

## Research Article

# Diagnosis of Alzheimer's Disease by Time-Dependent Power Spectrum Descriptors and Convolutional Neural Network Using EEG Signal

Morteza Amini <sup>1</sup>, Mir Mohsen Pedram <sup>2,3</sup>, AliReza Moradi <sup>4,5</sup>  
and Mahshad Ouchani <sup>6</sup>

<sup>1</sup>Department of Cognitive Modeling, Institute for Cognitive Science Studies, Shahid Beheshti University, Tehran, Iran

<sup>2</sup>Department of Electrical and Computer Engineering, Faculty of Engineering, Kharazmi University, Tehran, Iran

<sup>3</sup>Department of Cognitive Modeling, Institute for Cognitive Science Studies, Tehran, Iran

<sup>4</sup>Department of Clinical Psychology, Faculty of Psychology and Educational Science, Kharazmi University, Tehran, Iran

<sup>5</sup>Department of Cognitive Psychology, Institute for Cognitive Science Studies, Tehran, Iran

<sup>6</sup>Institute for Cognitive and Brain Science, Shahid Beheshti University, Tehran, Iran

Correspondence should be addressed to Mir Mohsen Pedram; pedram@khu.ac.ir

Received 2 March 2021; Revised 26 March 2021; Accepted 7 April 2021; Published 24 April 2021

Academic Editor: Venkatesan Rajinikanth

Copyright © 2021 Morteza Amini 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.

Using strategies that obtain biomarkers where early symptoms coincide, the early detection of Alzheimer's disease and its complications is essential. Electroencephalogram is a technology that allows thousands of neurons with equal spatial orientation of the duration of cerebral cortex electrical activity to be registered by postsynaptic potential. Therefore, in this paper, the time-dependent power spectrum descriptors are used to diagnose the electroencephalogram signal function from three groups: mild cognitive impairment, Alzheimer's disease, and healthy control test samples. The final feature used in three modes of traditional classification methods is recorded:  $k$ -nearest neighbors, support vector machine, linear discriminant analysis approaches, and documented results. Finally, for Alzheimer's disease patient classification, the convolutional neural network architecture is presented. The results are indicated using output assessment. For the convolutional neural network approach, the accurate meaning of accuracy is 82.3%. 85% of mild cognitive impairment cases are accurately detected in-depth, but 89.1% of the Alzheimer's disease and 75% of the healthy population are correctly diagnosed. The presented convolutional neural network outperforms other approaches because performance and the  $k$ -nearest neighbors' approach is the next target. The linear discriminant analysis and support vector machine were at the low area under the curve values.

## 1. Introduction

The term "dementia" refers to many neurodegenerative illnesses caused by neuronal failure and death that interrupt cognitive and behavioral activities. The most prevalent of the several types of dementia is Alzheimer's disease (AD), with about 70% of worldwide dementia cases. It affects the individual over 65 years, and the rate of occurrence increases exponentially at the age of 65 years [1–3]. To date, AD has not been resolved by palliative therapies, which have been

temporarily slow to deteriorate in patients and caregiver living [4]. Today, only postmortem diagnosis of definitive AD is possible after examining the structural brain injury that is typical of the condition. Accuracies of up to 90% have usually been recorded for modern testing procedures, such as neurological assessments and medical history. The National Institute on Aging and Alzheimer's Association has established the existing standards of clinical diagnosis of AD, and the Alzheimer's Association has established them [5]. These standards are an advancement in the previous guidelines,

which had been developed in 1984 by the National Institute of Neurological And Communicative Diseases and Stroke/Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) [6]. It is part of the NINCDS-ADRDA guideline. These revised suggestions require neuroimaging and the use of biomarkers and cerebrospinal fluid to diagnose AD for symptomatic people [5].

A guideline for diagnosing and monitoring AD [7] was established by the European Federation of Neurological Societies. The Mini-Mental State Assessment [8, 9] is the most used AD diagnosis method to test cognitive ability. The revised Montreal Cognitive Assessment [10] is commonly used in therapeutic functional applications and the revised Addenbrooke Cognitive Evaluation [11]. Another example of neurological testing is the Extreme Cognitive Disorder, Alzheimer's Cognitive Disease Evaluation Scale, Neuropsychological Assessment Battery, and Serious Impairment Battery. The Trail Making Test [12] and the clock drawing test [13], by contrast, focus not only on testing thinking abilities but also on concentrating and administrative work. In comparison, the visual learning test and the Rey Auditory Fluency Assessment assess all patient support practice skills [14]. In some instances, AD is also associated with other diseases that cause dementia as brain vascular injury, Lewy body disease, and Parkinson's disease [15]. The early diagnosis of AD and these problems is improved using methods that gain biomarkers as early signs overlap [16–19]. Electroencephalogram (EEG) is a technology that enables the recording by postsynaptic potentials of a thousand neurons of identical spatial orientation of the time of cerebral cortex electrical activity. Scalp-positioned electrodes measure the electrical potentials. EEG's spatial resolution refers to the number and location of electrodes on the scalp. The most used configuration is the international 10-20 system, which consists of 21 electrodes; the 10-20 system is often used for higher density versions, for instance, 10-10 and 10-5, usually 64 and 128 electrodes, and the Maudsley [20] and geodesic positions [21] alternating layouts. Reliable therapeutic methods have been shown in recent years for the diagnosis and analysis of disorders and cortical conditions like the Huntington syndrome, the autism spectrum disorder [22], epilepsy and seizure [23], brain ischemia [24], frontotemporal dementia [25], and Parkinson's dementia [26]. Furthermore, EEG evaluations were carried out on the comparative diagnosis of AD and other dementia-contributing diseases such as brain vascular injuries [27, 28] and Lewy disease [29, 30]. Theta ( $\theta$ ) 4–8 Hz, delta ( $\delta$ ) 0.1–4 Hz, beta ( $\beta$ ) 12–30 Hz, alpha ( $\alpha$ ) 8–12 Hz, and gamma ( $\gamma$ ) > 30 Hz are typically divided between 5 major frequency bands in the analytics. Also, more divisions into these bands (high alpha, low alpha, low beta) are considered, but the subband frequency limits are not uniform in all studies. The different data on brain function and synchronization are given in each frequency band [31–33]. There has been a comprehensive study of the possible use of electroencephalography to diagnose dementia and AD [34]. EEG is a high time resolution noninvasive, comparatively inexpensive, and potentially mobile technology (about milliseconds). It was studied primarily as an AD diagnostic tool when comparing EEG records in AD patients

with control subjects (healthy individuals) [35, 36]. AD is generally known to decrease the complexity of EEG signals and synchronous change in EEG.

These improvements have been used as discriminatory features for AD diagnosis in EEG recordings. Several methods of assessing the complexity of EEG signals have been established. The connection factor and the first positive exponent of Lyapunov have been used frequently [37–42]. EEG signals from AD patients have been shown to show lower (lower complexity) values of certain tests than signals from age-matched control subjects. Other information-theoretical methods, in specific entropy-based approaches, have appeared as theoretically useful EEG indicators for AD: epoch-based entropy [43, 44], sample entropy [45], Tsallis entropy [46], approximate entropy [47, 48], multiscale entropy [49], and complexity of Lempel-Ziv [50]. These approaches relate the strength of a signal to unpredictability: irregular signals are more complicated than regular ones because they are erratic. Different detection algorithms have been suggested in previous studies for epileptiform EEG data [51]. Current seizure detection methods use hand-built feature extraction techniques from EEG signals [52], including frequency domain, time domain, time-frequency domain, and nonlinear evaluation of signals [53, 54]. The features selected must be listed after the feature extraction to identify various EEG signals using all forms of classifiers [55]. Hamad et al. employed a differential wavelet transformation procedure to obtain the feature collection, then trained the radial reference algorithm with the support vector machine (SVM), demonstrating an epilepsy diagnosis with the suggested SVM gray wolf optimizer [56]. For the refinement of the SVM parameters based on genetic algorithms and particle swarm optimization, Subasi et al. developed a hybrid model. The proposed SVM hybrid model shows that neuroscientists use EEG as an essential method for diagnosing epileptic seizures [57]. However, the manual function selection criteria are not eradicated by these strategies [58]. The feature extraction is an important stage in classification determination, as it determines its specificity in large part. A system for classifying without the removal of complicated properties was suggested. Furthermore, recent advancements in deep learning have shown a new way of coping with this problem. Deep learning has in recent years reached the recognized form of computer vision and machine learning and has demonstrated that almost all human and superhuman functions such as pattern recognition and sequence learning perform numerous functions [59], among other things. Feature extraction before classification is more advantageous than entering raw EEG samples directly into the classifier. Nevertheless, several recent research types have not performed feature extraction, but instead, raw EEG signals were used for the deep learning model [60, 61].

In this paper, the time-dependent power spectrum descriptor (TD-PSD) method is utilized for feature extraction of the EEG signal from three categories of MCI, AD, and HC sample test. The final feature with labeling is used in three types of traditional classification methods, including  $k$ -nearest neighboring (KNN), SVM, and linear discriminant analysis (LDA) approaches, and the results are recorded.

TABLE 1: Summary of machine learning method for brain disease diagnosis with EEG signal.

Author	Year	Disease	Feature extraction	Classification	Results
Xin et al. [80]	2021	Epilepsy	Dimensionality reduction principal component analysis (PCA)	Convolution SVM	The method's accuracy, sensitivity, and specificity reach up to 99.56%, 99.72%, and 99.52%, respectively
Aliyu and Lim [81]	2021	Epilepsy	Discrete wavelet transforms (DWT)	LSTM network	Reduction of the number of LSTM trainable parameters needed to achieve extreme accuracy
Tuncer [82]	2021	Epileptic seizure	Nonlinear textural feature extraction (Hamsi hash)	$k$ -nearest neighborhood	This model has an accuracy in the EEG dataset of 99.20% for five classes and has 100.0% accuracy in other conditions
Cicalese et al. [83]	2020	AD	Pearson correlation coefficient-based feature selection (PCCFS)	LDA	The EEG-fNIRS feature set combination was expected to obtain greater precision (79.31%) by combining its supplementary properties as compared with the EEG (65.52%) or fNIRS alone (58.62%). Moreover, AD development is associated with the right and left parietal lobe
Ferri et al. [84]	2020	AD	rsEEG + sMRI	Low-resolution brain electromagnetic tomography (LORETA)	Classification accuracy of 80%, 85%, and 89% using rsEEG, sMRI, and rsEEG + sMRI features, respectively, discriminates against them
Trambaiolli et al. [85]	2017	AD	Feature selection (FS)	SVM classifier	Since eliminating $88.76 \pm 1.12\%$ of the initial elements, the filtered subset evaluator technique obtained the highest efficiency gain, both on a per-patient basis (91.18% accuracy) and on a per-epoch basis ( $85.29 \pm 21.62\%$ )
Nobukawa et al. [86]	2020	AD	Functional connectivity	SVM	A novel interpretation of neural network functions in healthy brains and unhealthy disorders can be provided by applying a mixture of both machine learning approaches to neurophysiological evidence
Kulkarni and Bairagi [87]	2017	AD	Extracting salient features that are spectral-, wavelet-, and complexity-based	SVM	The increased performance in AD diagnosis
Vecchio et al. [88]	2020	AD	—	SVM	A low-cost and noninvasive process uses readily available tools that, when integrated, achieve high sensitivity/specificity and optimum individual classification accuracy (0.97 of AUC)

Finally, an architecture of convolutional neural network (CNN) for AD patients' classification is provided. The results are indicated using performance analysis.

## 2. Literature Review

EEG signals' complex and nonlinear nature implies creating new methods to study machine and signal processing [62, 63]. Recent progress has been made to enhance high-level abstractive methods for the automated removal of complex data features in the field of deep learning methodologies [64–66]. In the last years, these deep learning methods are usually used in image processing, natural language processing [67–70], speech processing, and video games [71]. The biomedical area has also been identified with these methods [72–74]. Acharya et al. [60] suggested a deep, 13-layer neural CNN that distinguishes normal, preictal, and EEG signals of seizure. In the study, 300 EEG signals were used for registering a classification rate of 88.67%. The same group proposed

a deep neural network approach for an innovative EEG-based depression screening method [60]. This investigation's outcomes are reported in 15 regular and 15 depressive patients, 93.5% (left hemisphere) and 96.0% (right hemisphere). Oh et al. [75] suggested that further studies used EEG signals to diagnose Parkinson's disease. A 13-layer CNN model of 20 healthy and 20 Parkinson patients reached an accuracy of 88.25%. Hong et al. [76] propose a mathematical model employing Long Short-Term Memory (LSTM), a recurrent neural network (RNN) that predicts the mild cognitive impairment of AD. The data is taken in image form in this research, and the preprocessing is done by skull strip, normalization, registration, smoothing, and segmentation. The training is carried out by feeding sequential data with time steps to the model after preprocessing, and the model projects the state of the next six months. During model testing, when the feature data for the 18th and 24th months is presented, it forecasts the state of the subject for the 30th month. Similarly, Aghili et al. [77] suggest an RNN approach

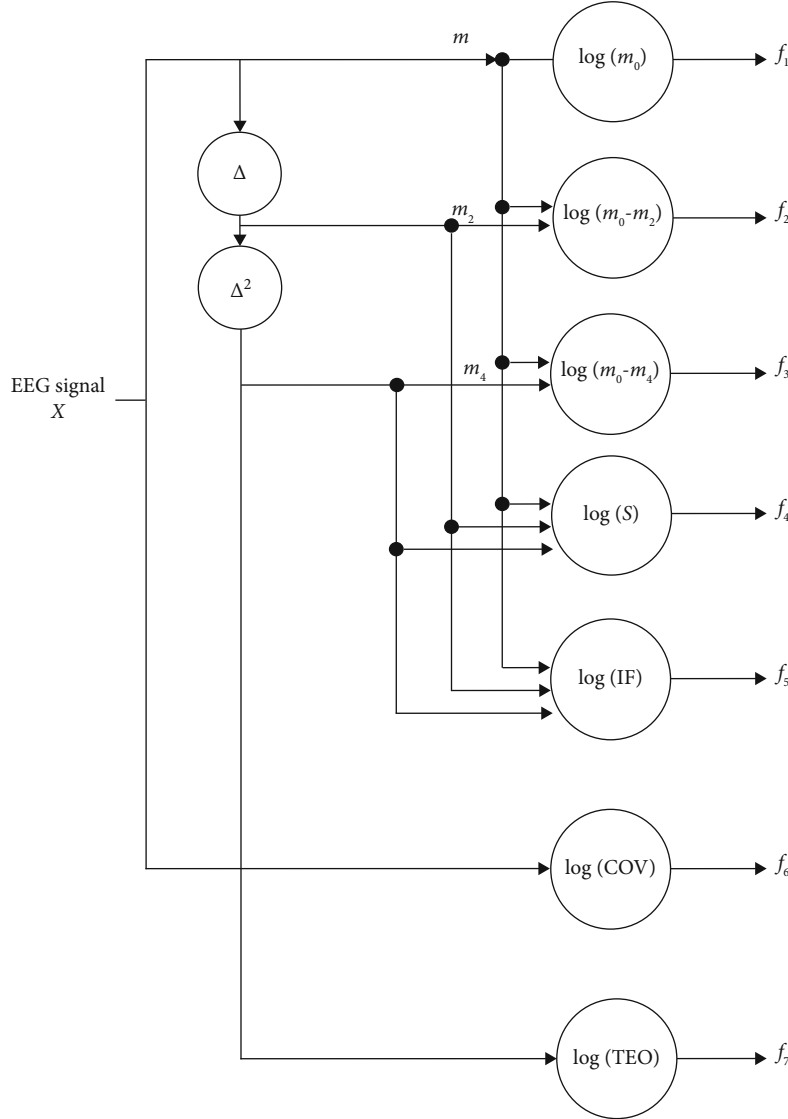


FIGURE 1: The block diagram of the TD-PSD feature extraction method for EEG signals.

to evaluate longitudinal data to differentiate between stable people against AD individuals. The input is preprocessed, and the function is normalized. It is fed into LSTM and gated recurring units after preprocessing the data. In the model of LSTM and gated regular units, each subject's time point data is provided to the corresponding cell along with its final diagnosis mark to learn the pattern of data transition. The effect models are contrasted with the effects of nonrecurrent networks, i.e., multilayer perceptron (MLP), for all data arrangements. For each patient, the data is fed into the MLP once. There are many trainable parameters in the LSTM models that need to be substantially trained for sequential data and are vulnerable to overfitting the training data. The purpose of unsupervised feature learning is to define AD using the principle of unsupervised feature learning. An approach used sparse filtering to learn the expressive characteristics of brain images [78]. The SoftMax regression is trained to classify the circumstances. In the first step, there are three phases: sparse filtering is trained, and its  $W$  weight matrix is calculated. The

sparse filtering learned is used to extract from each sample the local characteristics. In terms of negative matrix factorization (NMF) and SVM with limitations of certainty, Padilla et al. [79] offer a novel conclusion technique for the early determination of Alzheimer's disease. Through implementing the Fisher discriminant proportion and non-NMF for highlighting preference and extraction of the most important highlights, the single-photon emission computed tomography and positron emission tomography datasets are studied (see Table 1).

### 3. Research Methodology

The proposed study uses the EEG signal to describe the phases of the disorder. It is suggested that a deep CNN network architecture is learned to distinguish multichannel human EEG signal data into different stages and that increases the efficiency of classification. This work includes the modules below:

- (i) Preprocessing
- (ii) Feature extraction
- (iii) Classification

**3.1. Feature Extraction.** The EEG trace is expected to be explicated in a function of frequency  $X[k]$ , using the discrete Fourier transform, as a product of the sampled representation of the EEG signal as  $x[j]$  with  $j = 1, 2, \dots, N$ , length  $N$ , and sampling frequency  $f_s$  (Hz). Parseval's theorem explains that the function's total square is the complete square of its transformation; the process starts with the extraction of features.

$$\sum_{j=0}^{N-1} |x[j]|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X[k]X^*[k]| = \sum_{k=0}^{N-1} P[k]. \quad (1)$$

According to the above equation,  $P[k]$  is the phase-excluded power spectrum. It means that the frequency index is obtained by multiplying  $X[k]$  by the  $X^*[k]$  conjugate divided by  $N$ , where the phase-excluded power spectrograph is  $P[k]$ , i.e.,  $X[k]$  has its conjugate  $X^*[k]$ , separated by  $N$ , which is compounded by  $k$ , and frequency index. The full definition of the frequency as obtained by the Fourier transform is usually well-known to be symmetrical concerning zero frequency; i.e., it has similar sections extending to the frequencies, which are both positive and negative [89]. The whole spectrum, including positive and negative frequencies, is free from this symmetry. Access to spectral power from the time domain has not been completed. According to the concept of a one-minute  $m$  of the order  $nP[k]$  of the power spectral density, all irregular moments are also zero by the frequency distribution model's statistical approach.

$$m_n = \sum_{k=0}^{N-1} k^n P[k]. \quad (2)$$

In the latter equation, the Parseval theorem of Equation (1) may be used where  $n = 0$  is used, and the Fourier transform time-differentiation feature for nonzero quantities of  $n$  is used. Such a feature explicitly indicates then  $n'$  the equal to multiply the  $k$  augmented by the spectrum to then  $n'$  th power, the derivative of a time-domain function referred to as  $\Delta^n$  for various time signals.

$$F[\Delta^n x[j]] = k^n X[k]. \quad (3)$$

To this end, as seen in Figure 1, the description of the characteristics used is as follows:

**Root squared zero-order moment ( $\overline{m_0}$ ):** this is a function that shows the overall power of the frequency domain and is as follows:

$$\overline{m_0} = \sqrt{\sum_{j=0}^{N-1} x[j]^2}. \quad (4)$$

All channels may also standardize their corresponding

zero-order moments by dividing all channels into zero-order moments.

**Root squared second and fourth-order moments:** according to Hjorth [89], the second time is used as power, but then, a spectrum shifted  $k^2 P[k]$ , referring to the frequency function:

$$\overline{m_2} = \sqrt{\sum_{k=0}^{N-1} k^2 P[k]} = \sqrt{\frac{1}{N} \sum_{k=0}^{N-1} (kX[k])^2} = \sqrt{\sum_{j=0}^{N-1} (\Delta x[j])^2}. \quad (5)$$

A repetition of this procedure gives the moment:

$$\overline{m_4} = \sqrt{\sum_{k=0}^{N-1} k^4 P[k]} = \sqrt{\sum_{j=0}^{N-1} (\Delta^2 x[j])^2}. \quad (6)$$

The total energy of the signal is decreased by consideration of the second and fourth signals; thus, a power transformer is implemented to normalize the domain of  $m_0$ ,  $m_2$ , and  $m_4$  to minimize the noise effect on all moment-based features as follows:

$$\begin{aligned} m_0 &= \frac{\overline{m_0}^\lambda}{\lambda}, \\ m_2 &= \frac{\overline{m_2}^\lambda}{\lambda}, \\ m_4 &= \frac{\overline{m_4}^\lambda}{\lambda}. \end{aligned} \quad (7)$$

The experimental setting of  $\lambda$  is 0.1. From these parameters, consequently, the first three features extracted are described as follows:

$$\begin{aligned} f_1 &= \log(m_0), \\ f_2 &= \log(m_0 - m_2), \\ f_3 &= \log(m_0 - m_4). \end{aligned} \quad (8)$$

**Sparseness:** this feature measures the amount of vector energy in just several more elements. It is followed as

$$f_4 = \log\left(\frac{m_0}{\sqrt{m_0 - m_2} \sqrt{m_0 - m_4}}\right). \quad (9)$$

A feature represents a vector with all elements equivalent to a zero-sparseness index, i.e.,  $m_2$  and  $m_4 = 0$ , due to differentiation and  $\log(m_0/m_0) = 0$ , while it should require a value greater than zero for all other sparseness levels [90].

**Irregularity factor (IF):** a measure that expresses the ratio of peak numbers divided by zero crossings. According to [91], only in terms of their spectral instances can the number of upward zero crossings (ZC) and the number of peaks (NP) in a random signal be specified. It is necessary to write the

corresponding feature as

$$f_5 = \frac{ZC}{NP} = \frac{\sqrt{m_2/m_0}}{\sqrt{m_4/m_2}} = \sqrt{\frac{m_2^2}{m_0 m_4}} = \frac{m_2}{\sqrt{m_0 m_4}}. \quad (10)$$

*Covariance (COV)*: COV function is the ratio of the standard deviation on arithmetic averages as follows:

$$f_6 = \log \left( \frac{\sqrt{\left( \sum_{j=0}^{N-1} (x - \bar{x})^2 \right) / n}}{\bar{x}} \right). \quad (11)$$

*Teager energy operator (TEO)*: it mainly displays the magnitude of the signal amplitude and instantaneous changes that are very susceptible to minor changes. While TEO was proposed for nonlinear speech signal modeling, it was later used to process audio signals. It is formed as follows:

$$f_7 = \log (\Psi(x[j])) = \log \left( \sum_{j=0}^{N-1} x^2[j] - x[j-1]x[j+1] \right). \quad (12)$$

In conjunction with the schematics in Figure 1, from each EEG record  $x$ , first, the seven features are extracted. In the classification method, the features described by the corresponding vector  $f$  are used. These characteristics can be assumed to reflect the EEG behavior in the form of cepstrum. Contrary to the well-known voice cepstral features [92], our EEG features have been obtained as the orientation between characteristics derived from a nonlinear EEG record and an initial EEG record following the equation. In the case of EEG classifications at differing levels of force, orientation-based feature extraction processes have recently been demonstrated to be of considerable significance for research on intact-limbed subjects as force generation relies on multiple hand muscle coordination [93].

There have not been any prior attempts to test the efficacy of specific features in amputees. In the coming subsection, the suggested orientation-based feature is adequate for amputees to classify EEG signals with different classes. In the remainder of the essay, the last feature  $f$  is defined, together with the time-dependent descriptor spectrum, from all channels, given as TD-PSD.

**3.2. Convolutional Neural Network.** CNN is one of the learning networks inspired by the MLP in this type of neural network. This deep network comprises an input layer, an output layer, and a deeply hidden layer. Firstly, the problem's signal or data are identified and trained into the algorithm. The hidden weights of the output layer appear in many forms [94]. If the algorithm output includes numerous numerical components, such as a binary number or index (e.g., signal classification, normal = 1, abnormal = 2), then the algorithm presented is a classification or detection algorithm. That is, the outcomes are weighted after the training of several signals. When a new signal is added to the algorithm other than the training signals, the signal form is identified, for example,

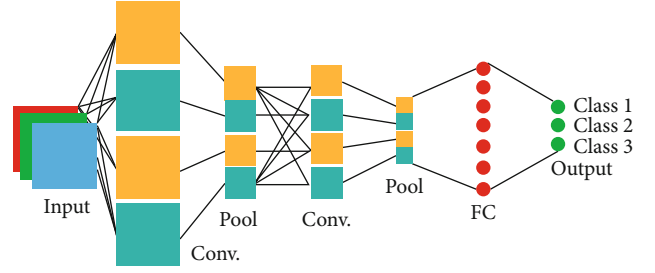


FIGURE 2: A conventional sublayer in CNN.

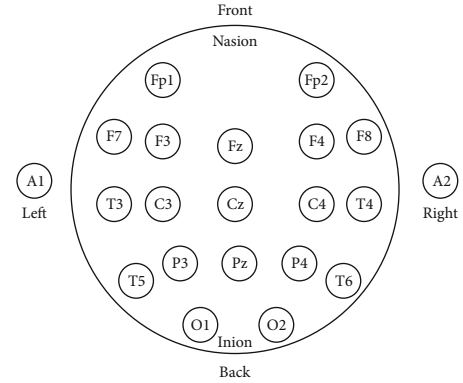


FIGURE 3: The place of electrodes of EEG signals.

whether a matrix of various kinds of signals is sent to the algorithm and trains the machine, signals of benign or malignant types of cancer, Alzheimer's, sarcoma, or brain tumor, for example. The type of disease can be identified by the algorithm with the weights obtained. CNN consists of numerous hidden sublayer forms that are explained as follows.

**3.3. Convolutional Sublayer.** The basic of the CNN is the convolutional sublayer, and it is possible to view its output matrix as a three-dimensional neuron matrix. For a deeper explanation of this, imagine traditional neural networks. Each layer was a little more than a list (one-dimensional as a rectangle) of neurons in regular neural networks in which each neuron generated its output, and gradually, a collection of outputs referring to each neuron was produced. However, instead of a single list, it is presented with a three-dimensional list (one cube) where the neurons are organized in three dimensions. Therefore, a three-dimensional matrix would also be the production of this cube. This principle and the distinction between the two are illustrated in the images below [95].

Let the size of the input matrix be  $20 \times 16 \times 16$ . Thus, utilizing a receptive field of  $3 \times 3$ , each neuron would have  $3 \times 3 \times 20 = 180$  connections to the input matrix in the conventional layer. Notice that space's connection is local (for example,  $3 \times 3$  here) but covers the maximum depth (Figure 2). An input displays the left image (for example, a  $3 \times 32 \times 32$  image). The neuron matrix is observed in the conventional blue sheet. In terms of spatial coordinates (length and width) in the input matrix, each neuron in the conventional layer is related to one local region only, but this connection extends

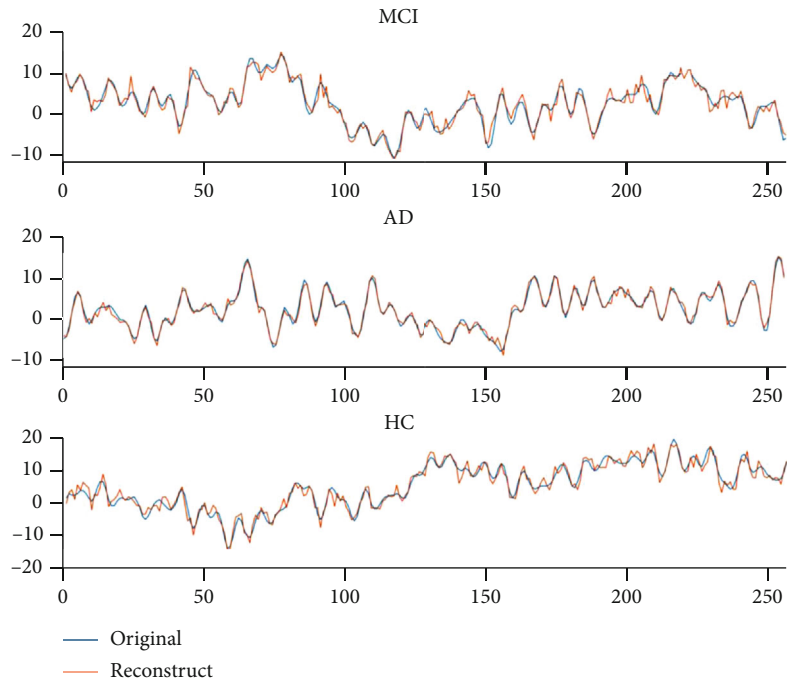


FIGURE 4: Plots of sampling from EEG signals for three categories of MCI, AD, and HC.

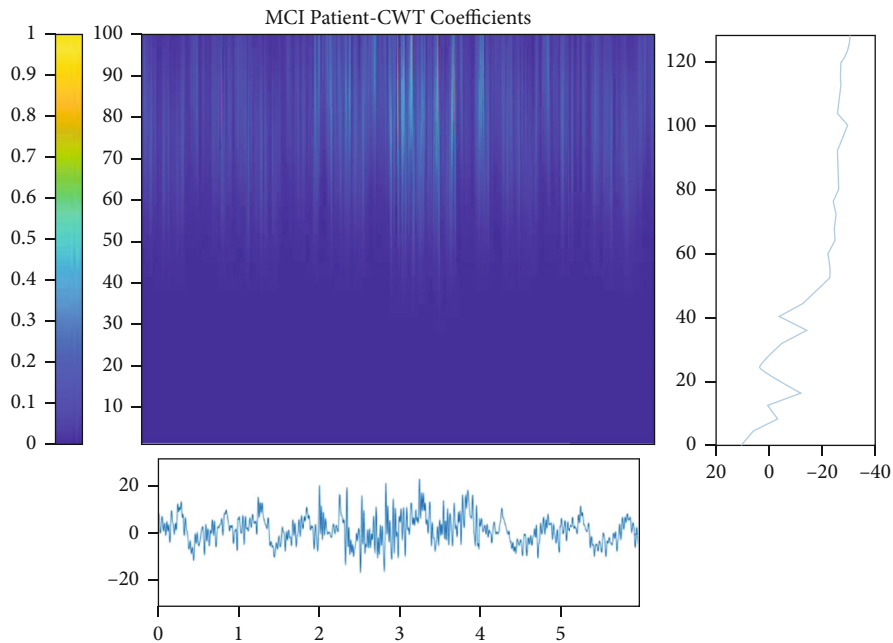


FIGURE 5: Time-frequency analysis of EEG signals for MCI patients.

in-depth (i.e., covers all color channels). There are depths (in this case, neurons) that all look at one place at the entrance.

The hyperparameter regulates the output matrix's dimension. The depth, stride, and zero-padding layer are these three parameters. The parameter can be used in the depth of the output matrix. This parameter regulates the number of neurons that bind to a region in the conventional layer's input matrix. This variable is analogous to multiple

neurons in a hidden layer all attached to one input in classical neural networks. All these neurons learn to function on various feedback features in which the deep columns around the spatial dimensions must be defined (width and height). When the stride is equal to 0, the spacing coordinates of just five spaced points are allocated to a new depth column of neurons. Also, in large output matrixes, this leads to receptive areas of overlap between columns. Alternatively, the

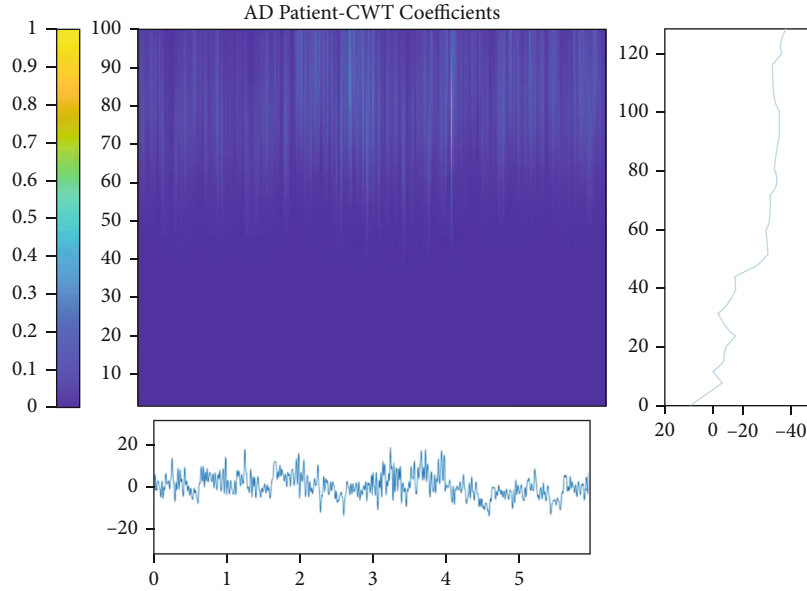


FIGURE 6: Time-frequency analysis of EEG signals for AD patients.

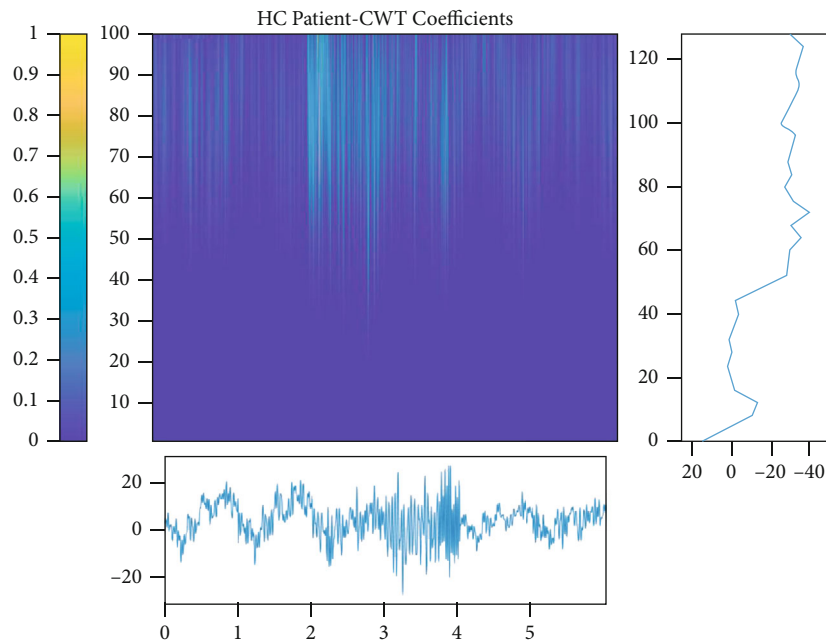


FIGURE 7: Time-frequency analysis of EEG signals for HC samples.

receptive areas are less frequent if the measures taken are bigger, and the output mass is smaller in the spatial dimension [96]. Overriding the inbound matrix with a zero pad is also more precise. In other terms, fill zero with the input picture circle. Our signal is put within a zero signal like inserting row 1 and column 2 at the beginning and end of a signal.

**3.4. Max Pool Sublayer.** One standard technique in conventional architecture is the positioning of a pooling layer between many successive layers. This layer's purpose is to minimize the matrix (input) size (width and height) by reducing the number of variables and calculations within

the grid and thereby overfitting the monitor. The pooling layer functions and uses it on each depth cut of the input matrix independently. The MAX function resizes the spot. The most typical way to utilize this layer is to use this layer with filters of  $2 \times 2$  sizes with phase  $S=2$  that eliminates any depth cuts at the input by deleting two elements from the width and two elements from the height and deleting 2% of the values [97].

**3.5. Activation Function.** Artificial neural networks' activation function determines the node's output node or "neuron" according to the input or group of inputs. In the next node,



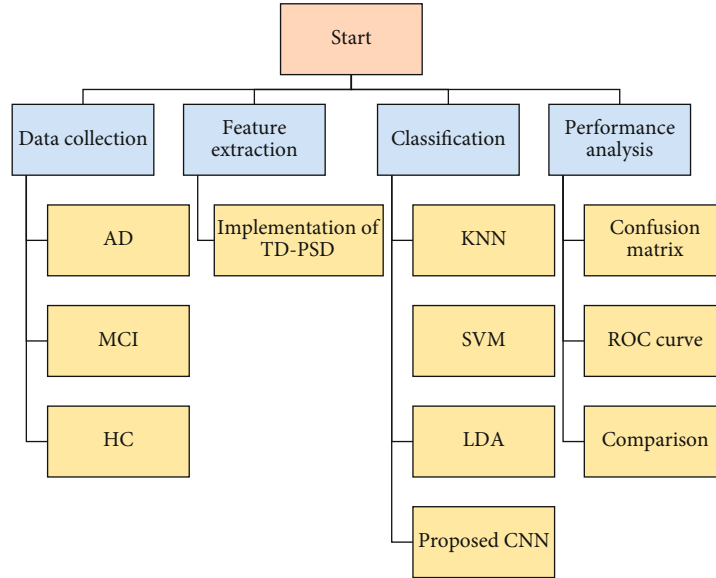


FIGURE 8: Conceptual diagram of the process used in this paper.

this output is known as the input. It follows before a solution to the issue is sought. The outcome values are translated to the target set, such as 0 to 1 or -1 to 1 (depending on the activation function selected). Using the logistic activation function, for example, transforms all inputs into true real ranges between 0 and 1. Another essential characteristic of an activation function is that it must be derivative to execute the optimization technique for backpropagation error and measure the weight gradient error in the network and use gradient descent or another optimal approach. Another optimization is optimizing weight to reduce weight. The rectified linear unit (ReLU) is used in this paper for the use of functions as follows:

$$f(x) = \begin{cases} 0, & x < 0, \\ x, & x \geq 0. \end{cases} \quad (13)$$

Some activation functions are also not unique to a single variable and refer to the vector or different variables used in this article, such as SoftMax [98]:

$$f_i(\vec{x}) = \frac{e^{x_i}}{\sum_{j=1}^J e^{x_j}}. \quad (14)$$

Also, to normalize input results, a batch normalization layer is applied to the network to speed up the training process and reduce network sensitivity between convolutional layers and nonlinearities. Also, to create an abnormal signal augmentation, a dropout layer on the fully connected layers is applied. Fully connected layers are used at the end of the hidden layer, which has been known to distinguish signals. The deep learning layer's outcome leads to a fully connected layer that drives the final classification judgment.

**3.6. Receiver Operating Characteristic (ROC) Curve.** In 2004, the ROC curves were developed, which were used to detect a radio noise signal [99]. These curves have recently been discovered to have important uses in medical decision-making. Presume having two kinds of individuals, one is normal and the other is a patient. It is a screening test on both our patients and healthy people, and the spectrum of values from the test ranges from 0 to the large number scale. In this case, the greater the test outcome, the greater the risk of the disease. (For certain things, the action can be the opposite.)

The ROC curve is established by the true-positive rate (TPR) projection in different threshold settings against the false-positive rate (FPR). The TPR is often identified in ML as sensitivity, recall or detection probability. Beginning from the ROC's left side, both the FPR and the TPR are zero at this point. (This argument indicates that the threshold line, which is the most significant number of test outcomes, is very large.) The TPR and FPR values in this case are measured and the next curve is drawn. The definition of TPR equals TP/Y, and the definition of FPR equals FP/N.

For a reduced number of previous values, let us reduce the threshold line. For lower values, the trend is replicated and eventually reaches the rightmost point of the ROC curve, which is equal to the baseline of the lowest value of the test outcome in this case. In this function, there is one TPR and one FPR. Accuracy relies entirely on allocating random errors and does not correspond with the real value or the value specified. In terms of bias, correctness is conveyed. A complete structural error can consist of one or more components of a systematic error. In the comparison value, a strong bias implies a large disparity. The two variables' sensitivity and specificity in statistics were used to determine the binary classification outcome (duality). When the data can be broken into positive and negative classes, using sensitivity and attribute indicators, the consistency of a test's

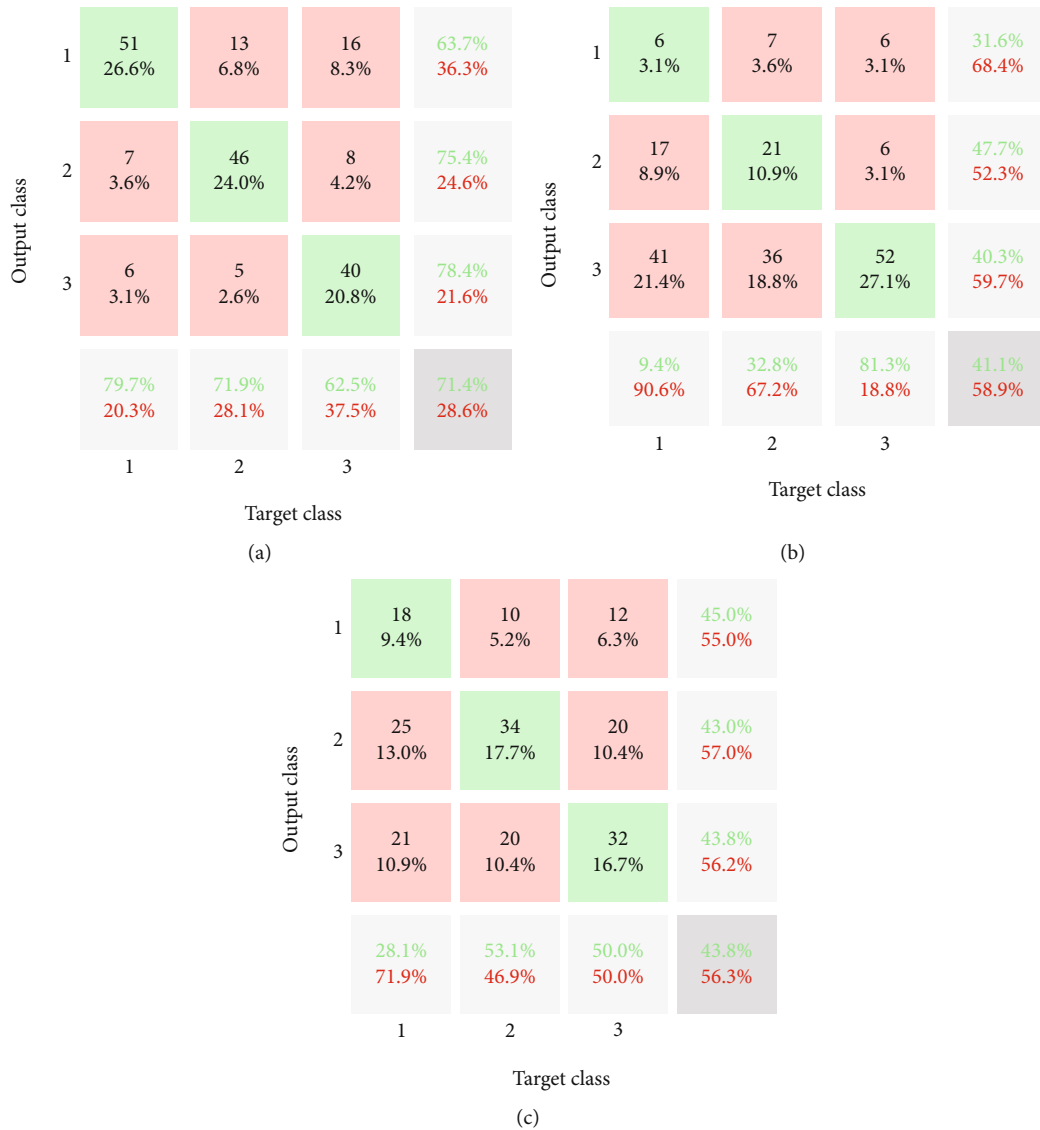


FIGURE 9: Confusion matrix of classification methods: (a) KNN, (b) SVM, and (c) LDA.

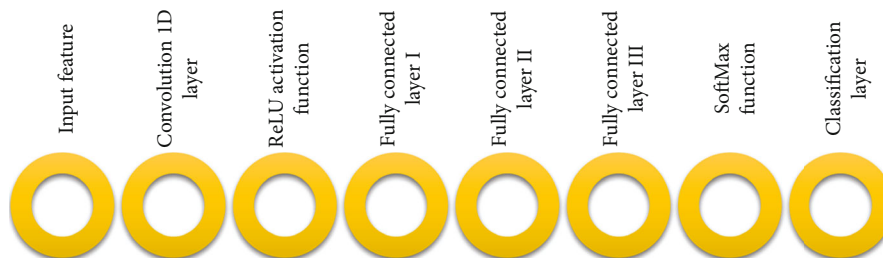


FIGURE 10: The architecture of the CNN layer for classification Alzheimer's disease.

outcomes that separate the information into these two divisions is observable and descriptive. Sensitivity means the number of positive cases that would be accurately checked as positive. Specificity means the number of negative cases that accurately label them as negative (positive = special illness, negative = other cases).

*True positive (TP):* the positive signal is accurately detected.

*False positive (FP):* the negative signal is detected with mistakes.

*True negative (TN):* the negative signal is detected accurately.

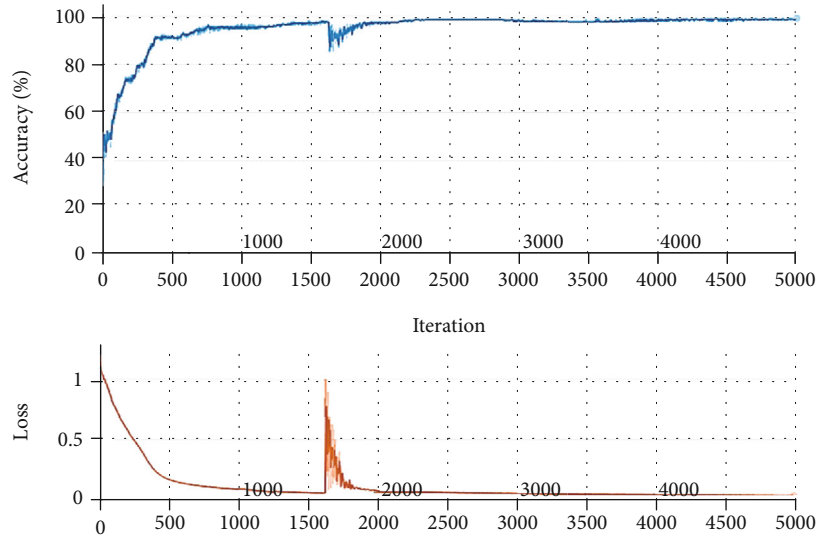


FIGURE 11: Accuracy and loss values for classification process of the presented CNN.

*False negative (FN):* the positive signal is detected with mistakes.

The sensitivity divides TP cases into the sum of true-positive and false-negative cases in statistical terms.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (15)$$

The confusion matrix is the role of the algorithms described in the field of artificial intelligence. Usually, for supervised learning algorithms, such a demonstration is being used, but it is often utilized in unsupervised learning. An instance of the predicted value is seen in each column of the matrix. Suppose it includes an actual (true) instance in each row. This matrix’s name is also gained, making it possible to mistake and mess with the results. This matrix is commonly called a contingency matrix or an error matrix outside of artificial intelligence.

#### 4. Results and Discussion

**4.1. Data Collection.** Multichannel EEG signals were captured using earlobe-electrode hallmark monopolar connections [100]. The location of the electrodes over the scalp was obtained according to the 10-20 International Electrode Positioning Method (i.e., Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, O2). In Figure 3, an example of EEG electrode placement is shown. The electrodes calculate the weak electrical potential in the microvolt range produced by brain activities. Recordings were conducted with closed eyes in a sleeping state. In this way, multiple brain regions can be believed to be governed by the exact hierarchical mechanism.

Data with a signal length of 300 seconds and a sampling frequency of 1024 and 256 samples per second were obtained. Just 180 seconds is extracted for each signal (i.e., from 60 to 240 seconds) to minimize the EEG context artifacts and convert each one to 256 samples per second. The

Output class	1	53 27.6%	3 1.6%	9 4.7%	81.5% 18.5%
	2	8 4.2%	57 29.7%	7 3.6%	79.2% 20.8%
	3	3 1.6%	4 2.1%	48 25.0%	87.3% 12.7%
		82.8% 17.2%	89.1% 10.9%	75.0% 25.0%	82.3% 17.7%
	1	2	3	Target class	

FIGURE 12: Confusion matrix of the presented CNN method for classification of patients suffering from Alzheimer’s.

sampling frequency, or sample rate, is the number of equal-spaced samples per unit of time. For instance, if there are 1024 equally spaced observations per second, the sampling rate is 1024/second or 1024 Hz. In this paper, for each second, 256 samples are used. Therefore, the length of the time series is  $300 * 256 = 76800$ . Moreover, the time interval from 60 to 240 seconds (of 300-second signals) is used with length  $240 - 60 = 180$ . Finally, the length of the signal is  $180 * 256 = 46080$ . To summarize, the EEG capacity of human samples belonging to three categories has been recorded:

- (1) Patients who have Alzheimer’s disease (AD)
- (2) Patients who suffer from mild cognitive impairment (MCI)
- (3) Healthy control (HC)

Figure 4 gives an example of extracted EEG recordings of 256 samples for each group.

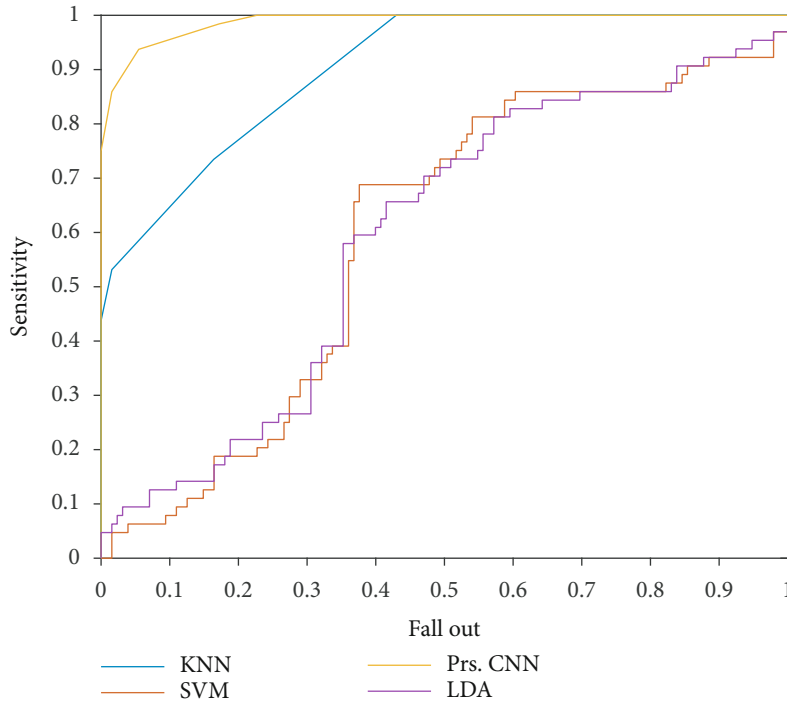


FIGURE 13: The ROC curve for the presented methods.

Figures 5–7 show each category’s time-frequency analysis based on continuous threshold wavelet transform. The EEG signal processing strategies are addressed in this section to extract the necessary quality information from clinical limitations and simultaneously determine patients’ status. The dispersion coefficient is a strong AD marker that can separate AD EEG from MCI and HC.

**4.2. Implementation of the Proposed Pattern Detection Method.** In this research, the automatic classification of normal and abnormal EEG signals was built in a deep learning model. EEG signals are used in the EEG database. Documents are separated into two sections of the EEG database: training and testing samples. During the learning stage, the training data were used, and during the evaluation stage, assessment data were used for training the algorithms with 80% used as validation data, while the remaining with 20%.

These data distributions have been randomly chosen. For validation purposes, many datasets have been used since the model parameter is set in several steps. However, experimental results with a specific random seed value are obtained to ensure that the model is reproducible and uniform. Figure 6 provides a thorough description of the data considered for this work. The block diagram of the process is described in Figure 8. Based on the process, firstly, TD-PSD is used for feature extraction of input EEG signals. The results of feature extraction are the produced seven features of values from 256 EEG samples.

Regarding this fact, the number of input variables for each person reduced from 256 to 7 features. It not only reduced the number of inputs but also increased the classification process speed. Signals with the time-dependent value

are needed to be transformed into a meaningful value. Therefore, these values are extracted features. Altogether, it is 64 persons for each category of MCI, AD, and HC that each one has seven features. Therefore, the input matrix for each category has  $7 \times 64$  elements.

In the next step, the output features are used for classification methods. For achieving this purpose, three powerful traditional methods of classification of EEG signal and diagnosis of AD and MCI patients from HC or healthy people are used. Used methods include KNN, SVM, and LDA. All the methods are chosen from powerful machine learning techniques. The results of classification for mentioned methods are as follows (Figure 9). Figure 9 shows the confusion matrix of utilized methods of classification. The labeling (1, 2, and 3) illustrates MCI, AD, and HC, respectively. Based on the confusion matrix. The green values show the number of persons diagnosed correctly from 64 people. Based on Figure 9(a), from 64 patients who suffered from MCI, 51 of them (79.7%) are diagnosed correctly.

Moreover, for people with AD, the sensitivity of the KNN is 71.9% (46 out of 64 patients). Moreover, 62.5% (40 persons) of healthy samples are detected correctly. Based on the results, the lower row of the matrix illustrates the sensitivity of the methods for the classification of each category. Moreover, red %ages indicated miss rates of each class. The right column of the matrix represents the precision of the KNN technique. Based on the results, 63.7% of the MCI-diagnosed patients are classified correctly. Other precision values are depicted in the right column. Finally, the accuracy of the methods is presented in the lower-right corner of the matrix. Based on the result, the accuracy of KNN, SVM, and LDA approaches is 71.4%, 41.1%, and 43.8%, respectively. In other words, the accuracy of the KNN is higher than

TABLE 2: Comparison of the diagnosis methods used in this paper.

	Sensitivity			Precision			AUC	Accuracy
	MCI	AD	HC	MCI	AD	HC		
KNN	79.7%	71.9%	62.5%	63.7%	75.4%	78.4%	0.902	71.4%
SVM	9.4%	32.8%	81.3%	31.6%	47.7%	40.3%	0.593	41.1%
LDA	28.1%	53.1%	50.0%	45.0%	43.0%	48.8%	0.594	43.8%
Presented CNN	82.8%	89.1%	75.5%	81.5%	79.2%	87.3%	0.988	82.3%

that of the other methods. On the other hand, SVM and LDA methods represent results with very low accuracy. It can be noticed that the classification with two categories of patients and healthy people illustrates higher accuracy in comparison with three categories.

In this paper, a novel deep learning process for the classification of EEG signals is introduced based on a CNN. The architecture of the presented method is illustrated in Figure 10. Input layer includes the following:

- (i) Seven features of 64 people for each category of MCI, AD, and HC
- (ii) Input matrix 4D [ $7 \times 1 \times 1 \times 192$ ]
- (iii) Output matrix 1D [ $1 \times 192$ ]

The labeling (1, 2, and 3) illustrates MCI, AD, and HC, respectively, output matrix. The process of the classification using the presented CNN architecture is shown in Figure 11. The maximum accuracy for the process reached almost 100%, and the loss value decreased to almost zero.

The confusion matrix of the presented method is indicated in Figure 12 without uncertainty value. The confident value of the accuracy for the CNN approach is 82.3%. In detail, 85% of the MCI patients are detected correctly; however, 89.1% of the AD and 75% of the normal sample are diagnosed correctly. For comparison of the resulted model of CNN with other KNN, SVM, and LDA techniques, the ROC plot is depicted in Figure 13. Regarding Figure 13, the value of FP rate versus TP rate is depicted based on output classification scores. Based on this graph, if the values would be higher for the TP rate and lower for the FP rate, it is better than the other graphs.

Moreover, the area under the curve (AUC) is one of the criteria for the classification method's performance analysis. Based on Figure 13, the presented CNN outperforms other methods, and the KNN method is in the following priority. The LDA and SVM reached low AUC values (see Table 2). To conclude, it can be seen that the presented CNN architecture is better and more accurate than other classification methods. Based on Table 2, the best method for diagnosing AD patients from EEG signals is the presented architecture's CNN approach.

## 5. Conclusion

In this article, the TD-PSD approach is used for EEG signal feature extraction from three groups of MCI, AD, and HC test samples. The final features used in three conventional

classification methods are registered: KNN, SVM, and LDA, and the effects are recorded. Finally, the CNN architecture is provided for AD patient classification. The findings are indicated using performance measurement. Data were obtained with a signal duration of 300 seconds and a sampling frequency of 1024 and 256 samples per second. To minimize the EEG background artifacts, it obtains 180 seconds for each signal (i.e., from 60 to 240 seconds), and it is translated to 256 samples per second. The EEG ability of human samples belonging to three groups, including AD, MCI, and HC, has been summarized. First, the TD-PSD method is utilized to feature input EEG signals based on the procedure. The results of the extraction of features are the generation of seven value characteristics from 256 EEG samples. For classification methods, the output features are used in the next step. Methods used include KNN, SVM, and LDA. All the strategies are picked from effective methods in machine learning.

Based on the findings, 51 of the 64 patients with MCI (79.7%) were correctly diagnosed. In comparison, the KNN sensitivity for AD persons is 71.9% (46 out of 64 patients). In comparison, 62.5% (40 individuals) of sound samples are appropriately classified. Also, 63.7% of the MCI diagnosed patients are appropriately categorized based on the findings. In comparison, KNN, SVM, and LDA methods have a precision of 71.4%, 41.1%, and 43.8%, respectively. The precision of KNN, in other words, is greater than that of other processes. Then, a new EEG signal classification architecture is implemented that is focused on a CNN. For the CNN approach, the accurate meaning of accuracy is 82.3%. 85% of MCI cases are accurately detected in-depth, but 89.1% of the AD and 75% of the healthy population are correctly diagnosed. The presented CNN outperforms other approaches based on performance, and the KNN approach is the next target. The LDA and SVM were at low AUC values. For potential outcomes, it is recommended to modify feature extraction with another EEG signal justification for classification with a lower number of features and training using the design provided.

## Data Availability

Data is extracted from the following ref: An integrated approach based on EEG signals processing combined with supervised methods to classify Alzheimer's disease patients.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] World Health Organization, *Alzheimer's Disease International. Dementia: A Public Health Priority*, vol. 2013, World Health Organization, Geneva, 2012.
- [2] B. Duthey, "Background paper 6.11: Alzheimer disease and other dementias," *A Public Health Approach to Innovation*, vol. 6, pp. 1–74, 2013.
- [3] M. Prince, A. Wimo, M. Guerchet et al., "The global impact of dementia: an analysis of prevalence, incidence, cost and trends," *Alzheimer's Disease International*, vol. 2015, 2015.
- [4] E. E. Lee, B. Chang, S. Huege, and J. Hirst, "Complex clinical intersection: palliative care in patients with dementia," *The American Journal of Geriatric Psychiatry*, vol. 26, no. 2, pp. 224–234, 2018.
- [5] C. R. Jack Jr., M. S. Albert, D. S. Knopman et al., "Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," *Alzheimer's & Dementia*, vol. 7, no. 3, pp. 257–262, 2011.
- [6] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, and E. M. Stadlan, "Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group\* under the auspices of department of health and human services task force on Alzheimer's disease," *Neurology*, vol. 34, no. 7, pp. 939–944, 1984.
- [7] J. Hort, J. T. O'Brien, G. Gainotti et al., "EFNS guidelines for the diagnosis and management of Alzheimer's disease," *European Journal of Neurology*, vol. 17, no. 10, pp. 1236–1248, 2010.
- [8] M. F. Folstein, S. E. Folstein, and P. R. McHugh, "Mini-mental state: a practical method for grading the cognitive state of patients for the clinician," *Journal of Psychiatric Research*, vol. 12, no. 3, pp. 189–198, 1975.
- [9] A. J. Mitchell, "A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment," *Journal of Psychiatric Research*, vol. 43, no. 4, pp. 411–431, 2009.
- [10] Z. S. Nasreddine, N. A. Phillips, V. Bédirian et al., "The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment," *Journal of the American Geriatrics Society*, vol. 53, no. 4, pp. 695–699, 2005.
- [11] P. S. Mathuranath, P. J. Nestor, G. E. Berrios, W. Rakowicz, and J. R. Hodges, "A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia," *Neurology*, vol. 55, no. 11, pp. 1613–1620, 2000.
- [12] P. Amodio, H. Wenin, F. del Piccolo et al., "Variability of trail making test, symbol digit test and line trait test in normal people. A normative study taking into account age-dependent decline and sociobiological variables," *Aging Clinical and Experimental Research*, vol. 14, no. 2, pp. 117–131, 2002.
- [13] K. I. Shulman, "Clock-drawing: is it the ideal cognitive screening test?," *International Journal of Geriatric Psychiatry*, vol. 15, no. 6, pp. 548–561, 2000.
- [14] N. K. al-Qazzaz, S. H. B. M. D. Ali, S. A. Ahmad, K. Chellappan, M. S. Islam, and J. Escudero, "Role of EEG as biomarker in the early detection and classification of dementia," *The Scientific World Journal*, vol. 2014, Article ID 906038, 16 pages, 2014.
- [15] T. J. Montine, C. H. Phelps, T. G. Beach et al., "National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease: a practical approach," *Acta Neuropathologica*, vol. 123, no. 1, pp. 1–11, 2012.
- [16] L. Wu, L. Wu, Y. Chen, and J. Zhou, "A promising method to distinguish vascular dementia from Alzheimer's disease with standardized low-resolution brain electromagnetic tomography and quantitative EEG," *Clinical EEG and Neuroscience*, vol. 45, no. 3, pp. 152–157, 2014.
- [17] B. Dubois, H. H. Feldman, C. Jacova et al., "Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria," *The Lancet Neurology*, vol. 13, no. 6, pp. 614–629, 2014.
- [18] D. Ferreira, L. Perestelo-Pérez, E. Westman, L. O. Wahlund, A. Sarría, and P. Serrano-Aguilar, "Meta-review of CSF core biomarkers in Alzheimer's disease: the state-of-the-art after the new revised diagnostic criteria," *Frontiers in Aging Neuroscience*, vol. 6, p. 47, 2014.
- [19] B. Olsson, R. Lautner, U. Andreasson et al., "CSF and blood biomarkers for the diagnosis of Alzheimer's disease: a systematic review and meta-analysis," *The Lancet Neurology*, vol. 15, no. 7, pp. 673–684, 2016.
- [20] M. Seeck, L. Koessler, T. Bast et al., "The standardized EEG electrode array of the IFCN," *Clinical Neurophysiology*, vol. 128, no. 10, pp. 2070–2077, 2017.
- [21] S. Hu, Y. Lai, P. A. Valdes-Sosa, M. L. Bringas-Vega, and D. Yao, "How do reference montage and electrodes setup affect the measured scalp EEG potentials?," *Journal of Neural Engineering*, vol. 15, no. 2, article 026013, 2018.
- [22] J. Wang, J. Barstein, L. E. Ethridge, M. W. Mosconi, Y. Takarae, and J. A. Sweeney, "Resting-state EEG abnormalities in autism spectrum disorders," *Journal of Neurodevelopmental Disorders*, vol. 5, no. 1, 2013.
- [23] O. Faust, U. R. Acharya, H. Adeli, and A. Adeli, "Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis," *Seizure*, vol. 26, pp. 56–64, 2015.
- [24] C. F. Muniz, A. V. Shenoy, K. L. O'Connor et al., "Clinical development and implementation of an institutional guideline for prospective EEG monitoring and reporting of delayed cerebral ischemia," *Journal of Clinical Neurophysiology*, vol. 33, no. 3, pp. 217–226, 2016.
- [25] K. Nishida, M. Yoshimura, T. Isotani et al., "Differences in quantitative EEG between frontotemporal dementia and Alzheimer's disease as revealed by LORETA," *Clinical Neurophysiology*, vol. 122, no. 9, pp. 1718–1725, 2011.
- [26] D.-H. Jeong, Y.-D. Kim, I.-U. Song, Y.-A. Chung, and J. Jeong, "Wavelet energy and wavelet coherence as EEG biomarkers for the diagnosis of Parkinson's disease-related dementia and Alzheimer's disease," *Entropy*, vol. 18, no. 1, p. 8, 2016.
- [27] E. Neto, E. A. Allen, H. Aurlien, H. Nordby, and T. Eichele, "EEG spectral features discriminate between Alzheimer's and vascular dementia," *Frontiers in Neurology*, vol. 6, 2015.
- [28] E. Neto, F. Biessmann, H. Aurlien, H. Nordby, and T. Eichele, "Regularized linear discriminant analysis of EEG features in dementia patients," *Frontiers in Aging Neuroscience*, vol. 8, 2016.
- [29] S. J. Colloby, R. A. Cromarty, L. R. Peraza et al., "Multimodal EEG-MRI in the differential diagnosis of Alzheimer's disease and dementia with Lewy bodies," *Journal of Psychiatric Research*, vol. 78, pp. 48–55, 2016.
- [30] H. Garn, C. Coronel, M. Waser, G. Caravias, and G. Ransmayr, "Differential diagnosis between patients with probable Alzheimer's disease, Parkinson's disease dementia,

- or dementia with Lewy bodies and frontotemporal dementia, behavioral variant, using quantitative electroencephalographic features,” *Journal of Neural Transmission*, vol. 124, no. 5, pp. 569–581, 2017.
- [31] P. L. Nunez and R. Srinivasan, *Electric Fields of the Brain: The Neurophysics of EEG*, Oxford University Press, New York, NY, USA, 2006.
- [32] M. Cohen, *Analyzing Neural Time Series Data: Theory and Practice*, MIT Press, 2014.
- [33] L. Sörnmo and P. Laguna, *Bioelectrical Signal Processing in Cardiac and Neurological Applications*, Academic Press, 2005.
- [34] J. Dauwels, K. Srinivasan, M. Ramasubba Reddy et al., “Slowing and loss of complexity in Alzheimer’s EEG: two sides of the same coin?,” *International Journal of Alzheimer’s Disease*, vol. 2011, article 539621, pp. 1–10, 2011.
- [35] A. Alberdi, A. Aztiria, and A. Basarab, “On the early diagnosis of Alzheimer’s disease from multimodal signals: a survey,” *Artificial Intelligence in Medicine*, vol. 71, pp. 1–29, 2016.
- [36] N. Malek, M. R. Baker, C. Mann, and J. Greene, “Electroencephalographic markers in dementia,” *Acta Neurologica Scandinavica*, vol. 135, no. 4, pp. 388–393, 2017.
- [37] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, “Alzheimer’s disease: models of computation and analysis of EEGs,” *Clinical and EEG Neuroscience*, vol. 36, no. 3, pp. 131–140, 2005.
- [38] B. Jelles, J. H. van Birgelen, J. P. Slaets, R. E. Hekster, E. J. Jonkman, and C. J. Stam, “Decrease of non-linear structure in the EEG of Alzheimer patients compared to healthy controls,” *Clinical Neurophysiology*, vol. 110, no. 7, pp. 1159–1167, 1999.
- [39] J. Jeong, S. Y. Kim, and S. H. Han, “Non-linear dynamical analysis of the EEG in Alzheimer’s disease with optimal embedding dimension,” *Electroencephalography and Clinical Neurophysiology*, vol. 106, no. 3, pp. 220–228, 1998.
- [40] J. Jeong, J. H. Chae, S. Y. Kim, and S. H. Han, “Non-linear dynamic analysis of the EEG in patients with Alzheimer’s disease and vascular dementia,” *Journal of Clinical Neurophysiology*, vol. 18, no. 1, pp. 58–67, 2001.
- [41] T. Takahashi, “Complexity of spontaneous brain activity in mental disorders,” *Progress in Neuro-psychopharmacology & Biological Psychiatry*, vol. 45, pp. 258–266, 2013.
- [42] T. Yagyu, J. Wackermann, M. Shigeta et al., “Global dimensional complexity of multi-channel EEG in mild Alzheimer’s disease and age-matched cohorts,” *Dementia and Geriatric Cognitive Disorders*, vol. 8, no. 6, pp. 343–347, 1997.
- [43] N. Houmani, F. B. Vialatte, C. Latchoumane, J. Jeong, and G. Dreyfus, “Stationary epoch-based entropy estimation for early diagnosis of Alzheimer’s disease,” in *2013 IEEE Faible Tension Faible Consommation*, pp. 1–4, Paris, France, 2013.
- [44] N. Houmani, F. B. Vialatte, and G. Dreyfus, “Epoch-based entropy for early screening of Alzheimer’s disease,” *International Journal of Neural Systems*, vol. 25, no. 8, article 1550032, 2015.
- [45] D. Aba’solo, R. Hornero, P. Espino, D. Alvarez, and J. Poza, “Entropy analysis of the EEG background activity in Alzheimer’s disease patients,” *Physiological Measurement*, vol. 27, no. 3, pp. 241–253, 2006.
- [46] T. J. De Bock, S. Das, M. Mohsin et al., “Early detection of Alzheimer’s disease using non-linear analysis of EEG via Tsallis entropy,” in *2010 Biomedical Sciences and Engineering Conference*, pp. 1–4, Oak Ridge, TN, USA, 2010.
- [47] D. Aba’solo, R. Hornero, P. Espino, J. Poza, C. I. Sanchez, and R. de la Rosa, “Analysis of regularity in the EEG background activity of Alzheimer’s disease patients with approximate entropy,” *Clinical Neurophysiology*, vol. 116, no. 8, pp. 1826–1834, 2005.
- [48] S. M. Pincus, “Approximate entropy as a measure of irregularity for psychiatric serial metrics,” *Bipolar Disorders*, vol. 8, no. 5p1, pp. 430–440, 2006.
- [49] J. Escudero, D. Aba’solo, R. Hornero, P. Espino, and M. Lopez, “Analysis of electroencephalograms in Alzheimer’s disease patients with multiscale entropy,” *Physiological Measurement*, vol. 27, no. 11, pp. 1091–1106, 2006.
- [50] D. Aba’solo, R. Hornero, and C. Gomez, “Analysis of EEG background activity in Alzheimer’s disease patients with Lempel-Ziv complexity and central tendency measure,” *Medical Engineering & Physics*, vol. 28, no. 4, pp. 315–322, 2006.
- [51] W. Chen, Y. Y. Lam, C. P. Shen, H. Y. Sung, M. J. Chiu, and F. Lai, “Ultra-fast epileptic seizure detection using EMD based on multi-channel electroencephalogram,” in *13th IEEE International Conference on BioInformatics and BioEngineering*, pp. 1–4, Chania, Greece, 2013.
- [52] G. Pei, J. Wu, D. Chen et al., “Effects of an integrated neuro-feedback system with dry electrodes: EEG acquisition and cognition assessment,” *Sensors*, vol. 18, no. 10, p. 3396, 2018.
- [53] G. Swapna, G. Swapna, G. Swapna, J. S. Suri, and J. S. Suri, “Automated EEG analysis of epilepsy: a review,” *Knowledge-Based Systems*, vol. 45, pp. 147–165, 2013.
- [54] T. Yan, W. Wang, T. Liu et al., “Increased local connectivity of brain functional networks during facial processing in schizophrenia: evidence from EEG data,” *Oncotarget*, vol. 8, no. 63, pp. 107312–107322, 2017.
- [55] G. Chen, W. Xie, T. D. Bui, and A. Krzyzak, “Automatic epileptic seizure detection in EEG using nonsubsampling wavelet-Fourier features,” *Journal of Medical and Biological Engineering*, vol. 37, no. 1, pp. 123–131, 2017.
- [56] A. Hamad, E. H. Houssein, A. E. Hassanien, and A. A. Fahmy, “A hybrid EEG signals classification approach based on grey wolf optimizer enhanced SVMs for epileptic detection,” in *International Conference on Advanced Intelligent Systems and Informatics*, pp. 108–117, Cairo, Egypt, 2017.
- [57] A. Subasi, J. Kevric, and M. Abdullah Canbaz, “Epileptic seizure detection using hybrid machine learning methods,” *Neural Computing and Applications*, vol. 31, no. 1, pp. 317–325, 2017.
- [58] B. Wang, T. Yan, S. Ohno, S. Kanazawa, and J. Wu, “Retinotopy and attention to the face and house images in the human visual cortex,” *Experimental Brain Research*, vol. 234, no. 6, pp. 1623–1635, 2016.
- [59] D. Ahmedt-Aristizabal, C. Fookes, K. Nguyen, S. Denman, S. Sridharan, and S. Dionisio, “Deep facial analysis: a new phase I epilepsy evaluation using computer vision,” *Epilepsy & Behavior*, vol. 82, pp. 17–24, 2018.
- [60] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, “Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals,” *Computers in Biology and Medicine*, vol. 100, pp. 270–278, 2018.
- [61] R. Hussein, H. Palangi, R. Ward, and Z. J. Wang, *Epileptic Seizure Detection: A Deep Learning Approach* 2018, <https://arxiv.org/abs/1803.09848>.
- [62] N. Kulkarni and V. Bairagi, *EEG-Based Diagnosis of Alzheimer Disease: A Review and Novel Approach for Feature*

*Extraction and Classification Techniques*, Academic Press, Cambridge, 2018.

- [63] M. Z. Parvez and M. Paul, "Seizure prediction using undulated global and local features," *IEEE Transactions on Bio-medical Engineering*, vol. 64, no. 1, pp. 208–217, 2017.
- [64] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, R. Wald, E. Muharemagic, and N. Seliya, "Deep learning applications and challenges in big data analytics," *Journal of Big Data*, vol. 2, no. 1, 2015.
- [65] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [66] M. Coşkun, Ö. Yildirim, A. Uçar, and Y. Demir, "An overview of popular deep learning methods," *European Journal of Technic*, vol. 7, no. 2, pp. 165–176, 2017.
- [67] J. Günther, P. M. Pilarski, G. Helfrich, H. Shen, and K. Diepold, "First steps towards an intelligent laser welding architecture using deep neural networks and reinforcement learning," *Procedia Technology*, vol. 15, pp. 474–483, 2014.
- [68] Y. LeCun, B. Boser, J. S. Denker et al., "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, no. 4, pp. 541–551, 1989.
- [69] A. Uçar, Y. Demir, and C. Güzeliş, "Object recognition and detection with deep learning for autonomous driving applications," *Simulation*, vol. 93, no. 9, pp. 759–769, 2017.
- [70] F. Beser, M. A. Kizrak, B. Bolat, and T. Yildirim, "Recognition of sign language using capsule networks," in *2018 26th IEEE signal processing and communications applications conference (SIU)*, Izmir, Turkey, 2018.
- [71] V. Mnih, K. Kavukcuoglu, D. Silver et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.
- [72] O. Yildirim, R. S. Tan, and U. R. Acharya, "An efficient compression of ECG signals using deep convolutional autoencoders," *Cognitive Systems Research*, vol. 52, pp. 198–211, 2018.
- [73] Ö. Yildirim, "A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification," *Computers in Biology and Medicine*, vol. 96, pp. 189–202, 2018.
- [74] O. Yildirim, P. Pławiak, R. S. Tan, and U. R. Acharya, "Arrhythmia detection using deep convolutional neural network with long duration ECG signals," *Computers in Biology and Medicine*, vol. 102, pp. 411–420, 2018.
- [75] S. L. Oh, Y. Hagiwara, U. Raghavendra et al., "A deep learning approach for Parkinson's disease diagnosis from EEG signals," *Neural Computing and Applications*, vol. 32, no. 15, pp. 10927–10933, 2020.
- [76] X. Hong, R. Lin, and C. Yang, "Predicting Alzheimer's disease using LSTM," *IEEE Access*, vol. 7, pp. 2169–3536, 2019.
- [77] M. Aghili, S. Tabarestani, and M. Adjouadi, "Predictive modeling of longitudinal data for Alzheimer's disease diagnosis using RNNs," in *International Workshop on Predictive Intelligence In MEDicine*, pp. 112–119, Granada, Spain, 2018.
- [78] F. Razavi, M. J. Tarokh, and M. Alborzi, "An intelligent Alzheimer's disease diagnosis method using unsupervised feature learning," *Journal of Big Data*, vol. 6, no. 1, 2019.
- [79] P. Padilla, M. López, J. M. Górriz, J. Ramírez, D. Salas-González, and I. Álvarez, "NMF-SVM based CAD tool applied to functional brain images for the diagnosis of Alzheimer's disease," *IEEE Transactions on Medical Imaging*, vol. 31, no. 2, pp. 207–216, 2012.
- [80] Q. Xin, S. Hu, S. Liu, X. Ma, H. Lv, and Y. D. Zhang, "Epilepsy EEG classification based on convolution support vector machine," *Journal of Medical Imaging and Health Informatics*, vol. 11, no. 1, pp. 25–32, 2021.
- [81] I. Aliyu and C. G. Lim, "Selection of optimal wavelet features for epileptic EEG signal classification with LSTM," *Neural Computing and Applications*, pp. 1–21, 2021.
- [82] T. Tuncer, "A new stable non-linear textural feature extraction method-based EEG signal classification method using substitution Box of the Hamsi hash function: Hamsi pattern," *Applied Acoustics*, vol. 172, article 107607, 2021.
- [83] P. A. Cicalese, R. Li, M. B. Ahmadi et al., "An EEG-fNIRS hybridization technique in the four-class classification of Alzheimer's disease," *Journal of Neuroscience Methods*, vol. 336, article 108618, 2020.
- [84] R. Ferri, C. Babiloni, V. Karami et al., "Stacked autoencoders as new models for an accurate Alzheimer's disease classification support using resting-state EEG and MRI measurements," *Clinical Neurophysiology*, vol. 132, no. 1, pp. 232–245, 2021.
- [85] L. R. Trambaiolli, N. Spolaôr, A. C. Lorena, R. Anghinah, and J. R. Sato, "Feature selection before EEG classification supports the diagnosis of Alzheimer's disease," *Clinical Neurophysiology*, vol. 128, no. 10, pp. 2058–2067, 2017.
- [86] S. Nobukawa, T. Yamanishi, S. Kasakawa, H. Nishimura, M. Kikuchi, and T. Takahashi, "Classification methods based on complexity and synchronization of electroencephalography signals in Alzheimer's disease," *Frontiers in Psychiatry*, vol. 11, 2020.
- [87] N. N. Kulkarni and V. K. Bairagi, "Extracting salient features for EEG-based diagnosis of Alzheimer's disease using support vector machine classifier," *IETE Journal of Research*, vol. 63, no. 1, pp. 11–22, 2017.
- [88] F. Vecchio, F. Miraglia, F. Alù et al., "Classification of Alzheimer's disease with respect to physiological aging with innovative EEG biomarkers in a machine learning implementation," *Journal of Alzheimer's Disease*, vol. 75, no. 4, pp. 1253–1261, 2020.
- [89] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalography and Clinical Neurophysiology*, vol. 29, no. 3, pp. 306–310, 1970.
- [90] R. N. Khushaba, M. Takruri, J. V. Miro, and S. Kodagoda, "Towards limb position invariant myoelectric pattern recognition using time-dependent spectral features," *Neural Networks*, vol. 55, pp. 42–58, 2014.
- [91] V. Wellekens and C. Tyagi, "On desensitizing the Mel-Cepstrum to spurious spectral components for robust speech recognition," in *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005*, pp. 529–532, Philadelphia, PA, USA, 2005.
- [92] D. G. Childers, D. P. Skinner, and R. C. Kemerait, "The cepstrum: a guide to processing," *Proceedings of the IEEE*, vol. 65, no. 10, pp. 1428–1443, 1977.
- [93] J. He, D. Zhang, X. Sheng, S. Li, and X. Zhu, "Invariant surface EMG feature against varying contraction level for myoelectric control based on muscle coordination," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 3, pp. 874–882, 2014.
- [94] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks,"



- Advances in Neural Information Processing System*, vol. 25, pp. 1097–1105, 2012.
- [95] S. Hassantabar, M. Ahmadi, and A. Sharifi, “Diagnosis and detection of infected tissue of COVID-19 patients based on lung x-ray image using convolutional neural network approaches,” *Chaos, Solitons & Fractals*, vol. 140, article 110170, 2020.
- [96] F. Milletari, N. Navab, and S. A. Ahmadi, “V-net: fully convolutional neural networks for volumetric medical image segmentation,” in *2016 Fourth International Conference on 3D Vision (3DV)*, pp. 565–571, Stanford, CA, USA, 2016.
- [97] R. Zhang, L. Bai, J. Zhang, L. Tian, R. Li, and Z. Liu, “Convolutional LSTM networks for vibration-based defect identification of the composite structure,” in *2020 International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD)*, pp. 390–395, Xi’an, China, 2020.
- [98] R. Liu, S. Wei, Y. Zhao, and Y. Yang, “Indexing of the CNN features for the large scale image search,” *Multimedia Tools and Applications*, vol. 77, no. 24, pp. 32107–32131, 2018.
- [99] S. H. Park, J. M. Goo, and C. H. Jo, “Receiver operating characteristic (ROC) curve: practical review for radiologists,” *Korean Journal of Radiology*, vol. 5, no. 1, p. 11, 2004.
- [100] H. H. Jasper, “The ten-twenty electrode system of the international federation,” *Electroencephalography and Clinical Neurophysiology*, vol. 10, pp. 371–375, 1958.