

## Research Article

# SchizoGoogLeNet: The GoogLeNet-Based Deep Feature Extraction Design for Automatic Detection of Schizophrenia

Siuly Siuly <sup>1</sup>, Yan Li <sup>2</sup>, Peng Wen <sup>3</sup> and Omer Faruk Alcin<sup>4</sup>

<sup>1</sup>Institute for Sustainable Industries & Liveable Cities, Victoria University, Melbourne, Australia

<sup>2</sup>School of Mathematics, Physics and Computing, University of Southern Queensland, Toowoomba, Australia

<sup>3</sup>School of Engineering, University of Southern Queensland, Toowoomba, Australia

<sup>4</sup>Department of Electrical and Electronics Engineering, Turgut Ozal University, Malatya, Turkey

Correspondence should be addressed to Siuly Siuly; [siuly\\_1976@yahoo.com](mailto:siuly_1976@yahoo.com)

Received 22 June 2022; Accepted 8 August 2022; Published 8 September 2022

Academic Editor: Abdul Rehman Javed

Copyright © 2022 Siuly Siuly et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Schizophrenia (SZ) is a severe and prolonged disorder of the human brain where people interpret reality in an abnormal way. Traditional methods of SZ detection are based on handcrafted feature extraction methods (manual process), which are tedious and unsophisticated, and also limited in their ability to balance efficiency and accuracy. To solve this issue, this study designed a deep learning-based feature extraction scheme involving the GoogLeNet model called “SchizoGoogLeNet” that can efficiently and automatically distinguish schizophrenic patients from healthy control (HC) subjects using electroencephalogram (EEG) signals with improved performance. The proposed framework involves multiple stages of EEG data processing. First, this study employs the average filtering method to remove noise and artifacts from the raw EEG signals to improve the signal-to-noise ratio. After that, a GoogLeNet model is designed to discover significant hidden features from denoised signals to identify schizophrenic patients from HC subjects. Finally, the obtained deep feature set is evaluated by the GoogLeNet classifier and also some renowned machine learning classifiers to find a sustainable classification method for the obtained deep feature set. Experimental results show that the proposed deep feature extraction model with a support vector machine performs the best, producing a 99.02% correct classification rate for SZ, with an overall accuracy of 98.84%. Furthermore, our proposed model outperforms other existing methods. The proposed design is able to accurately discriminate SZ from HC, and it will be useful for developing a diagnostic tool for SZ detection.

## 1. Introduction

Schizophrenia (SZ) is a devastating mental disorder and a progressive neurological disease that causes a significant impact on the quality of life of patients and their families, social environments, and healthcare systems [1]. This disorder affects a person’s perception of reality, social interactions, thought processes, and cognitive ability. Symptoms of SZ include hallucinations (hearing voices or seeing things that are not there), delusions (fixed, false beliefs), and thought disorders (unusual ways of thinking), as well as reduced expressions of emotions, reduced motivations to accomplish goals, difficulty in social relationships, motor impairments, and cognitive impairments

[2, 3]. SZ affects every 1 in 100 Australians and approximately 21 million people worldwide [4, 5]. It is associated with increased morbidity and mortality, and is the cause of disability and health costs worldwide [6]. SZ is treatable, but its treatment involves long-term medications, causing an extreme burden on healthcare systems and families. If the patients are not treated, SZ symptoms may be persistent which makes them disabled after a period of time. If early detection is possible, then patients can get timely treatments that can help those affected individuals and their families improve their lives [7]. Hence, there is a growing demand to develop an efficient and automatic diagnostic technique for the early detection of SZ patients from healthy control (HC) people.

Traditionally, the diagnosis of SZ is mainly performed through solely interviews and observations of patient behavior by a trained psychiatrist [8]. This diagnosis is a manual process that is time-consuming, burdensome, and subject to human errors and bias. Instead of this, recently, some imaging techniques, such as magnetic resonance imaging, computed tomography, positron emission tomography, and electroencephalography (EEG) have been introduced to diagnose SZ. Among these imaging techniques, currently, EEG has emerged as the reference standard for diagnosing SZ due to its high temporal resolution, noninvasiveness, and relatively low financial cost compared to other tests [2, 9]. EEG signals describe the electrical activity of the human brain recorded from the scalp via a set of electrodes. EEG recording generates a huge volume of data. Generally, this massive amount of data is analysed by visual inspection, which is time-consuming, error-prone, and reduces decision-making reliability. Thus, a computer-aided automatic data analysis system is required to make an accurate and reliable decision for the diagnosis of SZ.

*1.1. Related Works.* In recent years, many researchers have attempted to identify SZ from EEG data [10–18], but no efficient and reliable system has been developed yet for the reliable detection of SZ patients. Much research has been performed based on traditional handcrafted feature extraction methods, but they were inadequate in their ability to generate suitable performance for real-time applications. For example, Sabeti et al. [10] used autoregressive (AR) model coefficients, band power, and fractal dimension-based features for identifying SZ subjects. These features were fed to linear discriminant analysis (LDA), multiLDA, and adaptive boosting (AdaBoost) classifiers. An empirical mode decomposition (EMD) technique was introduced in [2] for the diagnosis of SZ from EEG signals to handle the behavior of nonstationary and nonlinear EEG signals. In [11], Ramos et al. employed power spectral density-based features in the maximum likelihood theory for classifying SZ and HC subjects.

Kaplan et al. [12] used spectral features, and the obtained features were classified using the “Kora-N” algorithm. A fast Fourier transformation (FFT)-based feature extraction process was reported by Buettner et al. in [13]. Finally, those features were used as input to a random forest (RF) classification method for identifying SZ and non-SZ. Approximate entropy, Shannon entropy, Kolmogorov complexity, and Lempel-Ziv complexity methods were proposed by Akar et al. in [14] for extracting features from EEG signals for identifying SZ. In their another work, Akar et al. [15] computed features using wavelet transformation (WT) and Welch power spectral density (PSD) methods for the detection of schizophrenia from EEG data. In [16], Li et al. used functional EEG networks to extract the inherent spatial pattern of the network (SPN) feature for brain states. The combined SPN features of the rest and task networks were used as the input to LDA and the support vector machine (SVM) to recognize SZ.

Only a few studies have been performed on deep learning for the detection of SZ using EEG data. In deep

learning, both the feature extraction and classification processes are conducted automatically, while traditional techniques require features to be extracted manually. Phang et al. [17] proposed a deep convolutional neural network (CNN) framework for the classification of SZ. In that model, the authors integrated the features from various domains and dimensions using different fusion strategies, and the model achieved 93.06% accuracy. An eleven-layered CNN model was introduced by Oh et al. in [18] to analyse the EEG signals for the diagnosis of schizophrenia. The model generated a classification accuracy of 81.26% for subject-based testing.

*1.2. Motivations.* Most of the existing research studies for SZ detection employ handcrafted feature extraction methods (e.g., WT, FFT, EMD, PSD, and entropy) before classification [10–16]. These feature extraction methods are manually chosen based on researchers’ expertise. The obtained performances of those methods are not satisfactory. The handcrafted feature extraction methods cannot form intellectual high levels of representations of EEG signals to discover deep concealed characteristics from data that can achieve better performance. This manual feature extraction process is time-consuming, labor-intensive, and has bias. Furthermore, if sizes of data are large, the method may not run properly and sometimes underperform.

As mentioned before, very few studies used deep learning methods for the detection of schizophrenia from EEG [17, 18]. The significant characteristic of deep learning is that the model can automatically extract effective features, which has certain advantages for large-scale data. The developed deep learning-based SZ detection methods are still limited in their ability to balance efficiency and accuracy. Hence, this study is motivated to introduce a new deep learning-based feature extraction scheme for automatic and efficient identification of SZ patients using EEG.

*1.3. Objectives of This Study.* The key objective of this study is to introduce a deep learning-based feature extraction scheme involving the GoogLeNet model called “SchizoGoogLeNet” for automatically and efficiently distinguishing schizophrenic patients from HC subjects using EEG data with improved performance. The reason for considering the GoogLeNet model in this study is that the GoogLeNet architecture has the ability to produce better performance than other deep learning models as it is designed to be a powerhouse with increased computational efficiency. This model trains faster than other network methods (e.g., AlexNet and VGGNet) and has fewer parameters and lower computational complexity than other models [19, 20]. Also, it has a relatively lower error rate than other deep learning methods.

To our knowledge, for the first time, GoogLeNet was introduced in SZ detection using EEG signals in this study. Our proposed “SchizoGoogLeNet” framework consists of several steps. First, this study employs an average filtering method to remove noise and artifacts from the raw EEG signals. Second, significant hidden features are extracted from EEG signals to identify schizophrenia patients from

HC subjects. Finally, a sustainable classification method for the obtained deep GoogLeNet features is explored. At this stage, deep learning and machine learning methods are tested for the obtained deep feature set. The performance of the proposed framework was tested using a publicly available EEG dataset.

*1.4. Our Contributions.* The main contributions of this work are summarized as follows: (1) For the first time, we introduce a GoogLeNet-based deep feature extraction scheme for automatic detection of SZ using EEG signals; (2) we investigate the impact of balance and unbalanced datasets on the proposed SZ detection model; (3) we discover a sustainable classifier for the obtained GoogLeNet deep feature set in a deep learning and machine learning environment; (4) we explore the performances of the deep feature set with a GoogLeNet classifier and several machine learning methods; (5) we achieve improved performances for our proposed model compared to those of the existing methods.

The rest of this paper is arranged as follows: The proposed methodology is described in Section 2. This section also provides the description of the data that is used in this study. Section 3 provides the experimental setup and the results with their corresponding discussions, followed by the conclusion in Section 4.

## 2. Proposed Methodology

This study proposes a GoogLeNet-based deep feature extraction scheme, “SchizoGoogLeNet,” for detecting SZ from EEG data automatically. Figure 1 graphically presents the general architecture of the proposed strategy. The proposed scheme involves four stages, such as data acquisition, removing noise and artifacts using average filtering, discovering deep features using GoogLeNet, and exploring sustainable classifiers for the obtained features. A detailed description of these steps is provided below.

*2.1. Data Acquisition.* This study used EEG data collected from 81 subjects, including 49 schizophrenia patients and 32 normal control persons, from the Kaggle data source. These data consist of EEG recordings of 14 female and 67 male subjects. The average age is 39 years, and the average education level is 14.5 years. The EEG data were recorded from 64 scalp sites and 8 external sites using a BioSemi active two system (<http://www.biosemi.com>). The data were continuously digitized at 1024 Hz and referenced offline to averaged earlobe electrodes. The details of the dataset can be seen online at <https://www.kaggle.com/broach/button-tone-sz> [21]. The description of the data is also available in [22].

*2.2. Removing Noise and Artifacts Using Average Filtering.* This section aims to introduce a technique for reducing noise and artifacts of EEG signals and improving the signal-to-noise-ratio (SNR) as the signal data are very noisy and often affected by artifacts. Due to noise, EEG signals have a low SNR that may lead to incorrect conclusions. The noise-

removing processes improve the data quality to create an appropriate model. In this study, we employ an average filter (AF) technique for removing noise and artifacts from EEG signals [23]. The reason for using this filtering is that it is simple, intuitive, and easy to implement for smoothing the signals and for dropping the amount of intensity variations that can reduce undesirable information from the signals.

In this study, we reduced noise smoothing of each signal ( $Sg$ ) by the AF technique with a kernel size of  $KS = 12$ , and then, we sampled the smoothed signal considering an interval length  $IL = KS = 12$  which is denoted as  $\hat{Sg}$ . We selected the values of  $KS$  and  $IL$  using a trial and error procedure to make the final matrix with a size of around 200. After the filtering, the raw signal data become the input matrix in the feature extraction stage of the GoogLeNet model (discussed in the next step). Figure 2 shows an example of the pattern of a raw EEG signal and a filtered EEG signal. It is apparent from Figure 2 that the shape of the filtered signal is the same as that of the raw signal. As can be seen in Figure 2, the number of data points has been reduced while still keeping the shape of the original curve of the EEG signal.

$$\overline{Sg}(t) = \frac{1}{KS} \sum_{i=t-KS/2}^{t+KS/2} Sg(i). \quad (1)$$

$$\hat{Sg}(t) = \sum_{i=0}^L \delta(i - t \cdot IL) \cdot \overline{Sg}(i). \quad (2)$$

*2.3. Discovering Deep Features Using GoogLeNet.* Technically, a feature represents a distinguishing property, a recognizable measurement, and a functional component obtained from a segment of a pattern [24]. Extracted features convey the most important information for the classification stage. Sometimes, traditional handcrafted features cannot convey meaningful information about the SZ detection due to manual choice of methods and also cannot handle big sizes of data. This section’s goal is to discover the significant feature set from EEG signals using the deep learning method that empowers efficiently the recognition of SZ from HC subjects. For this purpose, this study introduces the deep GoogLeNet-based architecture to extract representative features for identifying SZ from the denoised EEG signals automatically. To our knowledge, for the first time, GoogLeNet is employed in this study for the detection of SZ from EEG signals. GoogLeNet is a convolutional neural network (CNN)-based architecture designed by researchers at Google. It was the winner of ImageNet 2014, where it proved to be a powerful model.

The main objective of the GoogLeNet architecture is to achieve high accuracy with a reduced computational cost [19, 20]. The GoogLeNet model is constructed based on the inception architecture that introduced the new concept of the inception block in CNN, whereby it incorporates multiscale convolutional transformations using split, transform, and merge ideas [20]. A general strategy of the inception block is illustrated in Figure 3. The inception

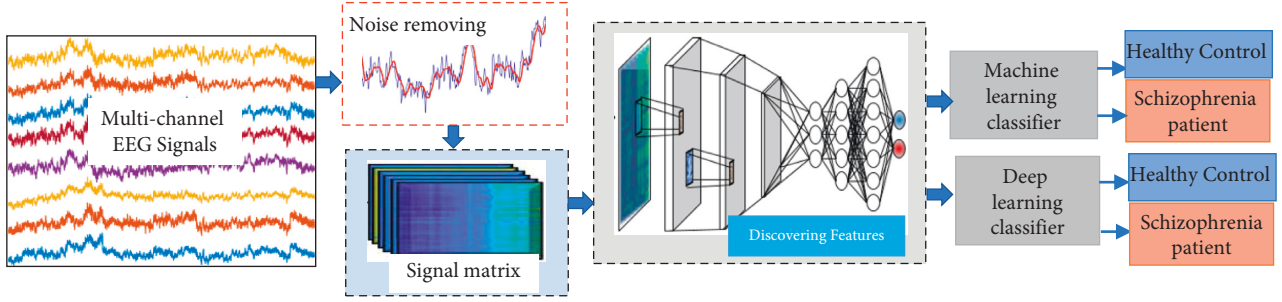


FIGURE 1: The overall architecture of the proposed “SchizoGoogLeNet” framework for automatic identification of SZ from EEG signals.

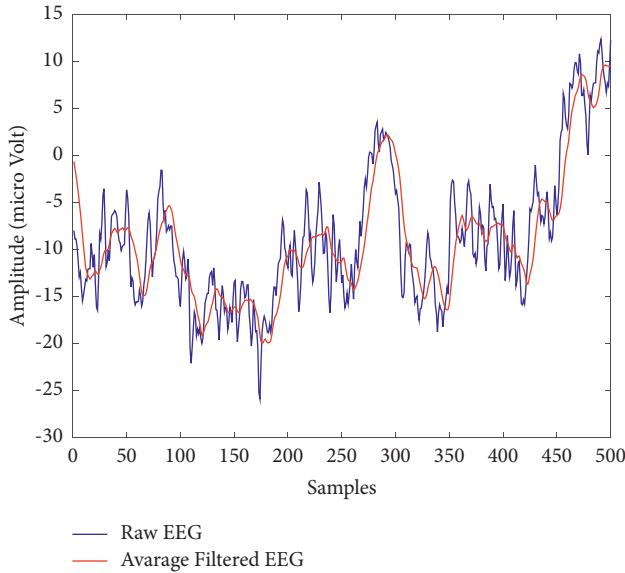


FIGURE 2: An illustration of the pattern of a raw EEG and filtered EEG signal.

module is different from other deep learning architectures where there is a fixed convolution size for each layer. In the inception module,  $1 \times 1$ ,  $3 \times 3$ , and  $5 \times 5$  convolutions and  $3 \times 3$  max pooling perform in a parallel way at the input, and the output of these is stacked together to generate the final output. In the GoogLeNet model, conventional convolutional layers are replaced with small blocks. These blocks condense filters of different sizes (e.g.,  $1 \times 1$ ,  $3 \times 3$ , and  $5 \times 5$ ) to capture spatial information at different scales, including both fine and coarse grain levels [19, 20]. As shown in Figure 3, multiple convolutions, with  $1 \times 1$  filters,  $3 \times 3$  filters, and  $5 \times 5$  filters, and  $3 \times 3$  max-pooling layers are organised in the GoogLeNet model.

The GoogLeNet model regulates the computations by adding a bottleneck layer of  $1 \times 1$  convolutional filters before employing large-size kernels.  $1 \times 1$  convolution is used to decrease the number of parameters (weights and biases) of the architecture. Furthermore, it uses sparse connections (not all the output feature maps are connected to all the input feature maps) to overcome the problem of redundant information and reduced costs by omitting feature maps that are not relevant [20]. Additionally, connection density is reduced by using global average pooling at the last layer

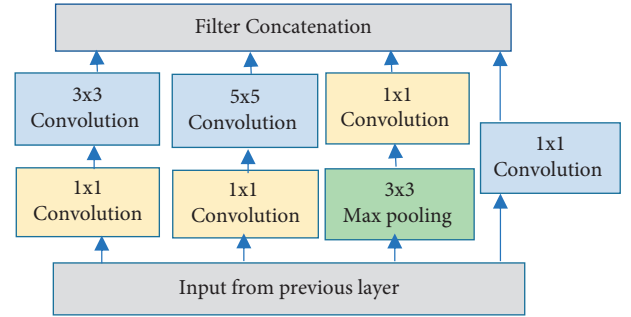


FIGURE 3: General architecture of the inception block.

instead of using a fully connected layer. These parameter tunings cause a significant decrease in the number of parameters [25].

In this study, we designed the structure of the GoogLeNet model for implementation of the SZ EEG database as illustrated in Table 1. This table presents layer-by-layer architectural details of the GoogLeNet model. “ $1 \times 1$  # $3 \times 3$  # $5 \times 5$ ” refers to various convolution filters used within the inception module. “ $3 \times 3$  reduce” and “ $5 \times 5$  reduce” symbolize the number of  $1 \times 1$  filters in the reduction layer used before related convolution layers. The number of  $1 \times 1$  filters in the projection layer after the built-in maximum pooling is shown in the “pool projection” column (denoted as “Pool proj”). “Max pool” stands for the maximum number of pooling layers. The purpose of these max-pooling layers is to downsample the input as it is fed forward through the network. All the convolution, reduction, and projection layers inside this architecture use rectified linear units (ReLUs) as their activation functions. This architecture is 22 layers deep without pooling (or 27 layers when we count pooling) [19, 20].

**2.4. Exploring Sustainable Classification Method for Identifying Schizophrenia.** This section’s aim is to discover a sustainable classifier for the obtained deep feature set to classify SZ and HC subjects and improve its performance. Unlike other deep learning models, GoogLeNet does not use fully connected (fc) layers for the final result of classification. Instead, the last convolutional map is subjected to channel-wise global average pooling, and the average activation values of each of the channels are used as the feature vector of the input image.

TABLE 1: Model architecture of GoogLeNet used in this study.

Layer	Patch size/stride	Depth	#1 × 1	#3 × 3 reduce	#3 × 3	#5 × 5 reduce	#5 × 5	Pool proj	Output size
Conv1	7 × 7/2	1							112 × 112 × 64
Max pool1	3 × 3/2	0							56 × 56 × 64
Conv2	3 × 3/1	2		64	192				56 × 56 × 192
Max pool2	3 × 3/2	0							28 × 28 × 192
Inception-3a		2	64	96	128	16	32	32	28 × 28 × 256
Inception-3b		2	128	128	192	32	96	64	28 × 28 × 480
Max pool3	3 × 3/2	0							14 × 14 × 480
Inception-4a		2	192	96	208	16	48	64	14 × 14 × 512
Inception-4b		2	160	112	224	24	64	64	14 × 14 × 512
Inception-4c		2	128	128	256	24	64	64	14 × 14 × 512
Inception-4d		2	112	144	288	32	64	64	14 × 14 × 528
Inception-4e		2	256	160	320	32	128	128	14 × 14 × 832
Max pool4	3 × 3/2	0							7 × 7 × 832
Inception-5a		2	256	160	320	32	128	128	7 × 7 × 832
Inception-5b		2	384	192	384	48	128	128	7 × 7 × 1024
Average pool5	7 × 7/1	0							1 × 1 × 1024
Dropout (40%)		0							1 × 1 × 1024
Fc		1							1 × 1 × 2
Softmax		0							1 × 1 × 2
Classification output									1 × 1 × 2

As can be seen in Table 1, the last four layers of our proposed architecture are as follows: a dropout layer set with a probability of 40% dropout, a fully connected (fc) layer, a softmax layer, and a classification output layer. The softmax layer is a final layer of the model that uses the softmax function, and an activation function is used to derive the probability distribution of a set of numbers within an input vector. The output of a softmax activation function is a vector in which its set of values represents the probability of a class or an event occurrence. The classification output layer is set to have the same size as the number of classes in the new dataset, which was two (e.g., SZ and HC) in our case. Before training GoogLeNet, training parameters were set after empirical evaluation. In our experiments, two parameters were used to access the performance of the networks: maximum epochs and batch size.

To determine an appropriate classifier for the obtained deep GoogLeNet feature set, this study also tested four machine learning classification methods: SVM, k-nearest neighbour (KNN), decision tree (DT), and linear discriminant analysis (LDA) for identifying SZ from HC subjects. The reason for the choice of these classifiers in this study is due to their popularity, simplicity, and effectiveness in implementation. They are also very powerful and fast learning algorithms that examine all their training inputs for classification. The description of these methods is available in [26–30].

**2.5. Performance Evaluation Parameters.** To fairly assess the performance of the proposed models, we computed all standard measurement parameters including the accuracy, sensitivity, specificity, positive predictive value, false alarm rate, F1-score, and the receiver operating characteristic curve (ROC) in this study. The descriptions of the mentioned measurements are available in [28, 31–35].

### 3. Experiments and Results

**3.1. Experimental Setting.** In this study, we performed all the experiments in MATLAB (2018b) on a PC with a six-core Intel i7 processor and 32 GB of memory. The server was equipped with an NVIDIA RTX 2060 GPU with 6 GB of memory. We run the GoogLeNet model in the MATLAB deep learning toolbox for our proposed design. As stated before, the dataset includes a total of 81 subjects, where 32 cases are normal control (with 3108 trials; 3072 samples per trial; 70 channels) and 49 cases are schizophrenia patients (with 4608 trials; 3072 samples per trial; 70 channels). In this study, we used all 70 channels' data for the proposed design. Here, we provide an example how we processed the raw EEG signal data of each subject and transformed the processed dataset for implementation in the experiments in this study. For example, for Subject 1, we had a dataset: 887808 × 70 (samples × channels), and we converted this dataset to a matrix sized 70 × 3072 × 289 (channels × window length × epoch) using the transpose process. Here, we obtained epoch = 289 dividing samples by window length (887808/3072). After the average filtering, the raw signal data matrix 70 × 3072 × 289 was moved to a reduced matrix size of 70 × 256 × 289. Afterward, the 70 × 256 × 289 matrix was resized to 224 × 224 × 3 × 289 to be compatible with the deep GoogLeNet input size for subject 1.

Following a similar process, for a total of 81 subjects, the whole dataset was transformed into an image matrix with a size of 224 × 224 × 3 × 23201 (height × weight × 3 symbolize color layer × image samples). The sizes of data for SZ and HC are 224 × 224 × 3 × 13975 and 224 × 224 × 3 × 9226, respectively, which shows that the sample points of SZ and the HC groups are not equal. Thus, we divided the dataset into two groups: balanced and unbalanced datasets to test the effect of equal and unequal sizes of sample points in SZ and HC categories. A balanced dataset is one that has the same

number of observations for each class in a classification dataset. An unbalanced dataset has the different number of observations for each class. Both SZ and HC categories in this study's balanced dataset have the same number of sample points; however, both categories in the unbalanced dataset have an unequal number of sample points. As seen in Table 2, the balanced dataset consists of 9226 sample points in each category of SZ and HC (including training, validation, and testing data), and total sample points for both categories are 18,452.

In the unbalanced dataset, the SZ category has 13,975 sample points and the HC category has 9,226 sample points (including training, validation, and testing data). The total sample point size for the unbalanced data is 23,201. Please note that the unbalanced dataset is the original dataset after data preprocessing. Thus, the balanced data size is  $224 \times 224 \times 3 \times 18452$  (height  $\times$  weight  $\times$  3 symbolize color layer  $\times$  image samples), and the unbalanced data size is  $224 \times 224 \times 3 \times 23201$  (height  $\times$  weight  $\times$  3 symbolize color layer  $\times$  image samples). Then, both datasets are divided into three parts: training, validation, and testing with a ratio of 70%, 10%, and 20%, respectively. The sizes of different parts of data are given in Table 2. In this study, the training dataset was used for the learning process in the proposed model, and the validation dataset was regarded as a part of the training set to tune the model. The validation set was used for tuning the parameters of the model and also for avoiding overfitting. Generally, the validation dataset helps provide an unbiased evaluation of the model's fitness. The testing dataset was used for the performance evaluation.

### 3.2. Feature Extraction Process and Hyperparameter Setting.

This section presents the process of how the features are extracted using the GoogLeNet model for the balanced and unbalanced datasets. Figure 4 shows the feature extraction process in the GoogLeNet model for the balanced dataset and unbalanced dataset. As seen in Figure 4, the proposed GoogLeNet model yields a deep feature set with a size of  $18452 \times 2$  for the balanced dataset and  $23201 \times 2$  for the unbalanced dataset. It means that two deep features are generated, including 18,452 sample points for the balanced dataset and 23,201 for the unbalanced dataset. In both datasets, two deep features are called deep feature1 and deep feature 2.

Figures 5 and 6 present the patterns of the distribution of the obtained two deep features (deep feature 1 and deep feature 2) for the balanced and unbalanced datasets through boxplots, respectively. As can be seen in both figures, the shape of the distribution in both schizophrenia and control groups are symmetrical and there are some outliers in each diagram. In both figures, it is observed that there is a significant difference between the central value of schizophrenia and control groups in both feature sets. The boxplot figures clearly demonstrate that there is a clear and significant difference in the values of the feature set in two groups that help in the efficient classification of schizophrenia and control.

To find the best model, the hyperparameters (e.g., the number of hidden units, the number of epochs, and the batch size) of the GoogLeNet model are optimized (tuned) by the training process. We run the data through the operations of the model, compare the resulting prediction with the actual value for each data instance, evaluate the accuracy, and adjust until we find the best values. We performed an extensive number of experiments to find appropriate values for different parameters. The configuration of hyperparameters of the proposed model is provided in Table 1. The table illustrates the layer-by-layer structural details of the proposed "SchizoGoogLeNet" model.

In this study, the SVM classifier with a linear kernel was used as an optimal kernel function after testing all the kernels (e.g., linear, polynomial, and radial basis kernels) because this kernel produced a better performance compared to others. The KNN classifier used the default distance metric "Euclidean distance" for the distance measure and  $K=1$  in the model after several experimental evaluations. For DT and LDA classifiers, the parameter values are considered which have been used in MATLAB default parameter settings as there are no specific guidelines for setting the values of the parameters for these classifiers.

**3.3. Results and Discussion.** This section provides the experimental results that are achieved using 10-fold cross-validation through MATLAB. For both balanced and unbalanced datasets, we used the obtained deep feature set as an input to the softmax classifier of GoogLeNet (deep learning (DL) classifier) and also four popular machine learning (ML) classifiers (SVM, KNN, DT, and LDA), separately, to find an optimal classifier. Tables 3 and 4 present the overall performance results for our proposed approaches in terms of the sensitivity (SEN), specificity (SPE), accuracy (ACC), positive predictive value (PPV), and  $F1$ -score for the balanced and unbalanced datasets, respectively.

As can be seen in Table 3, for the balanced dataset, the SVM classifier achieves the highest performances such as ACC (98.30%), SPE (98.27%), PPV (98.27%), and  $F1$ -score (98.30%), and the LDA classifier produces the highest sensitivity value (98.53%). The lowest performances (e.g., ACC (94.28%), SEN (92.15%), SPC (96.42%), PPV (96.26%), and  $F1$ -score (94.16%)) are obtained by the GoogLeNet classifier (DL classifier). For the unbalanced dataset, Table 4 reports that among the reported classifiers, the highest classification performances are attained by the SVM classifiers, which are 98.84% of ACC, 99.02% of SEN, 98.58% of SPE, 99.06% of PPV, and 99.04% of  $F1$ -score. On the other hand, the lowest performance (e.g., ACC (95.09%), SEN (93.81%), SPC (97.02%), (PPV 97.95%), and  $F1$ -score (95.83%)) are obtained by the GoogLeNet classifier like the balanced dataset.

In order to further assess, we also computed the false alarm rate (FAR) of the proposed classification models for the balanced and unbalanced datasets as shown in Figure 7. This figure also demonstrates that the SVM classifier produces better performance (a lower FAR indicates better

TABLE 2: The sizes of different parts of data.

Data	Category	Training data	Validation data	Testing data
Balanced	HC	$224 \times 224 \times 3 \times 6458$	$224 \times 224 \times 3 \times 922$	$224 \times 224 \times 3 \times 1846$
	SZ	$224 \times 224 \times 3 \times 6458$	$224 \times 224 \times 3 \times 922$	$224 \times 224 \times 3 \times 1846$
	Total	$224 \times 224 \times 3 \times 12916$	$224 \times 224 \times 3 \times 1844$	$224 \times 224 \times 3 \times 3692$
Unbalanced	HC	$224 \times 224 \times 3 \times 6458$	$224 \times 224 \times 3 \times 922$	$224 \times 224 \times 3 \times 1846$
	SZ	$224 \times 224 \times 3 \times 9782$	$224 \times 224 \times 3 \times 1397$	$224 \times 224 \times 3 \times 2796$
	Total	$224 \times 224 \times 3 \times 16240$	$224 \times 224 \times 3 \times 2319$	$224 \times 224 \times 3 \times 4642$

SZ = schizophrenia; HC = healthy control.

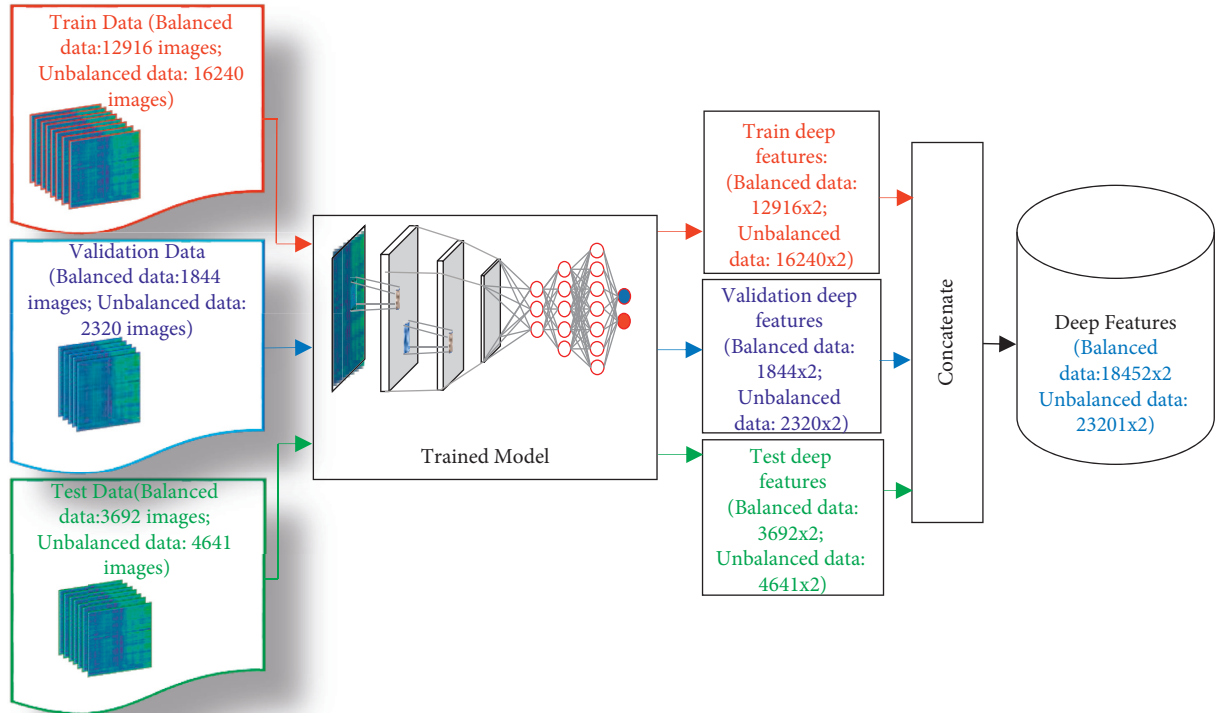


FIGURE 4: Feature extraction process in the proposed “SchizoGoogLeNet” model for the balanced and unbalanced dataset.

performance) with the obtained feature set than the GoogLeNet classifier (a higher FAR indicates lower performance). From Tables 3 and 4 and Figure 7, it is clearly apparent that the obtained feature set yields higher performance with ML classifiers than the DL classifier (e.g., GoogLeNet) for both balanced and unbalanced datasets. In both datasets, the SVM classifier is superior for the obtained deep feature set compared to other reported classifiers, and the GoogLeNet classifier with the same feature set achieved the worst performance.

Figure 8 and also Figure 7 show a comparison of the performances (in terms of ACC, SEN, SPE, and FAR) between the balanced and unbalanced datasets. The figures demonstrate that the performances of all classifiers are higher for the unbalanced dataset than those for the balanced dataset. The overall accuracy is increased by 3.75% for the unbalanced data and 4.02% for the balanced data for the ML-based classifier compared to the DL scheme. The

unbalanced dataset’s improved performance may be due to the fact that it is the original dataset, which includes all of the subjects’ data points. On the other hand, the balanced dataset was produced from the unbalanced dataset by eliminating some subjects. The performance was lower for the balanced dataset because some subjects’ data points had been eliminated. From the results, it can be considered that the obtained deep feature set with an SVM classifier is exceptional for the identification of SZ EEG signals from HC.

Figures 9(a) and 9(b) display an illustration of training and validation accuracy patterns for the deep GoogLeNet model in different iterations of the balanced and unbalanced datasets, respectively. For both datasets, the training accuracy and validation accuracy increase with the increase in the iteration numbers. It is notable that the accuracy of the training set does not deviate substantially from that of the validation set as observed in Figures 9(a) and 9(b). In the



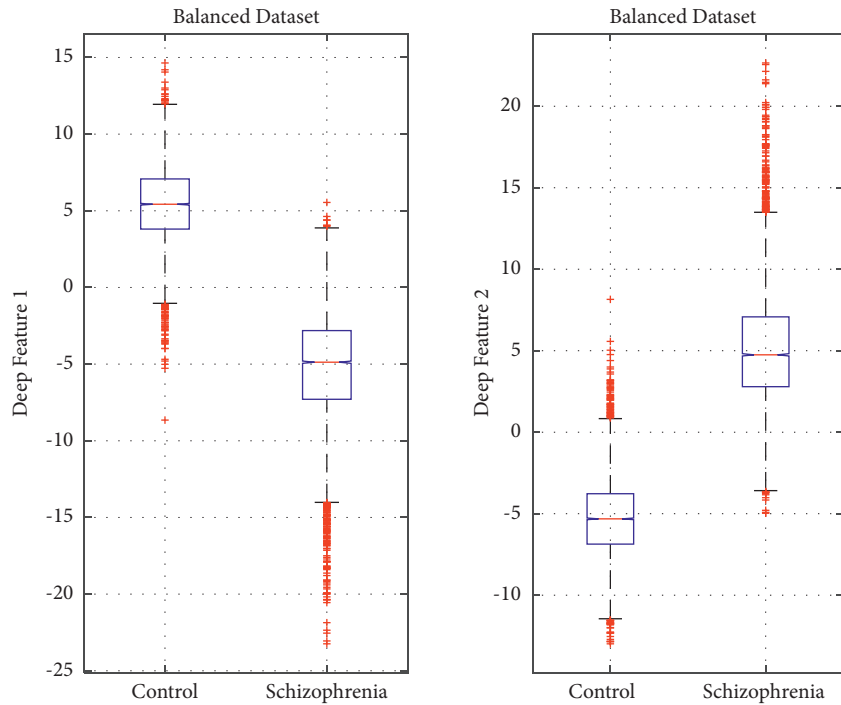


FIGURE 5: Distribution of two deep features for the balanced dataset.

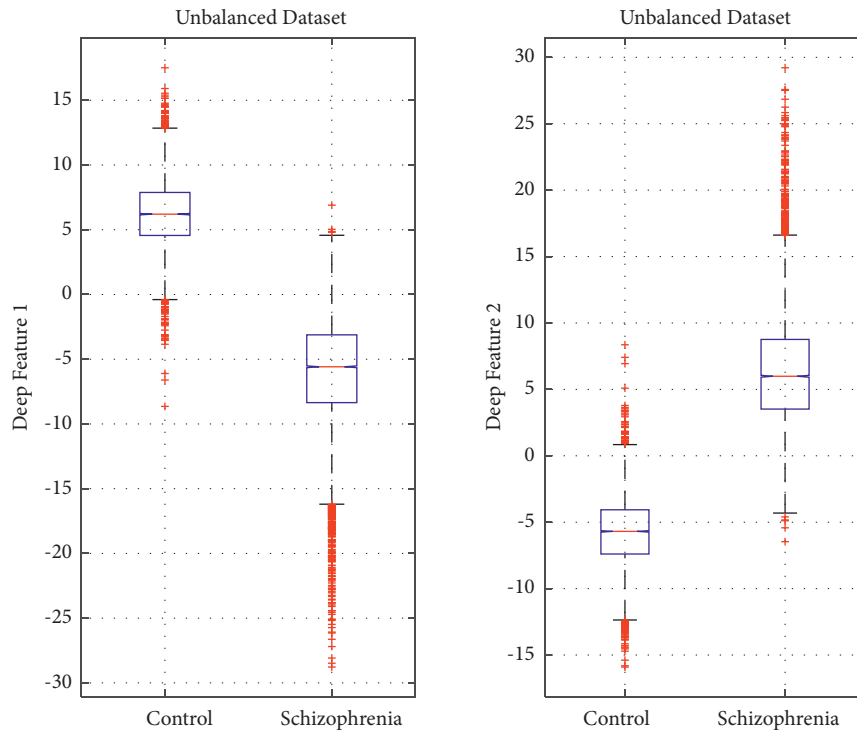


FIGURE 6: Distribution of two deep features for the unbalanced dataset.

TABLE 3: Overall performances of the proposed methods for the balanced dataset.

Classifier	Sensitivity (%)	Specificity (%)	Accuracy (%)	Positive predictive value (%)	F1-score (%)
GoogLeNet	92.15	96.42	94.28	96.26	94.16
SVM	98.33	98.27	98.30	98.27	98.30
KNN	97.29	97.27	97.28	97.27	97.28
DT	97.51	97.21	97.36	97.22	97.36
LDA	98.53	98.07	98.29	98.08	98.30



TABLE 4: Overall performances of the proposed methods for the unbalanced dataset.

Classifier	Sensitivity (%)	Specificity (%)	Accuracy (%)	Positive predictive value (%)	F1 score (%)
GoogLeNet	93.81	97.02	95.09	97.95	95.83
SVM	<b>99.02</b>	<b>98.58</b>	<b>98.84</b>	<b>99.06</b>	<b>99.04</b>
KNN	98.30	97.60	98.02	98.42	98.36
DT	98.55	97.62	98.18	98.43	98.49
LDA	99.13	98.18	98.75	98.80	98.97

Bold values represent highest performance.

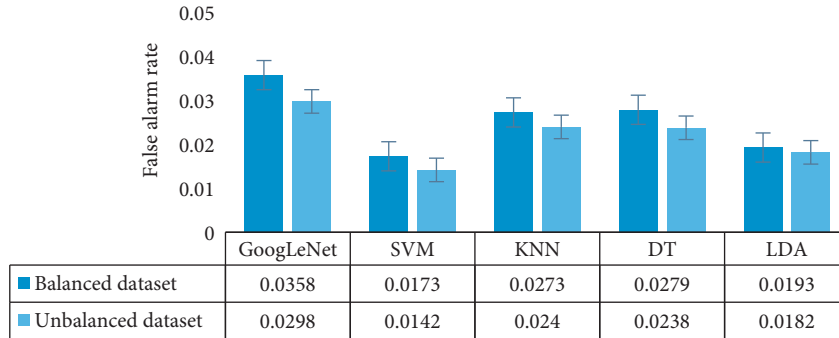


FIGURE 7: False alarm rate (FAR) for all of the reported classifiers for the balanced and unbalanced datasets.

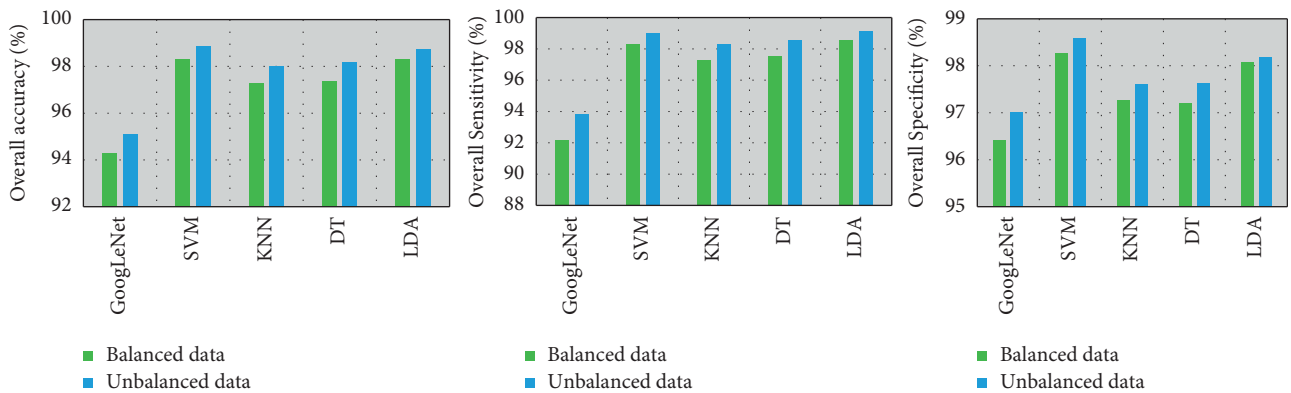


FIGURE 8: Comparison of performances between the balanced and the unbalanced datasets.

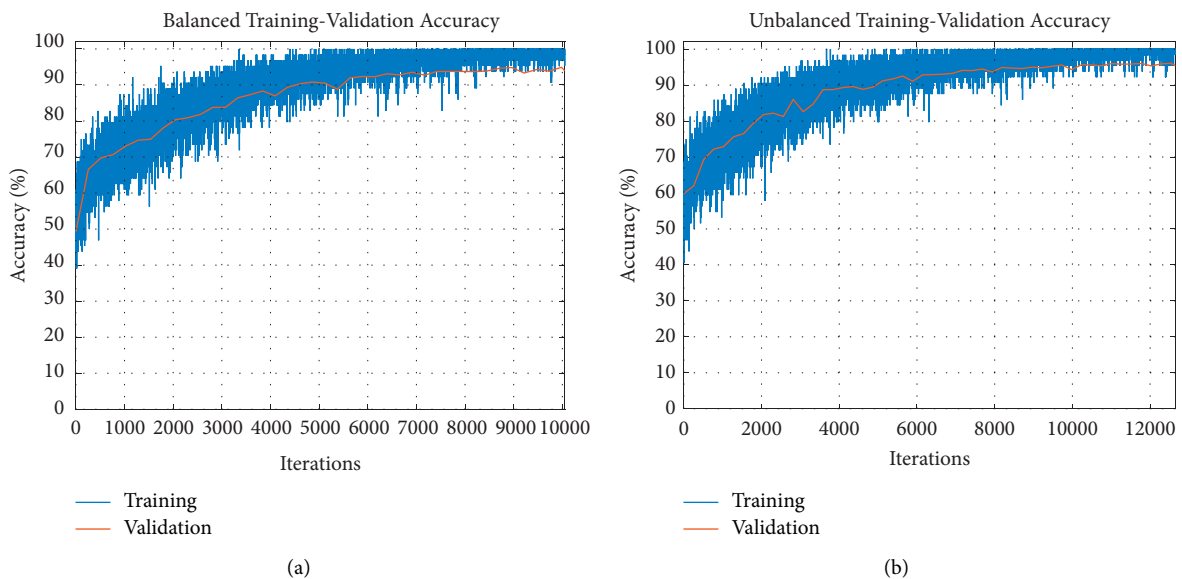


FIGURE 9: (a) Patterns of the training and validation accuracy of the GoogLeNet-based model for the balanced dataset. (b). Patterns of the training and validation accuracy of GoogLeNet-based model for the unbalanced dataset.

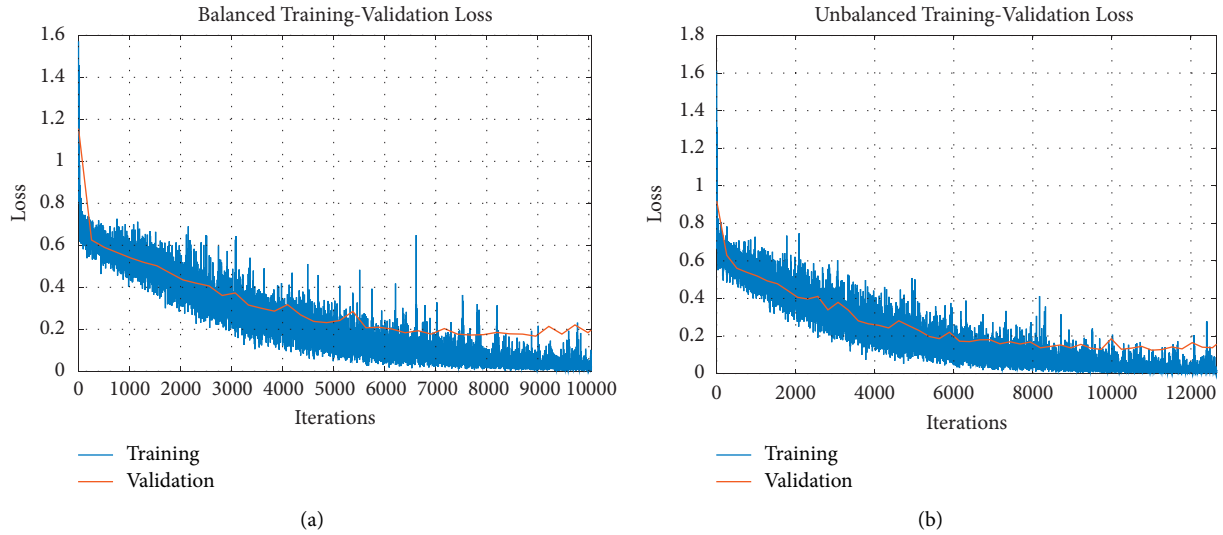


FIGURE 10: (a) Patterns of the training and validation loss information of the GoogLeNet-based model for the balanced dataset. (b). Patterns of the training and validation loss information of the GoogLeNet-based model for the unbalanced dataset.

training stage, the learning rate was set at 0.0001 and the batch number was one sample each time. The number of filters and kernel size was determined via the brute force technique.

The loss information of the training set and validation set with respect to different iterations is displayed in Figures 10(a) and 10(b) for the balanced dataset and unbalanced dataset, respectively. For both datasets, it is observed that the training loss and the validation loss decrease with the increase in the iteration numbers. The performance of the training set does not significantly diverge from that of the validation set, as shown in Figures 10(a) and 10(b).

To further assess the effectiveness of the GoogLeNet-based model, the ROC curves are drawn for different SZ detection models, where the input data were the deep feature set shown in Figures 11(a) and 11(b) for the balanced and unbalanced datasets, respectively. The corresponding performance measurements in every condition are shown in Table 5. Table 5 reports the area values under the ROC curve (AUC) for the reported classifiers. The AUC is the value of the area under the ROC curve that belongs to a value between 0 and 1 (a larger area indicates a better performance of the classifier). As can be seen in Table 5, the highest AUC is obtained by the SVM classifier, which is 0.9984 (close to 1) for the unbalanced dataset and 0.9973 for the balanced dataset. The KNN model produces the lowest AUC for both the balanced (0.9728) and the unbalanced (0.9795) datasets. Like the previous results, the results also indicate that the SVM classifier with the obtained deep feature set works better than other reported classifiers.

*3.4. Comparative Analysis Report for Our Proposed Method with Existing State-of-the-Art Methods.* A comparison of the prior EEG-based techniques used for SZ detection with our proposed model has been provided in Table 6. Until now, we found seven articles [2, 28, 35–39] in the literature for the

same database that we have used in this study. Table 6 shows the performance comparison of the proposed method with these published methods [2, 28, 35–39]. Kahre et al. [35] reported a method based on empirical wavelet transformation and SVM for the detection of SZ from EEG signals. Their method achieved an ACC of 88.70%, SEN of 91.13%, and SPE of 89.29. In [2], Siuly et al. introduced empirical mode decomposition (EMD)-based features with an ensemble bagged tree (EBT) for the detection of SZ using EEG signals. The ACC, SEN, and SPE scores of their method were 89.59%, 89.76%, and 89.32%, respectively. Guo et al. [38] reported a random forest (RF)-based machine learning algorithm for identifying schizophrenia patients from healthy control subjects using EEG signal data. In the designed plan, the author considered a number of features such as gender, age, education, and event-related potential (ERP) and the combination of the features.

RF yielded an accuracy of 81.1%. Khare et al. [36] introduced an automatic approach based on flexible tunable Q wavelet transform (F-TQWT) and a flexible least square support vector machine (F-LSSVM) classifier for the detection of SZ from EEG signals. The authors used the “Fisher score” method for the selection of the most discriminant channels. Their proposed method generated 91.39% accuracy, 92.65% sensitivity, and 93.22% specificity. In [37], Guo et al. proposed a scheme based on convolutional neural networks (CNNs) to characterize the difference in the distributed structure of data for identifying SZ from EEG. Their method achieved an accuracy of 92%. Khare et al. [38] designed a model involving a robust variational mode decomposition (RVMD) and an optimized extreme learning machine (OELM) algorithm. The experiment results reveal that the third mode’s chaotic features are responsible for generating the best performance (overall ACC of 92.30%). In [39], a time-frequency analysis-based convolutional neural network (CNN) model was proposed for identifying SZ from EEG signals by Khare and Bajaj. The authors used

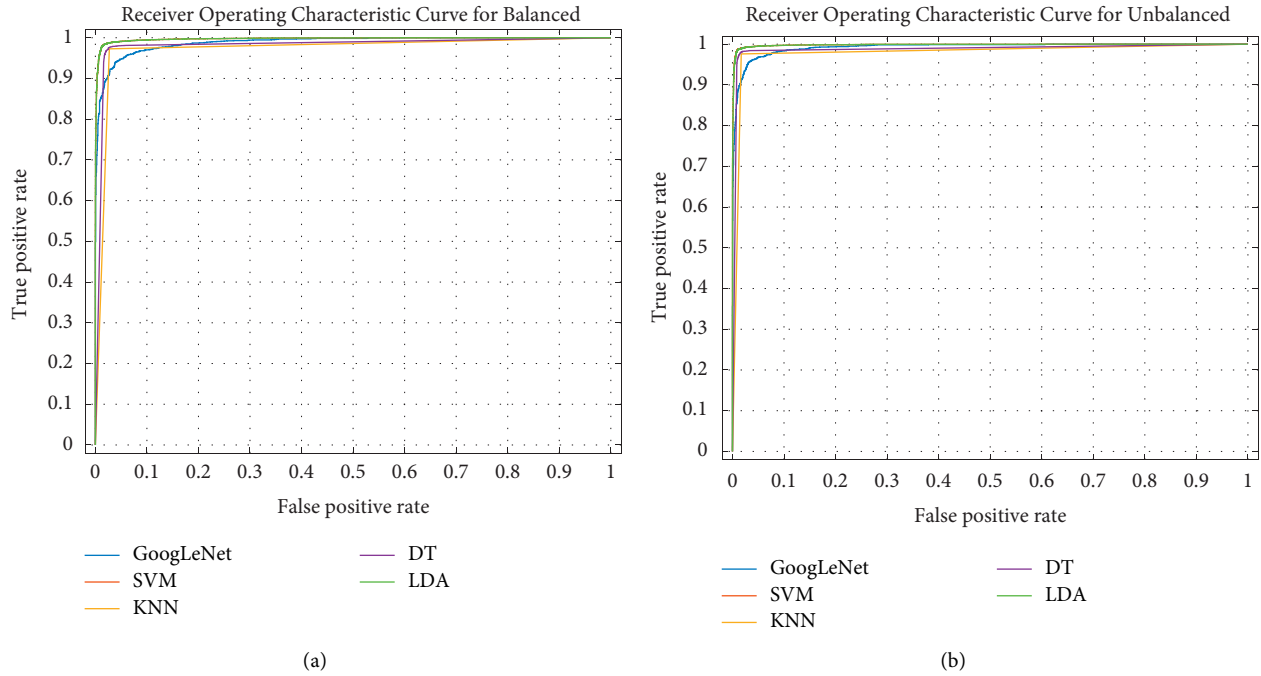


FIGURE 11: (a) The ROC curve for different classification models with the obtained deep feature set for the balanced dataset. (b) The ROC curve for different classification models with the obtained deep feature set for the unbalanced dataset.

TABLE 5: AUC values for the proposed SZ detection models.

Models	Balanced dataset	Unbalanced dataset
	Area under curve (AUC) values	Area under curve (AUC) values
GoogLeNet	0.9884	0.9923
SVM	0.9973	<b>0.9984</b>
KNN	0.9728	0.9795
DT	0.9815	0.9874
LDA	0.9973	0.9984

TABLE 6: The comparison of the proposed method with other methods for the same database.

Authors	Methods	ACC (%)	SEN (%)	SPE (%)
Khare et al. [35]	Empirical wavelet transformation with SVM	88.70	91.13	89.29
Siuly et al. [2]	EMD-based features with EBT	89.59	89.76	89.32
Guo et al. [38]	ERP features with RF	81.10	NA	NA
Khare and Bajaj [37]	F-TQWT-based scheme	91.39	92.65	93.22
Guo et al. [38]	Electrical marker with CNN	92.00	NA	NA
Khare et al. [38]	RVMD-based OELM method	92.93	97.15	91.06
Khare and Bajaj [39]	SPWVD-based TFR and CNN model	93.36	94.25	92.03
Proposed method	GoogLeNet-based deep features with an SVM model	<b>98.84</b>	<b>99.02</b>	<b>98.58</b>

\*NA = not available.

Bold values represent the highest performance.

continuous wavelet transform, short-time Fourier transform, and smoothed pseudo Wigner–Ville distribution (SPWVD) techniques to obtain scalogram, spectrogram, and SPWVD-based time-frequency representation (TFR) plots for SZ detection. Their method achieved an overall accuracy of 93.36% using the SPWVD-based TFR and CNN model.

It is apparent from Table 6 that our proposed model yielded the highest performance scores with an accuracy of

98.84%, sensitivity of 99.02%, and specificity of 98.58%, compared to performance scores of the other existing methods. The achieved accuracy improvement of our proposed model is 17.74% better than the accuracy score of Zhang [28] and 10.14% better than the accuracy score of Khare et al. [35]. In the end, it can be concluded based on the experimental results that deep GoogleNet features of EEG signals with an SVM classifier could serve as an applicable

measurement to correctly discriminate between schizophrenics and HC subjects. [40].

#### 4. Concluding Remarks

In this study, a GoogLeNet-based feature extraction scheme, called “SchizoGoogLeNet” is developed to efficiently identify SZ patients from HC subjects using EEG signal data. The proposed GoogLeNet model automatically extracted important hidden features which are advantageous for large-scale data. The obtained deep feature set was verified by the GoogLeNet classifier (DL classifier) and also four popular ML classifiers (e.g., SVM, KNN, DT, and LDA), separately. The performance of the proposed framework was evaluated on the benchmark SZ EEG database from Kaggle through extensive experimental evaluation. To check the effect of equal and unequal sample points in SZ and HC groups, we divided the dataset into two groups: balanced (the same number of sample points in SZ and HC) and unbalanced dataset (unequal sample points in SZ and HC) (original dataset). The experimental results show that the unbalanced set produces better performance compared to the balanced dataset. Among the reported classifiers, the SVM classifier with the obtained deep feature set yielded the highest performance (e.g., ACC 98.84%, SEN 99.02%, SPE 98.58%, PPV 99.06%, and F1-score 99.04%), while the lowest performances were obtained by the GoogLeNet classifier (e.g., ACC 95.09%, SEN 93.81%, SPC 97.02%, PPV 97.95%, and F1-score 95.83%). Moreover, our proposed model outperforms the existing methods. The findings of this study indicate that the obtained deep GoogLeNet features perform better with the SVM classifier in the SZ detection than the DL classifier (GoogLeNet classifier).

This study has some limitations, such as the fact that the study only considers two-class classification problems (SZ versus HC), while we intend to expand the application of the suggested approach to multiclass scenarios soon. Another flaw is that this study used a small SZ-based EEG dataset (81 total subjects: 49 schizophrenia (SZ) patients and 32 healthy control (HC) people). In the near future, we will broaden our method’s application to include huge clinical datasets.

#### Data Availability

The data can be obtained from the Kaggle website: <https://www.kaggle.com/broach/button-tone-sz>.

#### Conflicts of Interest

The authors declare that they have no conflicts of interest.

#### Acknowledgments

The authors would like to thank Kaggle EEG data collection team for the dataset: “EEG data from basic sensory task in Schizophrenia.” Funding for the data collection was supported by National Institute of Mental Health (Project Number: 5R01MH058262-16).

#### References

- [1] O. Baser, L. Xie, J. Pesa, and M. Durkin, “Healthcare utilization and costs of veterans health administration patients with schizophrenia treated with paliperidone palmitate long-acting injection or oral atypical antipsychotics,” *Journal of Medical Economics*, vol. 18, no. 5, pp. 357–365, 2015.
- [2] S. Siuly, S. K. Khare, V. Bajaj, H. Wang, and Y. Zhang, “A computerized method for automatic detection of schizophrenia using EEG signals,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 11, pp. 2390–2400, Nov 2020.
- [3] S. L. Oh, J. Vicnesh, E. J. Ciaccio, R. Yuvaraj, and U. R. Acharya, “Deep convolutional neural network model for automated diagnosis of schizophrenia using EEG signals,” *Applied Sciences*, vol. 9, no. 14, p. 2870, 2019.
- [4] Healthdirect, *Schizophrenia*, *Healthdirect*, <https://www.healthdirect.gov.au/schizophrenia> viewed 4 February 2020, 2018.
- [5] World Health Organization (Who), *Schizophrenia*, *WHO*, <https://www.who.int/news-room/fact-sheets/detail/schizophrenia> viewed 5 February 2020, 2019.
- [6] A. H. Mokdad, M. H. Forouzanfar, F. Daoud et al. “Global burden of diseases, injuries, and risk factors for young people’s health during 1990–2013: a systematic analysis for the global burden of disease study 2013,” *Lancet (North American Edition)*, vol. 387, no. 10036, pp. 2383–2401, 2016.
- [7] V. Jahmunah, S. Lih Oh, V. Rajinikanth et al., “Automated detection of schizophrenia using nonlinear signal processing methods,” *Artificial Intelligence in Medicine*, vol. 100, Article ID 101698, 2019.
- [8] C. Devia, R. Mayol-Troncoso, J. Parrini et al., “EEG classification during scene free-viewing for schizophrenia detection,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 6, pp. 1193–1199, 2019.
- [9] S. Siuly, O. F. Alçin, E. Kabir et al., “A new framework for automatic detection of patients with mild cognitive impairment using resting-state EEG signals,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 9, pp. 1966–1976, 2020.
- [10] M. Sabeti, R. Boostani, and T. Zoughi, “Using genetic programming to select the informative EEG-based features to distinguish schizophrenic patients,” *Neural Network World*, vol. 22, no. 1, pp. 3–20, 2012.
- [11] J. Ramos, L. F. Cerdán, M. A. Guevara, C. Amezcua, and A. Sanz, “Abnormal eeg patterns in treatment-resistant schizophrenic patients,” *International Journal of Neuroscience*, vol. 109, no. 1–2, pp. 47–59, 2001.
- [12] A. I. Kaplan, S. V. Borisov, and V. A. Zheligovskii, “Classification of the adolescent EEG by the spectral and segmental characteristics for normal,” *Zhurnal Vysshei Nervnoi Deiatelnosti Imeni I. P. Pavlova*, vol. 55, no. 4, pp. 478–486, 2005.
- [13] R. Buettner, D. Beil, S. Scholtz, and 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.
- [14] S. A. Akar, S. Kara, F. Latifoğlu, and V. Bilgiç, “Analysis of the complexity measures in the EEG of schizophrenia patients,” *International Journal of Neural Systems*, vol. 26, no. 02, Article ID 1650008, 2016.
- [15] S. A. Akar, S. Kara, F. Latifoğlu, and V. Bilgiç, “Wavelet-welch methodology for analysis of EEG signals of schizophrenia patients,” in *Proceedings of the 2012 Cairo International*

- Biomedical Engineering Conference (CIBEC) Cairo*, pp. 6–9, Egypt, December 20–21, 2012.
- [16] F. Li, J. Wang, Y. Liao et al., “Differentiation of schizophrenia by combining the spatial EEG brain network patterns of rest and task P300,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 4, pp. 594–602, 2019.
- [17] C.-R. Phang, F. Noman, H. Hussain, C. M. Ting, and H. Ombao, “A multi-domain connectome convolutional neural network for identifying schizophrenia from EEG connectivity patterns,” *IEEE J Biomed Health Inform*, vol. 24, no. 5, pp. 1333–1343, 2020.
- [18] S. L. Oh, J. Vicnesh, E. J. Ciaccio, R. Yuvaraj, and U. R. Acharya, “Deep convolutional neural network model for automated diagnosis of schizophrenia using EEG signals,” *Applied Sciences*, vol. 9, no. 14, p. 2870, 2019.
- [19] C. Szegedy, W. Liu, and Y. Jia, “Going deeper with convolutions,” in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9, 2015.
- [20] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5455–5516, 2020.
- [21] kaggle website, “Eeg data from basic sensory task in schizophrenia,” 2021, <https://www.kaggle.com/broach/button-tone-sz>.
- [22] J. M. Ford, V. A. Palzes, B. J. Roach, and D. H. Mathalon, “Did I do that? Abnormal predictive processes in schizophrenia when button pressing to deliver a tone,” *Schizophrenia Bulletin*, vol. 40, no. 4, pp. 804–812, 2014.
- [23] S. K. Khare, V. Bajaj, and U. R. Acharya, “SPWVD-CNN for automated detection of schizophrenia patients using EEG signals,” *IEEE Transactions on Instrumentation and Measurement*, vol. 70, p. 9, 2021.
- [24] Novel Egfr inhibitors attenuate cardiac hypertrophy, <http://pafit-phil.com/technically-a-feature-represents-a-distinguishing-property-a-recognizable-measurement-and/>, 2021.
- [25] Y. N. Dauphin, H. De Vries, and Y. Bengio, “Equilibrated adaptive learning rates for non-convex optimization,” *Advances in Neural Information Processing Systems*, pp. 1–10, 2015.
- [26] Y. L. Siuly and Y. Zhang, “EEG signal analysis and classification: techniques and applications,” *Health Information Science, Springer Nature*, US (ISBN 978-3-319-47653-7), Dec 2016.
- [27] M. N. A. Tawhid, S. Siuly, H. Wang, F. Whittaker, K. Wang, and Y. Zhang, “A spectrogram image based intelligent technique for automatic detection of autism spectrum disorder from EEG,” *PLoS One*, vol. 16, no. 6, Article ID e0253094, 2021.
- [28] M. T. Sadiq, H. Akbari, S. Siuly, A. Yousaf, and A. U. Rehman, “A novel computer-aided diagnosis framework for EEG-based identification of neural diseases,” *Computers in Biology and Medicine*, vol. 138, Article ID 104922, 2021.
- [29] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag, New York Inc, 2000.
- [30] T. Cover and P. Hart, “Nearest neighbor pattern classification,” *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967.
- [31] Y. Guo, G.-Q. Du, W.-Q. Shen, C. Du, P. N. He, and S. Siuly, “Automatic myocardial infarction detection in contrast echocardiography based on polar residual network,” *Computer Methods and Programs in Biomedicine*, vol. 198, Article ID 105791, 2021.
- [32] S. Taran, V. Bajaj, and S. Siuly, “An optimum allocation sampling based feature extraction scheme for distinguishing seizure and seizure-free EEG signals,” *Health Information Science and Systems*, vol. 5, no. 1, p. 7, 2017.
- [33] M. A. Alvi, S. Siuly, H. Wang, K. Wang, and F. Whittaker, “A deep learning based framework for diagnosis of mild cognitive impairment,” *Knowledge based System*, vol. 2022, 2022 (in press).
- [34] L. Farsi, S. Siuly, E. Kabir, and H. Wang, “Classification of alcoholic EEG signals using a deep learning method,” *IEEE Sensors Journal*, vol. 21, no. 3, pp. 3552–3560, 2021.
- [35] S. K. Khare, V. Bajaj, S. Siuly, and G. R. Sinha, “Classification of schizophrenia patients through empirical wavelet transformation using electroencephalogram signals,” in *Book: Modelling and Analysis of Active Biopotential Signals in Healthcare*, vol. 1, Publisher: IOP Science, 2020.
- [36] L. Zhang, “EEG signals classification using machine learning for the identification and diagnosis of schizophrenia,” in *Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4521–4524, EMBC, 2019.
- [37] S. K. Khare and V. Bajaj, “A self-learned decomposition and classification model for schizophrenia diagnosis,” *Computer Methods and Programs in Biomedicine*, vol. 211, Article ID 106450, 2021.
- [38] Z. Guo, L. Wu, Y. Li, and B. Li, “Deep neural network classification of EEG data in schizophrenia,” in *Proceedings of the 2021 IEEE 10th Data Driven Control and Learning Systems Conference, DDCLS*, pp. 1322–1327, 2021.
- [39] S. K. Khare and V. Bajaj, “A hybrid decision support system for automatic detection of Schizophrenia using EEG signals,” *Computers in Biology and Medicine*, vol. 141, Article ID 105028, February 2022.
- [40] A. Kawala-Sterniuk, M. Podpora, M. Pelc et al., “Comparison of smoothing filters in analysis of EEG data for the medical diagnostics purposes,” *Sensors*, vol. 20, no. 3, p. 807, 2020.