McConnell, et al., Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning, *Nat. Biomed. Eng.* 2 (3) (2018) 158–164.
- [41] S. Alhagry, A.A. Fahmy, R.A. El-Khoribi, Emotion recognition based on EEG using LSTM recurrent neural network, *Emotion* 8 (10) (2017) 355–358.
- [42] P. Wang, A. Jiang, X. Liu, J. Shang, L. Zhang, LSTM-based EEG classification in motor imagery tasks, *IEEE Trans. Neural Syst. Rehabil. Eng.* 26 (11) (2018) 2086–2095.
- [43] H. Zeng, C. Yang, G. Dai, F. Qin, J. Zhang, W. Kong, EEG classification of driver mental states by deep learning, *Cognit. Neurodyn.* 12 (6) (2018) 597–606.
- [44] [database] E. Olejarczyk, W. Jernajczyk, EEG in schizophrenia, RepOD, 2017, <https://doi.org/10.18150/repod.0107441>. V1.
- [45] A. Delorme, S. Makeig, EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, *J. Neurosci. Methods* 134 (1) (2004) 9–21.
- [46] R. Vicente, M. Wibral, M. Lindner, G. Pipa, Transfer entropy—a model-free measure of effective connectivity for the neurosciences, *J. Comput. Neurosci.* 30 (1) (2011) 45–67.
- [47] V.A. Vakorin, N. Kovacevic, A.R. McIntosh, Exploring transient transfer entropy based on a group-wise ICA decomposition of EEG data, *Neuroimage* 49 (2) (2010) 1593–1600.
- [48] M.H.I. Shovon, N. Nandagopal, R. Vijayalakshmi, J.T. Du, B. Cocks, Directed connectivity analysis of functional brain networks during cognitive activity using transfer entropy, *Neural Process. Lett.* 45 (3) (2017) 807–824.
- [49] M. Ursino, G. Ricci, E. Magosso, Transfer entropy as a measure of brain connectivity: a critical analysis with the help of neural mass models, *Front. Comput. Neurosci.* 14 (2020) 45.
- [50] C.S. Huang, N.R. Pal, C.H. Chuang, C.T. Lin, Identifying changes in EEG information transfer during drowsy driving by transfer entropy, *Front. Hum. Neurosci.* 9 (2015) 570.
- [51] C. Ciprian, K. Masychev, M. Ravan, J.P. Reilly, D. MacCrimmon, A machine learning approach using effective connectivity to predict response to clozapine treatment, *IEEE Trans. Neural Syst. Rehabil. Eng.* 28 (12) (Dec. 2020) 2598–2607.
- [52] Y. Gao, X. Wang, T. Potter, J. Zhang, Y. Zhang, Single-trial EEG Emotion recognition using granger causality/transfer entropy analysis, *Y. J. Neurosci. Methods* 346 (2020), 108904.
- [53] G. Niso, R. Bruña, E. Pereda, R. Gutiérrez, R. Bajo, et al., HERMES: towards an integrated toolbox to characterize functional and effective brain connectivity, *Neuroinformatics* 11 (4) (2013) 405–434.
- [54] A.S. Rifaioglu, H. Atas, M.J. Martin, R. Cetin-Atalay, V. Atalay, T. Dogan, Recent applications of deep learning and machine intelligence on *in silico* drug discovery: methods, tools and databases, *Briefings Bioinf.* 20 (5) (2019) 1878–1912.
- [55] J. Hong, Y. Luo, Y. Zhang, J. Ying, W. Xue, T. Xie, L. Tao, F. Zhu, Protein functional annotation of simultaneously improved stability, accuracy and false discovery rate achieved by a sequence-based deep learning, *Briefings Bioinf.* 21 (4) (2020) 1437–1447.
- [56] J. Hong, Y. Luo, M. Mou, J. Fu, Y. Zhang, W. Xue, T. Xie, L. Tao, Y. Lou, F. Zhu, Convolutional neural network-based annotation of bacterial type IV secretion system effectors with enhanced accuracy and reduced false discovery, *Briefings Bioinf.* 21 (5) (2020) 1825–1836.
- [57] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, *Adv. Neural Inf. Process. Syst.* 25 (2012) 1097–1105.
- [58] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, *arXiv preprint arXiv.1409* (2014) 1556.
- [59] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.
- [60] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2818–2826.
- [61] M. Tan, Q.V. Le, EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, 2019, p. 1194. ArXiv Preprint ArXiv.1905.
- [62] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4700–4708.
- [63] M. Sokolova, G. Lapalme, A systematic analysis of performance measures for classification tasks, *Inf. Process. Manag.* 45 (4) (2009) 427–437.
- [64] G.G. Brown, W.K. Thompson, Functional brain imaging in schizophrenia: selected results and methods, in: *Behavioral neurobiology of schizophrenia and its treatment*, 2010, pp. 181–214.
- [65] F. Li, J. Wang, Y. Liao, C. Yi, Y. Jiang, et al., Differentiation of schizophrenia by combining the spatial EEG brain network patterns of rest and task P300, *IEEE Trans. Neural Syst. Rehabil. Eng.* 27 (4) (2019) 594–602.
- [66] T. Liu, J. Zhang, X. Dong, Z. Li, X. Shi, Occipital alpha connectivity during resting-state electroencephalography in patients with ultra-high risk for psychosis and schizophrenia, *Front. Psychiatr.* 10 (2019) 553.