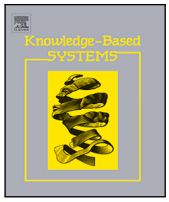
Knowledge-Based Systems 301 (2024) 112322

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys



MD-DCNN: Multi-Scale Dilation-Based Deep Convolution Neural Network for epilepsy detection using electroencephalogram signals

Mohan Karnati [a], Geet Sahu [b], Akanksha Yadav [c], Ayan Seal [c,d,e,f,g,*], Joanna Jaworek-Korjakowska [d], Marek Penhaker [e], Ondrej Krejcar [e,f,g]

[a] Department of Computer Science and Engineering, National Institute of Technology Raipur, Raipur, Chhatisgarh, 492010, India
[b] Department of Computer Science and Engineering, Siksha 'O' Anusandhan (Deemed to be University), Odisha, 751020, India
[c] Department of Computer Science and Engineering, PDPM Indian Institute of Information Technology, Design and Manufacturing Jabalpur, Jabalpur, Madhya Pradesh, 482005, India
[d] Department of Automatic Control and Robotics, and Center of Excellence in Artificial Intelligence, AGH University of Krakow, Kraków, 30059, Poland
[e] Faculty of Electrical Engineering and Computer Science (FEEC) VSB Technical University of Ostrava 17 Listopadu 2172/15, Ostrava, Czech Republic
[f] Center for Basic and Applied Science, Faculty of informatics and management, University of Hradec Kralove, Rokitanskeho 62, Hradec Kralove, 50003, Czech Republic
[g] Skoda Auto University, Na Karmeli 1457, 293 01 Mlada Boleslav, Czech Republic

ARTICLE INFO

Keywords:
Deep convolutional neural network
Epilepsy disease
Electroencephalography
Brain-computer interface

ABSTRACT

Approximately 65 million individuals experience epilepsy globally. Surgery or medication cannot cure more than 30% of epilepsy patients.However, through therapeutic intervention, anticipating a seizure can help us avoid it. According to previous studies, aberrant activity inside the brain begins a few minutes before the onset of a seizure, known as a pre-ictal state. Many researchers have attempted to anticipate the pre-ictal condition of a seizure; however, achieving high sensitivity and specificity remains challenging. Therefore, deep learning-based early diagnostic tools for epilepsy therapies using electroencephalogram (EEG) signals are urgently needed. Traditional methods perform well in binary epilepsy scenarios, such as normal vs. ictal, but poorly in ternary situations, such as ictal vs. normal vs. inter-ictal. This study proposes a multi-scale dilated convolution-based network (MD-DCNN) to predict seizures or epilepsy. Traditional DCNNs for epilepsy classification overfit due to insufficient training data (fewer subjects). Windowing 2-sec EEG recordings and extracting the frequency sub-band from each window prevents overfitting in deep networks, which lack training data. We convert each segmented window and its sub-bands into scalogram images and input them into MD-DCNN. The proposed MD-DCNN combines data from several scales without narrowing the acquisition domain. Integrating detailed information into high-level semantic features improves network interpretation and classification. The proposed MD-DCNN is evaluated for two-class, three-class, and cross-database strategy problems using three publicly accessible databases. Experiments show that the MD-DCNN statistically performs better than 13 other current approaches. This demonstrates its potential for developing equipment capable of measuring, monitoring, and recording EEG signals to diagnose epilepsy.

1. Introduction

Approximately 65 million individuals experience epilepsy globally, a frequent neurological disorder characterized by the sudden, irregular discharge of brain neurons regularly [1]. Seizures are the main symptoms of epilepsy. Unprovoked seizures, coupled with rapid abnormal neuronal discharges in the brain [2], are the main characteristics of epilepsy that can cause various health concerns, including cognitive disorders and death [3]. The quality of life and social standing of a person could be affected. An inexpensive, non-invasive method of monitoring, identifying, and diagnosing epilepsy is an electroencephalogram (EEG) of the scalp, which records the electrical activity of brain impulses over time [4]. The information in the electrophysiological signals generated by the neurons represents the overall brain activity in the cerebral cortex. EEG recording mostly contains signals from multi-channels that monitor the brain's electrical activities by placing several electrodes at different locations on the scalp. Recent studies have shown that two

* Corresponding author at: Department of Computer Science and Engineering, PDPM Indian Institute of Information Technology, Design and Manufacturing Jabalpur, Jabalpur, Madhya Pradesh, 482005, India.
E-mail address: ayan@iiitdmj.ac.in (A. Seal).

https://doi.org/10.1016/j.knosys.2024.112322
Received 21 February 2024; Received in revised form 12 July 2024; Accepted 31 July 2024
Available online 2 August 2024

methods exist for recording EEG signals, namely non-invasive scalp EEG and invasive intracranial EEG. The former is preferred over the second in most EEG seizure detection investigations, as the second one records signals by surgically implanting electrodes directly onto the surface of the brain [5]. Slight effort and time are required for an expert to visually analyze EEG patterns and recognize sicknesses, such as pre-ictal and ictal waves. Moreover, expert proficiency is required. Therefore, neurologists are under pressure to analyze the EEG brain waves of patients, which impedes their effectiveness. These constraints have prompted efforts to build and develop automated solutions to help neurologists differentiate between epileptic and non-epileptic EEG brain signals. As an alternative to manual seizure detection, intelligent epileptic EEG recognition has shown promise. Nonetheless, the nonstationary and non-linear nature of EEG waves makes this a challenging task [6]. Recent studies have published several articles on automatic EEG seizure identification as a classification problem [7].

Since Adeli et al. [8] were the first to develop automatic epilepsy detection utilizing EEG signals, several studies have begun investigating automatic epilepsy detection. Traditional epilepsy detection primarily employs machine learning (ML) approaches, split into two parts. Manually extracting characteristics from EEG waves is the initial part, for which non-linear [9], entropy [10], and complexity [11] are three widely used approaches. The second step involves sending the retrieved features to various classifiers for classification. Furthermore, principal component analysis (PCA) [12] is a statistical approach for exploring different features present in a signal through orthogonal transformation and has been widely used to reduce duplication in feature space and increase EEG epilepsy detection accuracy [13]. While traditional approaches successfully categorize EEG signals by extracting their characteristics in a specific way, they often require sophisticated methods to manually extract aspects of the EEG data, leading to significant information loss.

1.1. Motivation and contributions

Computer vision and data mining has witnessed a surge in the use of deep learning (DL), driving significant interest in employing it for EEG signal categorization [14–16]. Moreover, in the medical field, several methods have been employed using deep convolutional neural networks (DCNN) [17–26]. DL algorithms are more general in training better discriminative features than hand-crafted feature learning approaches. However, in medical settings, distinguishing between epileptic and non-epileptic EEG waves remains challenging. Typically, only a tiny percentage of epilepsy data is accessible for model training owing to the rarity of seizures. Additionally, noise and distortions in the data make it challenging to learn the neural activity associated with healthy, ictal, and non-ictal circumstances. This problem is exacerbated by variations in seizure patterns across patients [27]. Most existing strategies perform well for one problem but fall short when applied to others. For instance, they accurately distinguish seizure (ictal) vs. non-seizure (pre-ictal) cases but fail spectacularly in pre-ictal vs. ictal vs. inter-ictal circumstances [28]. Furthermore, most known approaches were tested using only one database. Moreover, the existing DL models for diagnosing epilepsy are far too simplistic. With less complicated databases, the straightforward approach works well. However, as complexity increases, its performance deteriorates (maybe owing to the underfitting issue). Hence, these strategies are unsuitable for use in the real world. Conversely, tweaking the hyperparameters of the architecture can enhance the performance of CNN-based methods. According to the above observations, epilepsy detection remains a challenging topic for three main reasons:

- Existing models perform well on a single database (train and test sets from the same database) but fail to perform well on other databases (trained on one database and tested on another). Thus, a lack of generalizability exists.
- Limited labeled data and low efficacy.
- Traditional methods perform well in binary epilepsy scenarios, such as normal vs. ictal, but poorly in ternary situations, such as ictal vs. normal vs. inter-ictal.

So, we need a generalized automated system that performs well enough with fewer training instances to help and support doctors and specialists [29]. Feature engineering determines the performance of an epilepsy system. Developing new features from existing ones can enhance the effectiveness of a system, encouraging further work in this direction. The primary contributions of this study are as follows:

- We conduct windowing for 2 s, followed by the frequency sub-band extraction of each window, to increase the data size and prevent deep networks from having insufficient training data. Then, the continuous wavelet transform (CWT) turns each segmented signal and its sub-bands into scalogram images, which we then use to train a new MD-DCNN model.
- The first two layers of the proposed MD-DCNN extract low-level features from scalogram images to improve the interpretation and classification capabilities of the network. Then, we feed these features into multi-scale dilated convolution blocks (MSDB) to aggregate multi-scale information without minimizing the acquiring domain. This process integrates the inherent detailed information into the high-level semantic facets.
- We investigate two-class, three-class, and cross-database evaluation approaches to determine the resilience of the proposed model. We compare the performance of the proposed strategy to 13 state-of-the-art (SOTA) approaches using various statistical classification measures such as precision, recall, and F1-score, addressing the accuracy paradox.
- We employ the cross-database strategy to check the generalizability of the proposed MD-DCNN. Regarding the three databases, the experimental results show that the proposed model performs exceptionally well and is more general than some of the baseline approaches.
- However, the performance metrics mentioned above alone are not sufficient to elucidate the proposed model. The MD-DCNN is lucid because of its interpretability, which solves the black box problem. Therefore, we visualize the t-distributed stochastic neighbor embedding (t-SNE) plots of the facet vectors of the proposed method on three datasets, along with the probability scores of five compared DL methods and the proposed MD-DCNN.
- To further investigate the proposed methodology, Wilcoxon's rank sum non-parametric statistical significance test for independent data items is performed at a 5% level of significance. The results of this experiment show that, in most situations, the MD-DCNN is statistically more significant than SOTA approaches.

The article follows this structure: we briefly review some existing literature on epilepsy categorization in Section 2. Section 3 illustrates the proposed technique. Section 4 presents the proposed model results and other baseline approaches. Finally, in Section 5, we conclude our research.

2. Related work

This section presents a comprehensive overview of the research conducted to formulate the proposed methodology, encompassing the utilization of DL and ML for epileptic seizure detection. The categorization of epileptic and normal EEG signals is challenging. Subsequently, a classification algorithm was applied to the discriminatory features extracted from EEG signals. In [30], handcrafted facets, like energy, standard deviation, and Shannon entropy, were retrieved. A general regression neural network classifier classified the aforementioned handcrafted feature on the University of Bonn database. The two-class

evaluation yielded good results. However, the three-class and cross-database evaluations were unsuccessful. Following the use of decision trees and linear support vector machine (SVM) classifiers, a few researchers reported using fast Fourier transforms (FFT) to fetch facets from single-channel EEG and multi-channel scalp EEG signals to efficiently identify epileptic seizures [31,32]. Researchers developed a short-time Fourier transform to address the inability of the FFT to fetch useful facets owing to insufficient temporal resolution. Linear kernel SVM and artificial neural networks were used in [33,34] to effectively sort epileptic seizures from EEG segments using SVM and rational discrete short-time Fourier transform (STFT), respectively. Then, wavelet transform (WT) was designed to overcome the constraints of STFT. A later study by [1,4,35,36] used dual-tree complex WT, harmonic wavelet packet transform, empirical WT, and discrete WT to get the differentiating information of EEG signals. This information was then fed into the SVM, random forest, relevance vector machine, and complex-valued neural network to detect epileptic seizures automatically. The Stockwell transform (ST) can mitigate the performance degradation of WT in noisy situations. ST is used with the k-nearest neighbor (kNN) classifier to obtain useful information from EEG signals so that epileptic seizures can be identified accurately [37]. Later, Zhang et al. [38] demonstrated the of genetic algorithm SVM and local mean decomposition to find epileptic seizures in EEG signals. In [39], the bandwidth or frequency spread of the filter was traditionally defined as the second-order moment of the frequency response of the filter to identify epileptic seizures. In [40], Fisher vector encoding based on gray-level co-occurrence matrix texture features and a sparse multi-scale radial basis function network were shown to be effective in the automatic diagnosis of epileptic convulsions. In [41], an automated system for identifying epileptic seizures was created using a signal processing technique known as complete ensemble empirical mode decomposition with adaptive noise on a single channel of EEG data. In [42], CNN was employed for feature extraction and classified using SVM. However, these methods performed well on a single database but failed to perform well on other databases when the complexity of the database increased. Moreover, conventional approaches sometimes require sophisticated methods to manually extract aspects of EEG data, leading to significant information loss.

Unlike handcrafted feature-based engineering, researchers have recently developed DL-based methods to extract unique elements from raw EEG data without utilizing any standard signal processing procedures. In [43–48], the CNN was presented for classifying images, voice sounds, and signal data with remarkable classification accuracy. Later, the DCNN was used to extract useful features from EEG recordings to help identify epileptic seizures [49,50]. To successfully find epilepsy, a pyramidal one-dimensional CNN was built to obtain useful information from EEG signals [51]. To find epileptic seizure events, wavelet packet decomposition with CNN and long-term recurrent CNN was used to correctly guess how the seizures would move [52–54]. In [54], Akyol et al. introduced extracting unique features from EEG data and feeding them into the fully connected layer for accurate epileptic event identification. Unlike CNN, DCNN, and long-short memory (LSTM), a deep convolutional autoencoder with an LSTM classifier and a semi-supervised stacking autoencoder was used to rebuild and categorize signals systematically [55,56]. Some researchers converted EEG signals into scalogram images, which were then fed into CNN [57]. In [58], a future fusion network, with the help of dilated CNN, was presented for classifying epilepsy. In [59], the EEG signals were converted into time-frequency images using continuous WT and fed into Visual Geometry Group 16 (VGG16) for detecting epilepsy. In [60], Xin et al. presented an attention mechanism-based wavelet CNN to detect epilepsy. In [61], Lih et al. introduced a transformer model for identifying epilepsy accurately using EEG signals. In [62], Albaqami et al. presented a multi-path seizure classification network (MP-SeizNet) comprising a CNN and a Bi-LSTM with an attention mechanism. Most of the above-mentioned DL-based methods require more training and testing time to produce satisfactory results owing to incorrect filter size selection, inappropriate initiation of different layers (more hidden layers), undesirable selection of the activation function, inappropriate activation size, and other factors. However, the methods mentioned above provided satisfactory performance on a single database and binary classification, but poor performance on multi-class classification and other databases. Therefore, existing methods lack generalizability.

3. Material and methods

3.0.1. Database-1: University of bonn database

This database [63] contained a single-channel EEG recording from five patients; five EEG sets were extracted and labeled A, B, C, D, and E. Each set had 100 EEG recordings, each lasting 23.6 s, with a sampling rate of 173.61 Hz. The data was obtained from a 128-channel data collection system, covering a spectral bandwidth range of 0.5–85 Hz. Sets A and B display the surface EEG of pre-ictal individuals with their eyes closed and open, respectively. Sets C and D represent intracerebral EEG recordings from epileptic subjects taken in the seizure-generating area and outside of it during seizure-free periods, respectively. Set E represents the intracranial EEG of an epileptic patient during an epileptic episode. Each set had 100 text files containing 4097 ASCII-coded samples from a single EEG time series.

3.0.2. Database-2: Neurology and sleep center, New Delhi EEG database

This database contained 5.12 s of EEG data [64]. A Grass Telefactor Comet AS40 amplification system with 57 EEG channels sampled at 200 Hz was used to record the data, with the spectral bandwidth ranging from 0.5 to 70 Hz. The time-series EEG recordings comprised three main MATLAB file folders: the ictal, pre-ictal, and inter-ictal stages. Each MAT file contained 1024 samples.

3.0.3. Database-3: Epileptic EEG database

This database contained multi-channel, long-term data on six focal epilepsy patients [65]. Before their epilepsy surgery, they were undergoing pre-operative tests. The data contained several seizure phases, including ictal, pre-ictal, inter-ictal, and onsets. The signals were recorded at a sampling rate of 500 Hz and stored as EDF files. The database contained labeled and classified data points (training and test sets) for complex partial electrographic and video-detected seizures. A band-pass filtering of 1–70 Hz was applied to all EEG signals, with a cutoff frequency of 50 Hz.

3.1. Proposed method

3.1.1. Preprocessing (Windowing)

Fig. 2 illustrates the step-by-step preprocessing process. We first preprocess the EEG signals from the previously described databases by taking a few steps. The first method is windowing, in which the channels of the EEG signal are split into 2-s frames with a 75% overlap, as shown in Fig. 2. Consequently, the dimensions of an EEG signal frame are (#channels ×2-s data points). Next, we used the standardization procedure to normalize each EEG frame to a zero mean and a unit variance. Conventional ML and DL models perform better and are more accurate when EEG inputs are normalized. The following justifies the selection of a 2-s time window with a 75% overlap:

- Two seconds have 348 samples at a frequency rate of 173.61 Hz, which is sufficient to denote frequency components, according to the Nyquist theorem. Consequently, all of the required properties may be extracted by the network [66].
- After segmentation, the increased database is sufficient for training, testing, and validating a binary and categorical classification.
- The 75% overlap can increase the number of data points while removing any phase shift influence in the data without using any extra data augmentation techniques [66].

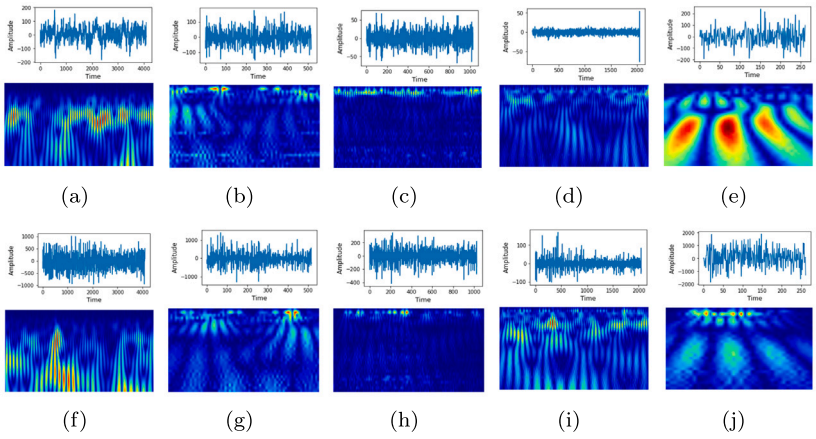


Fig. 1. The pictorial representation of raw EEG signals, frequency bands, and their scalogram images for healthy and epilepsy samples. The first row represents the healthy patient and their scalogram data, whereas the second row indicates the data of epilepsy patients and their scalograms. (a) Original signals (b) alpha band (c) beta band (d) gamma band (e) theta band (f) Original signals (g) alpha band (h) beta band (i) gamma band (j) theta band.

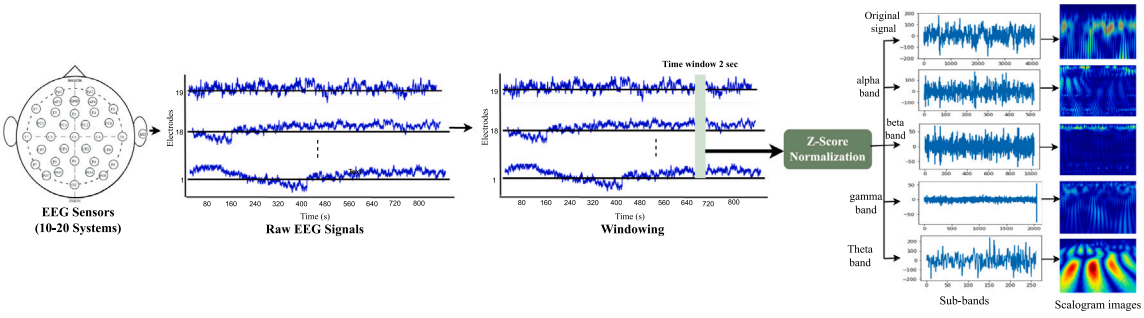


Fig. 2. EEG signals preprocessing steps.

Table 1
Statistical information of the databases.

Database	Number of samples					Total
	Two-class		Three-class			
	pre-ictal	ictal	pre-ictal	ictal	inter-ictal	
Before windowing						
Database-1	400	100	200	200	100	500
Database-2	50	100	50	50	50	150
Database-3	3895	3895	–	–	–	7790
After windowing						
Database-1	6000	1500	3000	3000	1500	7500
Database-2	100	200	100	100	100	300
Database-3	3895	3895	–	–	–	7790

Table 1 displays the number of samples generated after windowing. The total number of samples available before and after the windowing of three databases is represented in Table 1. Since database-3 collects signals at 500 Hz and each channel has 500 points, the windowing strategy is not employed.

3.1.2. Feature extraction using wavelet decomposition

Extracting more relevant information from the temporal and time-frequency domains is challenging to extract from the non-stationary and non-linear EEG data is challenging. In this study, we apply WT approaches to enhance the time-frequency analysis and effective time. WT is used to extract information from each segmented EEG data. The raw EEG signal is windowed to produce a larger set of EEG data samples. In [67], different regions were considered to increase the training data to avoid overfitting, inspiring the decision to crop the segmented signals into four frequency bands, which prevented overfitting by increasing the training data volume for the proposed model. The WT-based method divides the segmented EEG signals into four sub-frequency zones using a mother wavelet of fourth-order Daubechies. These are alpha, beta, gamma, and theta. Four frequency regions of theta (θ), alpha (α), beta (β), and gamma (γ) are employed for a more accurate analysis of the EEG signal. The selected frequency filter regions were 4–8 Hz, 8–16 Hz, 16–32 Hz, 32–64 Hz, and 64–128 Hz, respectively. We examine these frequency ranges to ensure sufficient data for the proposed network to distinguish between pre-ictal and ictal patients. Brain waves in the theta band indicate a disoriented, sleepy mind frame linked to mental inefficiency. Theta brain wave activity, which represents the transitional region between awake and sleep, is extremely relaxed when it occurs at relatively low levels. The slower, more expansive alpha band is linked to a relaxed mood and symbolizes the brain going into standby mode, ready to react when necessary.Alpha brain waves increase when we close our eyes and begin picturing something peaceful. The beta band comprises small, faster brainwaves associated with a state of mental and intellectual activity and outwardly focused concentration. We are often in this state of vigilance. The quickest and most subdued brain wave is the gamma band. Gamma rhythms alter consciousness and perception. Consequently, the proposed network has sufficient information (all subbands along with the original signal) to learn significant features relevant to individual classes, free from the overfitting problem.

3.1.3. Continuous wavelet transform (CWT)

With CWT, the time-frequency evaluation of a non-stationary signal has proven to be particularly successful. Because EEG signals are susceptible to noise, current research has focused on translating signals into a frequency domain, which is more efficient for evaluating noisy

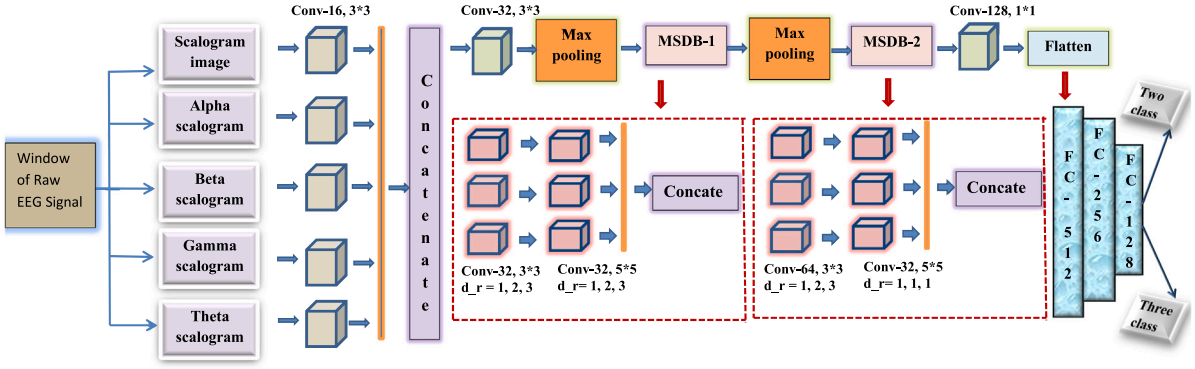


Fig. 3. The architecture of the proposed network.

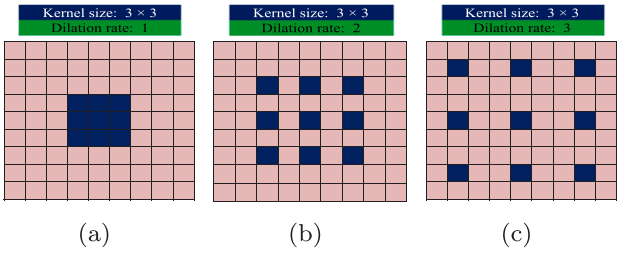


Fig. 4. Convolutional kernel size of 3 × 3 with dilation rates of 1, 2, and 3. (a) Kernel size 3 × 3 with dilation rate 1; (b) Kernel size 3 × 3 with dilation rate of 2; and (c) Kernel size 3 × 3 with dilation rate of 3.

signals. A wavelet transforms the 1-D signal from the spatial to the frequency domain, yielding a 2-D matrix for multiple-resolution studies. The spatial-domain representation of the signal is limited because it cannot provide an indefinite description. Currently, frequency-time analysis allows us to observe a signal's frequency and phase dispersion, simplifying the description and analysis of complex signals. Eq. (1) represents the mathematical measure of CWT.

$$CWT_s^{\phi}(\sigma,\tau) = \int_{-\infty}^{+\infty} s(x)\frac{1}{\sqrt{\sigma}}\phi(\frac{x-\tau}{\sigma})dt, \tag{1}$$

where τ and $s(x)$ are the translation parameter and the real signal, respectively. $\sigma > 0$ is the scale parameter, which is set to the range of 1–128, and $\phi(x)$ is the mother wavelet function (Complex Morlet wavelet), used to generate the wavelets for different scales and translations. Here, τ denotes the time shift for the translation, which could be considered the time interval around which the signal is processed. Essentially, τ takes on the values in the x array. For each time point in x, the wavelet is centered at that point, and the CWT is computed. CWT provides extensive information regarding the signal at approximately the instant, that is, higher frequency components with smaller σ values; conversely, with a higher σ, CWT delivers lesser frequency at components at approximately the instant. A two-dimensional image called a scalogram is employed to describe the square of CWT, $mid{-}CWT_s^{\phi}(\sigma,\tau)\,|^2$. Since the examined signal is digital, a discrete approximation of $CWT_s^{\phi}(\sigma,\tau)$ is computed [68]. We visualize the approximated scalogram using a matrix, where rows and columns represent different scales and translation parameters τ, respectively. However, in the time-frequency representation, frequency is more conventional than scale. We converted scale to frequency using the formula $f = \frac{1}{\sigma}$. The pictorial representation of original and frequency band signals and their scalogram images, is shown in Fig. 1. For further processing, these scalogram images are resized into 128 × 128.

3.2. Proposed MD-DCNN

Although the DCNN is used for feature extraction on EEG scalogram images [57,69–71], only the basic construction of a DCNN is investigated. We present a multi-scale dilated convolutional network for strong feature learning fully employing the feature learning power of various CNNs with various topologies. The proposed MD-DCNN structure is shown in Fig. 3. The MD-DCNN network has two main parts that work together to handle changes in feature scaling in scalogram images: a CNN for basic feature extraction and an MSDB block for dealing with scale change issues. The MSDB module uses dilated convolution to effectively combine multi-scale information without narrowing the acquisition domain. This turns important data into high-level semantic features that improve the ability of the network to understand and sort data. In dilated convolution, we use a "perforated" filter to successively convolve and concatenate the layers. This layer increases the receptive region while reducing the number of parameters and computational costs. In other words, we can create a larger receptive region with the same number of layers we used without incorporating more operations than standard convolution. The kernel size with the $p \times p$ filter is enhanced to $p + (p-1)(d_r - 1)$ with the dilation rate d_r in dilated convolution. Consequently, this enables the combination of multi-scale, relevant information while maintaining the same resolution. Convolution yields a 3 × 3 receptive region, while the dilated convolution with 2 and 3 dilation rates yields a 5 × 5 and a 7 × 7 receptive region, respectively, as shown in Fig. 4. The receptive region of a 3 × 3 convolution kernel with a dilation rate 2 is the same as that of a 5 × 5 convolution kernel; however, the number of learnable parameters is only 9 for the former, which is 36% of the 36% of the 5 × 5 kernel parameters. Similarly, for kernel size of 3 × 3 and a dilation rate of 3, the receptive region is the same as 7 × 7; however, the learnable parameters are also 9, representing 18% of the 7 × 7 kernel parameters.

3.2.1. Network structure

The proposed MD-DCNN consists of seven convolutional, two-max-pooling, two MSDB blocks, two fully connected layers, and a softmax layer. Initially, we apply a convolutional operation to scalogram images of four frequency bands and the original scalogram images to extract the low features, such as edges and lines. Next, we apply the concatenating operation to combine the extracted features. This will add the extracted features to the channel dimensions. Assuming the nth input scalogram image is S_n, the latent model output O_j of the convolutional layer is written as Eq. (2).

$$O_j = \sum_n W_{nj} * S_n + b_j, \tag{2}$$

where $*$ indicates the 2-D convolution operation, W_{nj} is the learnable kernel, i.e., weight connecting the input scalogram image S_n to the output O_j, and b_j is the learnable bias. After the convolution layer, a

max-pooling layer with a size of 2 × 2 is employed to reduce redundant feature information and noise. We then apply two MSDB blocks for noise invariance and robust feature extraction. We apply one max-pooling layer after the first MSDB block to reduce the dimensionality of facet maps. Finally, a convolutional layer with a kernel size of 1 × 1 is employed, which acts as an identity function. We then pass the obtained facet maps to fully connected layers to reduce dimensionality, extract meaningful features, and perform classification. The fully connected layers employ a softmax layer to learn the probabilities of a scalogram image sample for each class. The softmax layer produces a probability vector with the same dimension as the number of classes. This study examines two types of epileptic categorization problems: pre-ictal vs. ictal detection and pre-ictal vs. inter-ictal vs. ictal classification. Hence, the dimensions of the probability vector of the softmax layer are 2 and 3, respectively. The outcome of the softmax layer is written as Eq. (3).

$$P_k = \frac{e^{pk}}{\sum_m e^{pm}}, \tag{3}$$

where m (= 1, 2, …, M) is the class index and M indicates the total number of classes, pk is the input of the kth neuron of the softmax layer, and P_k is the associated probability to the kth class.

3.2.2. Activation function

The activation function is crucial in traditional artificial neural networks. The activation function regulates the activation level by translating its input into a specific range. Recently, sigmoid and hyperbolic tangent activation functions have dominated most shallow structural neural networks since the advent of DL. However, studies have shown that these activation functions face a saturation area problem. In large data applications, achieving remarkable outcomes often requires an extremely large DCNN with millions or even more parameters. The saturated region of standard activation functions enables the learning parameters with the gradient descendent-based error backpropagation method; however, this can lead to a model that is too quick and not very good. The rectified linear unit (ReLU), commonly known as the ramp function, was created for CNNs to address this shortcoming [72]. ReLU application is based on the idea that the information processing of biological neurons is frequently dispersed and sparse. Furthermore it has a high computing efficiency and is crucial for resolving the gradient vanishing problem. Researchers have previously created several versions of ReLU to enhance performance. For example, the Leaky ReLU was developed to propagate extremely small gradients to prevent zero gradients [73]. Therefore, this work employs the Leaky ReLU.

3.3. Parameter selection of MD-DCNN

We established the significant network architecture parameters of all the individuals to show that the network is ubiquitous, irrespective of how individual characteristics may change the optimal parameters of the network. Owing to time and computing constraints, we did not apply a complicated optimization technique to optimize the process parameters; instead, standard training was used to estimate the parameters.

- Adam optimization is chosen as the optimizer algorithm of the network, and the parameter is specified as the initial value.
- Categorical cross-entropy is chosen as a loss function for a three-class problem and binary cross-entropy is selected for a two-class problem.
- The learning rate is set to 0.0001.
- In each convolutional and fully connected layer, Leaky ReLU is employed as an activation function. The alpha value is set to 0.02.

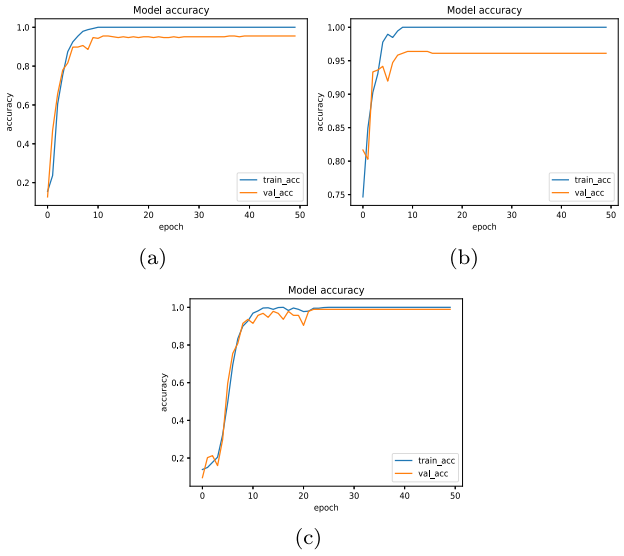


Fig. 5. Training and validation accuracy of the MD-DCNN on three databases for a two-class classification problem. From left to right, the graphs depict the accuracy for (a) Database-1, (b) Database-2, and (c) Database-3.

4. Experimental results & analysis

This section summarizes the results of the proposed method. Section 4.1 describes the experimental setup of the MD-DCNN. Further, Sections 4.2, 4.3, and 4.4 describe the two-class classification, three-class classification, and cross-database evaluation, respectively. Section 4.5 summarizes the compared results, and Section 4.6 details the ablation study. Finally, Section 4.7 describes the robustness of the MD-DCNN.

4.1. Experimental setup

Three types of experiments are performed in this study: two-class classification, three-class classification, and cross-database evaluation. In a two-class classification scenario, ictal vs. pre-ictal evaluation is performed on the MD-DCNN. Similarly, in three-class classification, the databases are divided into three classes viz., ictal, pre-ictal, and inter-ictal. In cross-database evaluation, training is performed in one database and tested on another database. The classification performance is examined using a tenfold cross-validation scheme and the average of all outcomes is used to calculate the final classification performance. All experiments are performed using Python and TensorFlow with 50 epochs and a mini-batch size of 32. The performance of the proposed method is evaluated using four traditional evaluation metrics, namely accuracy ($\mu_A(\%) = \frac{t_{tp}+t_{tn}}{t_{tp}+t_{tn}+t_{fp}+t_{fn}} \cdot 100$), F1-score ($\mu_{F1}(\%) = 2 \cdot \frac{\mu_P \cdot \mu_R}{\mu_P+\mu_R} \cdot 100$), precision ($\mu_P(\%) = \frac{t_{tp}}{t_{tp}+t_{fp}} \cdot 100$), and recall ($\mu_R(\%) = \frac{t_{tp}}{t_{tp}+t_{fn}} \cdot 100$), where t_{tp}, t_{tn}, t_{fp}, and t_{fn} correspondingly stand for "true positive", "true negative", "false positive", and "false negative".

4.2. Two-class classification

This experiment evaluates the MD-DCNN using three databases, each divided into ten equal parts through a ten-fold cross-validation procedure. This approach effectively removes sample bias and produces more comprehensive findings. We compare the proposed method to recent studies that use ML and DL to diagnose epilepsy, including [42, 52,57–62,74–78]. We evaluated most of these approaches on a single database, however, we also evaluated all of them on three databases to ensure consistency. Table 2 shows the results of the proposed approach and the SOTA methods.

Table 2
Comparison of the proposed method with 13 SOTA methods for a two-class classification problem.

Method	Year	Database-1				Database-2				Database-3			
		μ_A	μ_P	μ_R	μ_{F1}	μ_A	μ_P	μ_R	μ_{F1}	μ_A	μ_P	μ_R	μ_{F1}
Omer et al. [57]	2019	0.8708	0.87	0.87	0.87	0.5501	0.55	0.55	0.55	0.7728	0.80	0.77	0.77
Cao et al. [74]	2019	0.9403	0.94	0.94	0.94	0.8640	0.85	0.86	0.85	0.9004	0.90	0.90	0.90
Tian et al. [52]	2019	0.9556	0.96	0.96	0.96	0.8892	0.90	0.89	0.89	0.9375	0.94	0.94	0.94
Qin et al.[58]	2020	0.8503	0.85	0.85	0.85	0.9063	0.94	0.83	0.87	0.8996	0.90	0.90	0.90
Saidi et al. [42]	2021	0.9436	0.93	0.93	0.92	0.8805	0.88	0.87	0.87	0.7782	0.80	0.78	0.78
Rashed et al. [59]	2021	0.7084	0.35	0.50	0.41	0.6250	0.52	0.53	0.52	0.4902	0.48	0.50	0.48
Theekshana et al. [75]	2021	0.9573	0.96	0.96	0.96	0.8460	0.85	0.85	0.85	0.8948	0.90	0.89	0.89
Wang et al. [76]	2021	0.8427	0.86	0.84	0.85	0.8391	0.84	0.84	0.84	0.8806	0.88	0.88	0.88
Shoeibi et al. [77]	2022	0.8633	0.86	0.85	0.85	0.8019	0.80	0.80	0.80	0.8951	0.90	0.91	0.90
Gao et al. [78]	2022	0.9421	0.94	0.96	0.95	0.9008	0.92	0.90	0.91	0.9533	0.95	0.95	0.95
Xin et al. [60]	2022	0.9510	0.94	0.97	0.95	0.9281	0.93	0.93	0.93	0.9701	0.97	0.97	0.97
Lih et al. [61]	2023	0.9387	0.94	0.94	0.94	0.9410	0.95	0.92	0.93	0.9633	0.97	0.95	0.96
Albaqami et al. [62]	2023	0.9551	0.95	0.97	0.96	0.9316	0.93	0.93	0.93	0.9709	0.97	0.97	0.97
MD-DCNN	–	**0.9761**	**0.98**	**0.98**	**0.98**	**0.9611**	**0.97**	**0.95**	**0.96**	**0.9968**	**1.00**	**1.00**	**1.00**

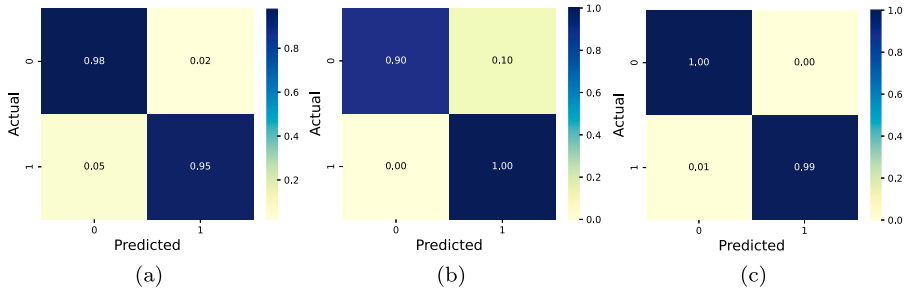


Fig. 6. Confusion matrices of the proposed methodology on three databases. From left to right, the confusion matrix of (a) Database-1, (b) Database-2, and (c) Database-3. Here, 0 and 1 indicate the ictal and pre-ictal databases.

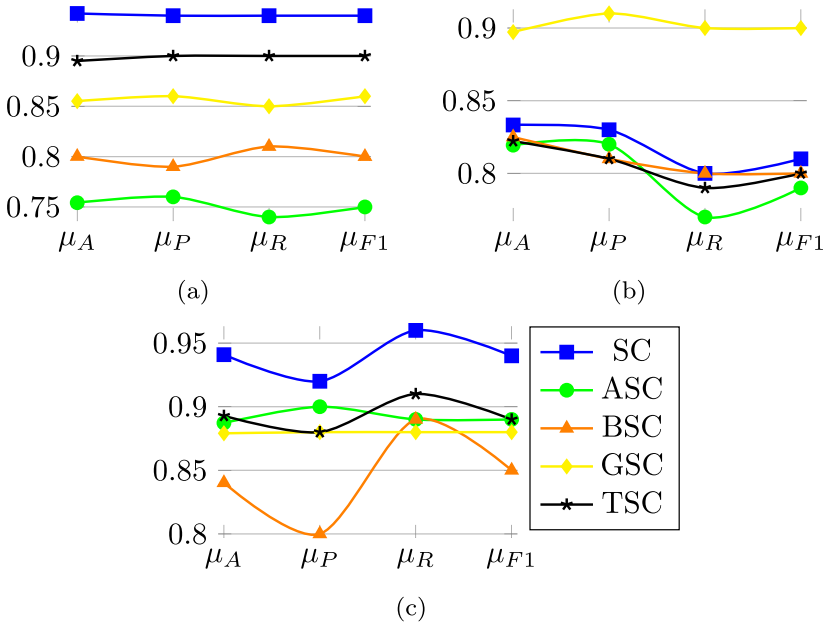


Fig. 7. The performance of the proposed method on individual frequency bands and original signal scalogram images of three databases in a two-class classification procedure. The graph represents (a) Database-1, (b) Database-2, and (c) Database-3. Here, SC → original signal scalogram, ASC → alpha scalogram, BSC → beta scalogram, GSC → gamma scalogram, and TSC → theta scalogram.

In terms of the four metrics, the result shows an average value of tenfold. For every database, the performance of the proposed method surpasses that of all other compared approaches. From Table 2, we can observe that most existing approaches worked effectively on one database but poorly on others. Here, the existing methods suffer from a generalizability problem. Furthermore, the training and validation plots of MD-DCNN are shown in Fig. 5. The proposed model exhibits strong generalization and resists overfitting issues. Furthermore, confusion matrices of the testing efficiency on the three databases are displayed in Fig. 6. The suggested MD-DCNN model worked well in all three databases, as shown in Fig. 6. The actual and predicted classes are represented by the vertical and horizontal dimensions, respectively, in Fig. 6. The diagonal values depict the classification rate of each class, whereas the $(i, j)^{th}$ off-diagonal values display the percentage

Table 3
Comparison of the proposed method with 13 SOTA methods for a three-class classification problem.

Reference	Year	Database-1				Database-2			
		μ_A	μ_P	μ_R	μ_{F1}	μ_A	μ_P	μ_R	μ_{F1}
Omer et al. [57]	2019	0.8603	0.86	0.87	0.86	0.5746	0.53	0.52	0.52
Cao et al. [74]	2019	0.8992	0.90	0.90	0.90	0.8403	0.84	0.85	0.84
Tian et al. [52]	2019	0.8002	0.80	0.80	0.80	0.7829	0.80	0.78	0.78
Qin et al.[58]	2020	0.8667	0.88	0.88	0.88	0.9337	0.93	0.92	0.93
Saidi et al. [42]	2021	0.9302	0.93	0.92	0.92	0.8802	0.88	0.88	0.88
Rashed et al. [59]	2021	0.7325	0.71	0.71	0.71	0.6428	0.51	0.54	0.48
Theekshana et al. [75]	2021	0.8092	0.81	0.81	0.81	0.8536	0.85	0.85	0.85
Wang et al. [76]	2021	0.8245	0.82	0.82	0.82	0.8449	0.86	0.83	0.84
Shoeibi et al. [77]	2022	0.7906	0.79	0.79	0.79	0.8118	0.82	0.81	0.81
Gao et al. [78]	2022	0.8998	0.90	0.90	0.90	0.8911	0.89	0.89	0.89
Xin et al. [60]	2022	0.9133	0.92	0.90	0.91	0.9001	0.90	0.90	0.90
Lih et al. [61]	2023	0.8986	0.90	0.90	0.90	0.9026	0.90	0.90	0.90
Albaqami et al. [62]	2023	0.8862	0.88	0.85	0.86	0.8704	0.87	0.87	0.87
MD-DCNN	–	**0.9542**	**0.95**	**0.95**	**0.93**	**0.9445**	**0.94**	**0.94**	**0.94**

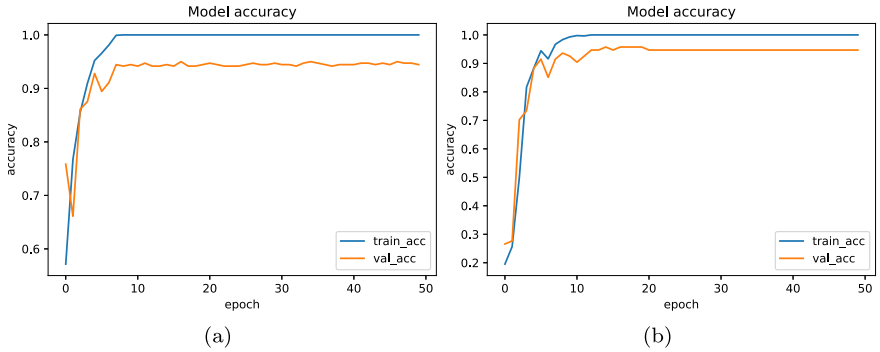


Fig. 8. Training and validation accuracy of the MD-DCNN on two databases for a three-class classification problem, (a) shows the accuracy on Database-1 and (b) shows the accuracy for Database-2.

of samples from the ith category that are incorrectly categorized to the jth class. We also tested the suggested network using separate sub-bands and the original signal scalogram images shown in Fig. 7 for all three databases. The original signal scalogram images obtained better performance on databases 1 and 3 when compared to frequency-band scalogram images; however, for database 2, the gamma-band scalograms obtained higher accuracy than others.

4.3. Three-class classification

The three-class classification method evaluates the proposed MD-DCNN using three databases comprising three classes, namely pre-ictal, inter-ictal, and ictal. Here, a similar procedure of two-class classification is followed, and the outcomes are reported in Table 3. From Table 3, we can see that the proposed method outperforms the existing approaches in a three-class classification scenario. A large margin was observed between the proposed method and the existing methods, except for Rashed et al. [59]. Database-3 is not considered in the experiment because it does not contain an inter-ictal class. Furthermore, the training and validation accuracy of MD-DCNN on three-class classification is depicted in Fig. 8. Furthermore, Fig. 9 shows the confusion matrices of the proposed methodology on two databases. The suggested network is also tested on each sub-band and on the original signal scalogram image. Fig. 10 shows the results for both Database-1 and Database-2. It can be stated from Fig. 10 that original signal scalogram images obtained better performance for Database-1 and gamma-band scalograms obtained higher efficiency for Database-2. However, Databases 1 and 2 perform very poorly in alpha-band scalograms.

4.4. Cross-database evaluation

After training the model on one database (the source database), this approach involves testing it on multiple databases (the target database). The trained models must perform well in all scenarios and not have overfitting issues to be used in therapeutic settings in the future. Therefore, a validation database is required to estimate the behavior of the system in real time. Cross-database shows similar accuracy on the test database compared to the results of within- database tests (two-class and three-class classifications). This indicates that the trained models can adapt to new situations and not suffer too much from over-learning. As shown in Tables 4 and 5, the MD-DCNN performed well compared to other existing approaches. We have conducted two experiments here: In the first experiment, we use Database-1 for training and use Database-2 and Database-3 for testing. Table 4 provides a full summary of the experimental findings. In the second experiment, we test the proposed approach on Database-1 and Database-2, and train it on Database-3. The empirical results are shown in Table 5. We solely use Database-2 for testing owing to its smaller sample size.

In addition, Fig. 11 shows the individual analysis of the MD-DCNN with sub-bands and the original signal scalogram images. Database-2 and Database-3 are considered for testing, while database-1 is used for training, as shown in the first row of Fig. 11. Similarly, Database-3 is employed for training, whereas Database-1 and Database-2 are used for testing, as seen in the second row of Fig. 11. The original signal scalogram-based MD-DCNN performed well compared to other sub-band scalogram images. However, when the original signal was combined with sub-band scalograms, the proposed MD-DCNN worked better than other SOTA methods.

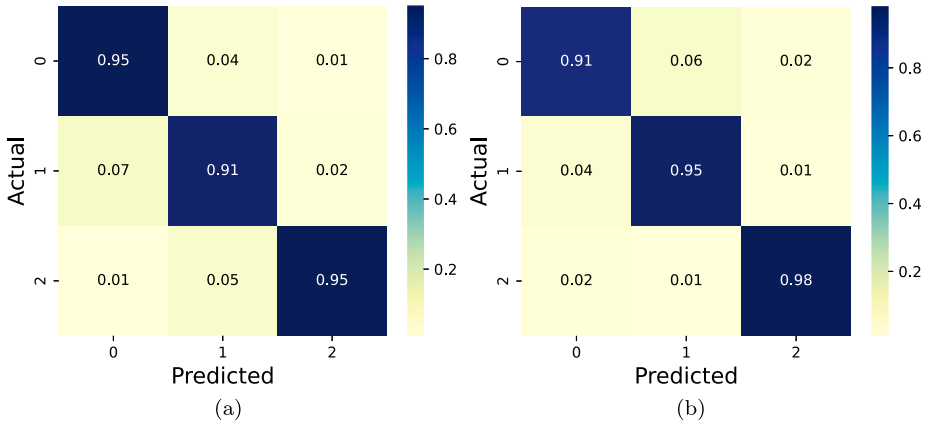


Fig. 9. Confusion matrices of the proposed methodology on two databases: (a) Database-1 and (b) Database-2, for three-class classification. Here, 0, 1, and 2 represent pre-ictal, inter-ictal, and ictal.

Table 4
Comparative analysis of the proposed method with 13 SOTA methods. Database-1 is used for training and is tested on Database-2 and Database-3. Here, Database-3 includes two-class problems only.

Reference	Year	Database-2								Database-3			
		Two-class				Three-class				Two-class			
		μ_A	μ_P	μ_R	μ_{F1}	μ_A	μ_P	μ_R	μ_{F1}	μ_A	μ_P	μ_R	μ_{F1}
Omer et al. [57]	2019	0.7837	0.78	0.78	0.78	0.6621	0.65	0.67	0.66	0.7539	0.75	0.75	0.75
Cao et al. [74]	2019	0.8329	0.83	0.83	0.83	0.8274	0.83	0.83	0.83	0.8001	0.80	0.80	0.80
Tian et al. [52]	2019	0.7420	0.75	0.74	0.74	0.5905	0.59	0.59	0.59	0.6611	0.66	0.66	0.66
Qin et al.[58]	2020	0.8163	0.82	0.82	0.82	0.8009	0.80	0.81	0.80	0.8328	0.83	0.83	0.83
Saidi et al. [42]	2021	0.6636	0.66	0.66	0.66	0.5291	0.54	0.52	0.52	0.5990	0.60	0.60	0.60
Rashed et al. [59]	2021	0.7101	0.71	0.71	0.71	0.7244	0.73	0.74	0.73	0.8011	0.80	0.80	0.80
Theekshana et al. [75]	2021	0.6993	0.70	0.70	0.70	0.7433	0.75	0.74	0.74	0.8333	0.83	0.83	0.83
Wang et al. [76]	2021	0.7836	0.76	0.79	0.77	0.7511	0.75	0.74	0.74	0.8009	0.80	0.80	0.80
Shoeibi et al. [77]	2022	0.7611	0.76	0.76	0.76	0.7208	0.72	0.72	0.72	0.7816	0.78	0.79	0.78
Gao et al. [78]	2022	0.7748	0.77	0.77	0.77	0.7287	0.73	0.74	0.73	0.7918	0.79	0.81	0.80
Xin et al. [60]	2022	0.8413	0.84	0.84	0.84	0.8016	0.81	0.80	0.90	0.8406	0.84	0.84	0.84
Lih et al. [61]	2023	0.7933	0.79	0.79	0.79	0.7462	0.75	0.75	0.75	0.7009	0.70	0.70	0.70
Albaqami et al. [62]	2023	0.8048	0.80	0.80	0.80	0.7916	0.79	0.80	0.79	0.8261	0.83	0.83	0.83
MD-DCNN	–	**0.9092**	**0.91**	**0.91**	**0.91**	**0.8805**	**0.87**	**0.89**	**0.88**	**0.8991**	**0.90**	**0.90**	**0.90**

Table 5
Comparative analysis of the proposed method with other SOTA methods. Database-3 is used for training and testing on Database-1 and Database-2. Here, only a two-class classification problem is analyzed since Database-3 contains two classes.

Reference	Year	Database-1				Database-2			
		Two-class				Two-class			
		μ_A	μ_P	μ_R	μ_{F1}	μ_A	μ_P	μ_R	μ_{F1}
Omer et al. [57]	2019	0.6113	0.59	0.61	0.60	0.7743	0.77	0.77	0.77
Cao et al. [74]	2019	0.8118	0.81	0.81	0.81	0.8404	0.84	0.84	0.84
Tian et al. [52]	2019	0.8441	0.85	0.84	0.84	0.8551	0.86	0.86	0.86
Qin et al.[58]	2020	0.8344	0.83	0.83	0.83	0.8661	0.87	0.88	0.87
Saidi et al. [42]	2021	0.5560	0.56	0.57	0.56	0.6722	0.68	0.67	0.67
Rashed et al. [59]	2021	0.7810	0.78	0.78	0.78	0.7008	0.69	0.70	0.69
Theekshana et al. [75]	2021	0.7911	0.79	0.79	0.79	0.8333	0.83	0.83	0.83
Wang et al. [76]	2021	0.7593	0.76	0.77	0.76	0.7816	0.78	0.78	0.78
Shoeibi et al. [77]	2022	0.7363	0.76	0.74	0.75	0.7556	0.76	0.78	0.77
Gao et al. [78]	2022	0.7493	0.77	0.75	0.76	0.7810	0.78	0.78	0.78
Xin et al. [60]	2022	0.8661	0.87	0.87	0.87	0.8592	0.86	0.86	0.86
Lih et al. [61]	2023	0.8509	0.85	0.85	0.85	0.8716	0.87	0.87	0.87
Albaqami et al. [62]	2023	0.8493	0.85	0.85	0.85	0.8774	0.88	0.88	0.88
MD-DCNN	–	**0.8961**	**0.90**	**0.90**	**0.90**	**0.9011**	**0.90**	**0.90**	**0.90**

4.5. Discussion

We compare the suggested MD-DCNN to existing SOTA techniques to determine its effectiveness and excellence. This study explores three types of assessments: two-class classification, three-class classification, and cross-database evaluation. According to the results, MD-DCNN outperformed all other methods in every database, with its performance nearly identical to that of the existing approach. However, as Tables 3, 4, and 5 demonstrate, the existing approaches do not perform well in three-class classification and cross-database evaluation. According to the previously described findings and discussion, the MD-DCNN can extract more representative and hidden information by reducing cross-term interference of time and frequency. When we compare hand-crafted feature extraction and classification with simultaneous deep feature extraction and classification using CNN, it significantly reduces the manual effort. The suggested MD-DCNN network had the best

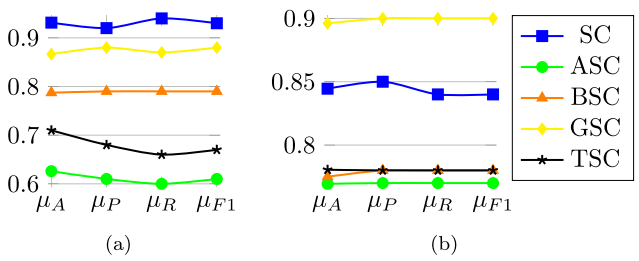


Fig. 10. The performance of the proposed method on the frequency band and original signal scalogram images of two databases: (a) Database-1 and (b) Database-2 in a three-class classification. Here, SC → original signal scalogram, ASC → alpha scalogram, BSC → beta scalogram, GSC → gamma scalogram, and TSC → theta scalogram.

Table 6
The performance analysis of the MD-DCNN with different window sizes.

Window size	Database	μ_A	μ_P (%)	μ_R (%)	μ_{F1} (%)
2 s	Database-1	98	98	98	98
	Database-2	96	97	95	96
	Database-3	100	100	100	100
4 s	Database-1	97	96	98	97
	Database-2	93	94	91	92
	Database-3	94	94	94	94
6 s	Database-1	97	97	97	97
	Database-2	92	90	93	91
	Database-3	96	97	95	96
8 s	Database-1	90	91	90	90
	Database-2	89	89	89	89
	Database-3	91	91	91	91

classification rate, even with just a few layers. The following are the principal advantages of our method:

- The resultant model is stable because it is created using tenfold cross-validation.
- The resulting model is precise because it is validated against three databases.
- The proposed method performs well even on small databases without leading to an overfitting problem.
- As far as we know, we are the first to achieve the best accuracy in three separate evaluations (two-class, three-class, and cross-database evaluation) for epilepsy identification, utilizing three public databases. Consequently, our methodology is clearly well-generalized compared with other SOTA methods.

4.6. Ablation study

Building a DCNN model from scratch for a specific problem is challenging, especially when considering a database with a limited number of images and data scarcity issues. An ablation study is necessary to fully understand the role of each component of the proposed MD-DCNN. Therefore, we conducted an ablation study to optimize the architecture of the proposed MD-DCNN and ensure its effective operation on unknown data. This subsection evaluates the impact of changing the model parameters on the performance of the suggested MD-DCNN. In the first experiment, we analyzed variations of hyperparameters, such as learning rates and activation functions. The empirical results of the MD-DCNN on various learning rates are shown in Fig. 12 for two-class, three-class, and cross-database strategies. Since Database-3 does not contain inter-class information, we consider databases 1 and 2 for three-class classification in all experiments to maintain consistency. For cross-database evaluation, we use databases 1 and 2 for training and evaluation on databases 2 and 3. The cross-database strategy only represents two-class classification results. Fig. 12 shows that our MD-DCNN performed well on all three evaluations at a learning rate of 0.0001. Moreover, we also evaluated the effect of activation functions in the hidden layers of MD-DCNN, and the examined outcomes are displayed in Fig. 13.

From Fig. 13, we can observe the performance of the MD-DCNN peak at the activation function of Leaky ReLU compared to other activation functions. We examine the effect of multi-scale dilation in the second experiment. The experimental results of MD-DCNN with and without multi-scale dilations are represented in Fig. 14. We can see from Fig. 14 that without MSDB blocks, MD-DCNN is less effective than the other existing methods. Therefore, the MSDB block of the MD-DCNN is crucial for extracting prominent features with a smaller number of parameters. In the third experiment, we analyzed the effect of the MD-DCNN with a varying number of MSDB blocks. The empirical results are shown in Fig. 15. From Fig. 15, we can see that the MD-DCNN with two MSDB blocks obtained higher performance, but after increasing the number of blocks, the performance degraded. As we can see from Fig. 15, the proposed MD-DCNN with four MSDB blocks obtained has a performance that is approximately lower than 80%, indicating overfitting. When four MSDB blocks are used in the MD-DCNN, the number of learnable parameters increases. This causes overfitting because the training data is not sufficient for all the parameters.

The fourth experiment evaluates the effect of various window sizes on MD-DCNN, and the results are reported in Table 6. Table 6 shows the performance for varied window sizes, with the 2-s window yielding the best result. Table 6 further shows that the proposed MD-DCNN performs worse for other window sizes. Here, only two-class classification results are mentioned in Table 6.

In the fifth experiment, we compare the spectrogram images from STFT, the scalogram images estimated from CWT, and the seizure images from segmented data. The seizure image is a color image created using the *pcolormesh* function in Python's Matplotlib package. The results are shown in Table 7, and the proposed model performs better on scalogram (CWT) images. It indicates that scalogram images can highlight more relevant features than seizure images of segmented data and a spectrogram. Here, only two-class classification results are mentioned in Table 7.

We analyzed various dilation rates in the MSDB block in the last experiment. The experimental results reported in Table 8. According to Table 8, the combination of dilation rates 1, 2, and 3 with the MSDB block achieved the best results compared to other combinations of dilation rates. However, other combinations also provided the nearest to the best combination of dilation rates. Here, only two-class classification results are mentioned in Table 8.

4.7. Robustness of the MD-DCNN

One cannot only rely on performance metrics to understand or interpret the proposed model. Interpreting the MD-DCNN makes it transparent and resolves the black box issue. Therefore, this section continues with two more tests. The probability vector produced by the softmax layer is assessed using current methodologies [75–78] to demonstrate the effectiveness of the feature learning capability of the proposed technique. The experiment validates its efficiency using Database-1, which contains samples from the pre-ictal, inter-ictal, and ictal stages. Fig. 16 shows the plotting of the softmax layer outputs produced by the proposed method and a comparison of other methods.

As shown in Fig. 16, the proposed strategy delivers the highest probabilities for all three correct classes, however, the maximum outputs from the other existing approaches are inconsistent with the appropriate classes for a sample. Furthermore, Fig. 17 presents the feature visualization of the MD-DCNN. We display the features in a 2-D graphic using a dimensional reduction approach called t-distributed stochastic neighbor embedding, which enables the visualization of a high-dimensional database in a low-dimensional context. From Fig. 17 (which belongs to two classes), a high correlation is observed between

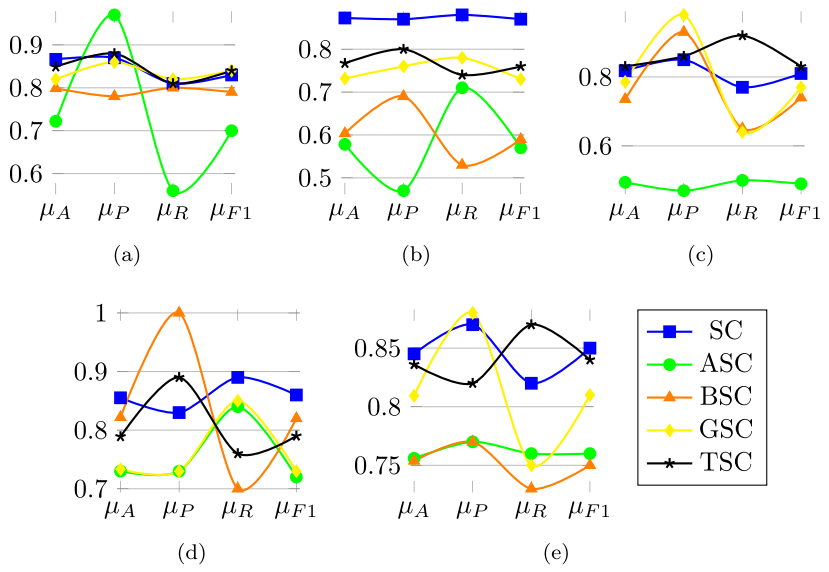


Fig. 11. The individual performance of the proposed method on frequency bands and original signal scalogram images. (a)–(b) indicates the two-class classification performance and (c) represents the three-class classification performance trained on Database-1 and tested on Database-2 and Database-3; (d)–(e) indicates the two-class classification performance trained on Database-3 and tested on Database-1 & 2. Here, SC → original signal scalogram, ASC → alpha scalogram, BSC → beta scalogram, GSC → gamma scalogram, and TSC → theta scalogram.

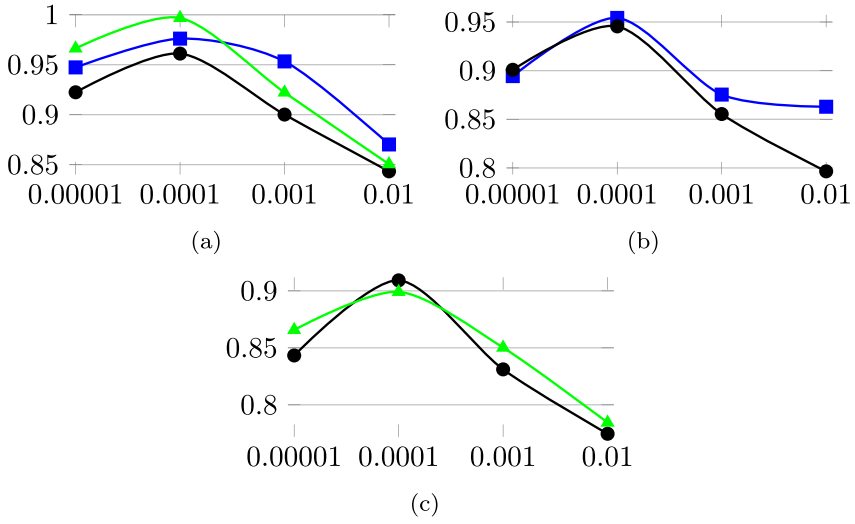


Fig. 12. The performance of the MD-DCNN in terms of accuracy while varying the learning rates. (a) For two-class classification, (b) for three-class classification, and (c) for cross-database evaluation. Here, the blue, black, and green colors indicate Database-1, Database-2, and Database-3.

Table 7
The classification report of MD-DCNN on different images like spectrogram, scalogram, and seizure pictorial method representation.

Image Representation	Database	μ_A (%)	μ_P (%)	μ_R (%)	μ_{F1} (%)
Spectrogram	Database-1	95	97	95	96
	Database-2	91	91	91	91
	Database-3	97	98	96	97
Seizure pictorial representation	Database-1	96	96	96	96
	Database-2	93	95	92	93
	Database-3	96	96	96	96
Scalogram	Database-1	98	98	98	98
	Database-2	96	97	95	96
	Database-3	100	100	100	100

the original properties of the three databases. The MD-DCNN can distinguish between traits that belong to different classes. We employ complete training data to demonstrate the significance of the proposed network. Consequently, the proposed structure performs exceptionally well in terms of epilepsy diagnosis.

4.8. Statistical significance of MD-DCNN

For independent data objects, Wilcoxon's rank-sum, a non-parametric statistical significance test, is performed at a 5% significance level [79]. This experiment employs distinct test sets, which can

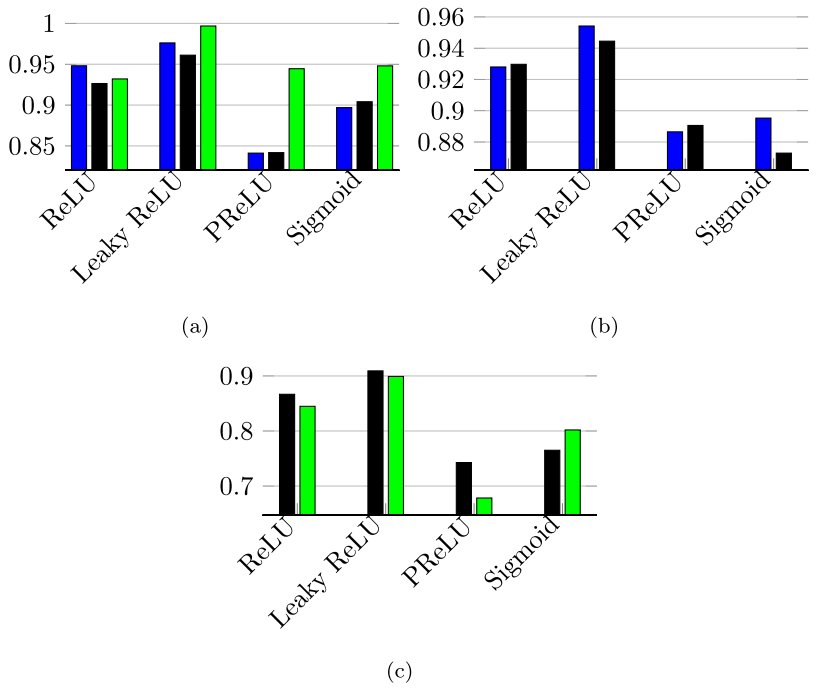


Fig. 13. The performance of the MD-DCNN in terms of accuracy with different activation functions. (a) For two-class classification, (b) three-class classification, and (c) cross-database evaluation. Here, the blue, black, and green colors indicate Database-1, Database-2, and Database-3, respectively.

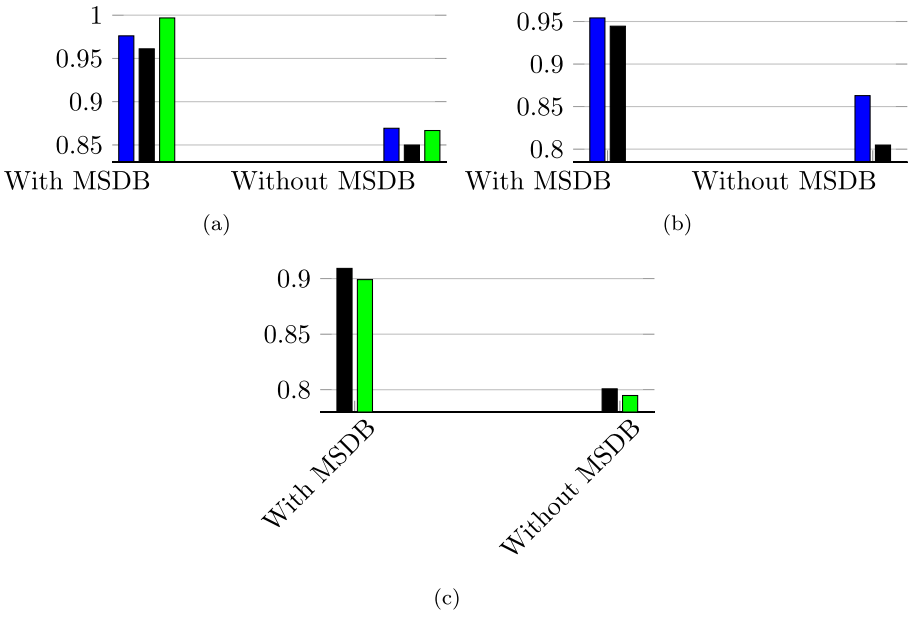


Fig. 14. The performance of the MD-DCNN in terms of accuracy with and without MSDB blocks (a) for two-class classification, (b) three-class classification, and (c) cross-database evaluation. Here, the blue, black, and green colors indicate Database-1, Database-2, and Database-3.

be considered unpaired data for the Wilcoxon rank-sum test, to validate each approach. The significance level serves as a signal of the caliber of the scientific backing of the data object before rejecting the null hypothesis and determining that the impact is statistically significant. For each database, we create 13 pairs, each containing the scores of classification assessment metrics determined through tenfold cross-validation. Table 9 provides the p-values that were determined using the Wilcoxon rank-sum test based on the accuracy results of the two- and three-class classification evaluations. Let E_j be the median value of every pair produced using a particular method, where $1 \leq j \leq 13$. The null hypothesis states that no statistically significant variation exists between an individual method and the median values of MD-DCNN. We consider the established hypothesis to be true until statistical evidence contradicts it and supports a different theory. Mathematically, $H_0 : E_1 = E_l$ vs. $H_a : E_1 > E_l$, where $l \in \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13\}$. For the most part, the p-values in Table 9 are below the 0.05 or 5% significance level in the three-class classification of Database-1 [42]. This means that strong evidence exists against the null hypothesis, demonstrating that the median values of the assessment metrics, as determined by MD-DCNN, are statistically significant and not the result of random variation. At a 5% confidence level, the null hypothesis is thus ruled out. Table 9 also shows that the proposed MD-DCNN does statistically perform better than the 13 SOTA techniques shown in Sections 4.2, 4.3, and 4.4.

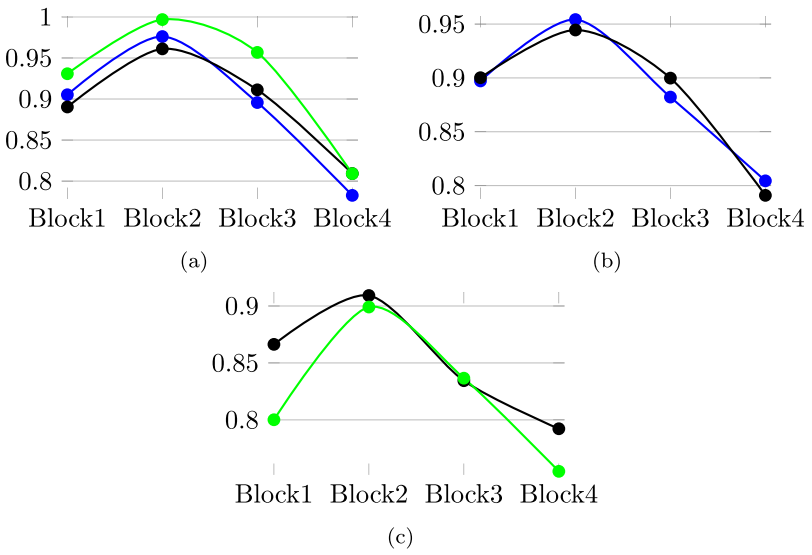


Fig. 15. The performance of the MD-DCNN in terms of accuracy while varying the number of MSDB blocks. (a) For two-class classification, (b) for three-class classification, and (c) for cross-database evaluation. Here, the blue color indicates database-1, the black color represents database-2, and the green color indicates database-3.

Table 8
The classification report of MD-DCNN on different combinations of dilation rates.

Dilation rates	Database	μ_A (%)	μ_P (%)	μ_R (%)	μ_{F1} (%)
1, 2, 3	Database-1	98	98	98	98
	Database-2	96	97	95	96
	Database-3	100	100	100	100
2, 3, 4	Database-1	97	97	97	97
	Database-2	94	95	94	94
	Database-3	97	98	96	97
3, 4, 5	Database-1	95	97	94	94
	Database-2	93	93	93	93
	Database-3	96	97	95	96
4, 5, 6	Database-1	90	90	90	90
	Database-2	91	92	90	91
	Database-3	91	91	91	91

5. Conclusion

This study proposes a dilated multi-scale DL model for detecting epilepsy using three publicly available EEG databases. The EEG signals are converted into scalogram images to train the MD-DCNN model. We train the model using publicly available databases classified into two or three classes. Furthermore, the proposed MD-DCNN is compared with 13 SOTA methods using various experiments, including two-class, three-class, and cross-database strategies. We perform feature visualization to explore the features learned for classification, and evaluate the softmax probability score to determine the consistency of the MD-DCNN. The study results show that the proposed model performs well in all experiments, and it also shows consistency when compared with other approaches. Our method can generalize well in all situations and does not overfit the training data. The findings might help neurologists diagnose epilepsy more quickly and accurately during screening. In the future, we plan to integrate this software system with the Internet of Things, enabling remote access for anyone. This study does not explicitly analyze the complexity of the proposed method. However, we believe it is high compared to some SOTA methods, depending on the number of frequency bands. Therefore, we could reduce the complexity of the proposed method by concentrating on fewer regions requiring further research.

CRediT authorship contribution statement

Mohan Karnati: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Geet Sahu:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Akanksha Yadav:** Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ayan Seal:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Conceptualization. **Joanna Jaworek-Korjakowska:** Writing – review & editing, Supervision, Project administration, Conceptualization. **Marek Penhaker:** Writing – review & editing, Project administration, Methodology, Conceptualization. **Ondrej Krejcar:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization.

Declaration of competing interest

The author(s) declare(s) that there is no conflict of interest.

Data availability

The authors do not have permission to share data.

Acknowledgments

This work was supported by the Program “Excellence Initiative Research University” for the AGH University, Krakow, and the ARTIQ Project, under Grant ARTIQ/0004/2021. This work was also supported in part by the European Union under the LERCO CZ.10.03.01/00/22_003/0000003 project via the Operational Programme Just Transition; in part by the Project “Smart Solutions in Ubiquitous Computing Environments”, (ID: UHK-FIMSPEV 2102-2024), University of Hradec Kralove, Faculty of Informatics and Management, Czech Republic. We are also grateful for the support of Ph.D. student Michal Dobrovolny for consultations. This work was also supported by the European Union under the LERCO CZ.10.03.01/00/22 003/0000003 project via the Operational Programme Just Transition and Project No. SP2024/071 Engineering systems XX. This work has also been partly supported by the project (2024/2204), Grant Agency of Excellence, University of Hradec Kralove, Faculty of Informatics and Management, Czech Republic.

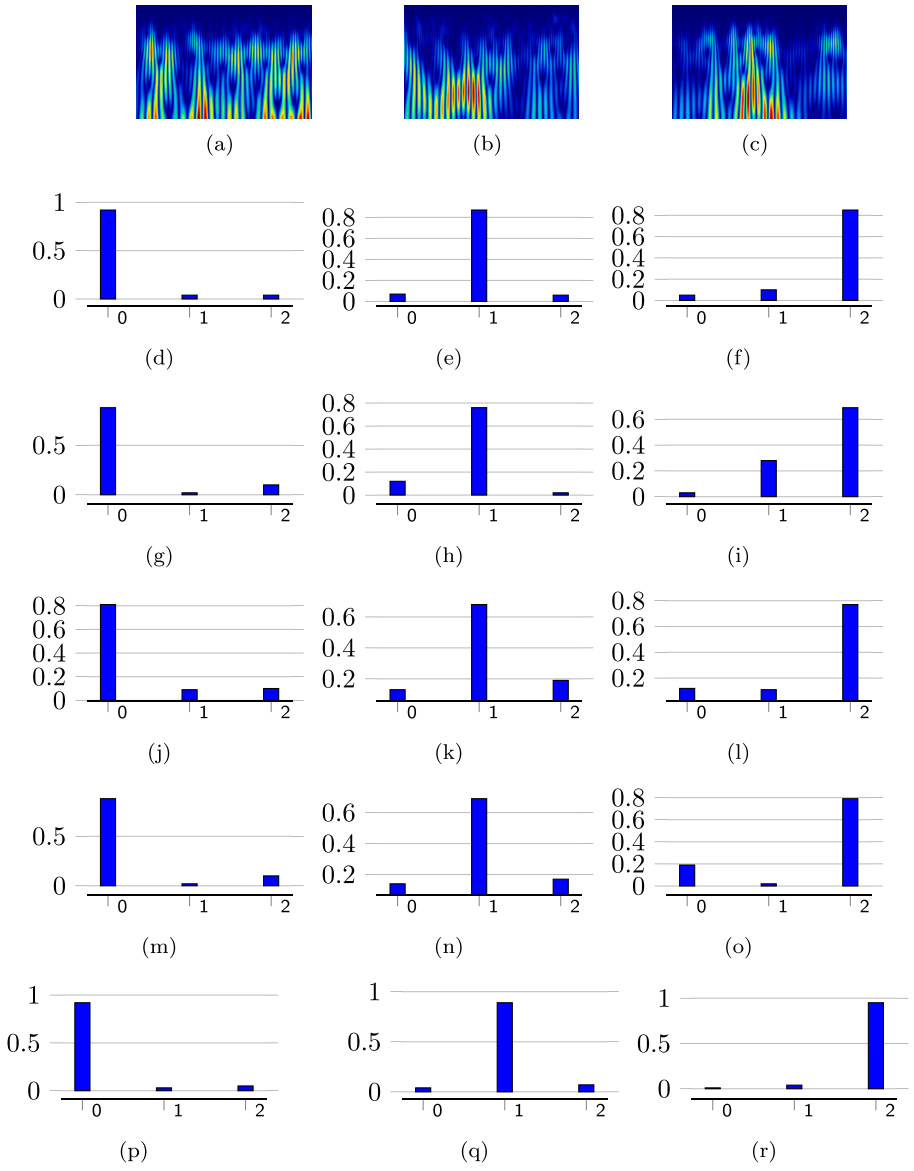


Fig. 16. The comparative probability scores of the MD-DCNN along with those of existing methods. (a), (b), and (c) represent the pre-ictal, inter-ictal, and ictal scalogram images. Each row represents the probability scores of three class scalogram images. (d), (e), and (f) are for [75], (g), (h), and (i) are for [76], (j), (k), and (l) are for [77], (m), (n), and (o) are for [78], (p), (q), and (r) are for the proposed method. Here, 0, 1, and 2 represent Pre-ictal, Inter-ictal, and Ictal.

Table 9
Significance test of the MD-DCNN.

Method vs. MD-DCNN	Database-1		Database-2		Database-3	
	Two-class	Three-class	Two-class	Three-class	Two-class	Three-class
Omer et al. [57]	0.001953	0.009765	0.001953	0.001953	0.003906	–
Cao et al. [74]	0.048828	0.019531	0.003906	0.019531	0.001953	–
Tian et al. [52]	0.037109	0.003906	0.001953	0.001953	0.009765	–
Qin et al.[58]	0.003906	0.001953	0.005859	0.048828	0.001953	–
Saidi et al. [42]	0.048828	0.160156	0.001953	0.048828	0.001953	–
Rashed et al. [59]	0.001953	0.005859	0.003906	0.003906	0.001953	–
Theekshana et al. [75]	0.001953	0.001953	0.001953	0.013671	0.020879	–
Wang et al. [76]	0.001953	0.001953	0.003906	0.027343	0.001953	–
Shoeibi et al. [77]	0.001953	0.001953	0.027343	0.003906	0.003906	–
Gao et al. [78]	0.048828	0.001953	0.009765	0.027343	0.010862	–
Xin et al. [60]	0.048828	0.005859	0.027343	0.0195313	0.027343	–
Lih et al. [61]	0.0371098	0.160156	0.001953	0.027343	0.010862	–
Albaqami et al. [62]	0.003906	0.001953	0.009765	0.019531	0.0273432	–

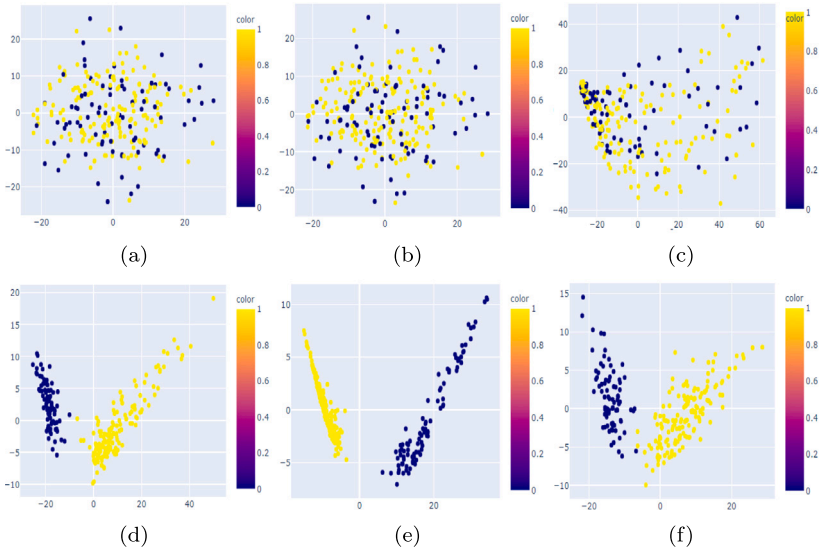


Fig. 17. Feature visualization of the proposed MD-DCNN. (a), (b), and (c) represent the original feature plots, and (d), (e), and (f) represent the features extracted by the MD-DCNN.

References

[1] A. Bhattacharyya, R.B. Pachori, A multivariate approach for patient-specific EEG seizure detection using empirical wavelet transform, IEEE Trans. Biomed. Eng. 64 (9) (2017) 2003–2015.

[2] Q. Zhu, H. Li, H. Ye, Z. Zhang, R. Wang, Z. Fan, D. Zhang, Incomplete multi-modal brain image fusion for epilepsy classification, Inf. Sci. 582 (2022) 316–333.

[3] S. Tong, N.V. Thankor, Quantitative EEG analysis methods and clinical applications, Artech House, 2009.

[4] L.S. Vidyaratne, K.M. Iftekharuddin, Real-time epileptic seizure detection using EEG, IEEE Trans. Neural Syst. Rehabil. Eng. 25 (11) (2017) 2146–2156.

[5] T.N. Alotaiby, S.A. Alshebeili, T. Alshawi, I. Ahmad, F.E. Abd El-Samie, EEG seizure detection and prediction algorithms: a survey, EURASIP J. Adv. Signal Process. 2014 (1) (2014) 1–21.

[6] Z. Deng, P. Xu, L. Xie, K.-S. Choi, S. Wang, Transductive joint-knowledge-transfer TSK FS for recognition of epileptic EEG signals, IEEE Trans. Neural Syst. Rehabil. Eng. 26 (8) (2018) 1481–1494.

[7] A.H. Shoeb, J.V. Guttag, Application of machine learning to epileptic seizure detection, in: ICML, 2010.

[8] H. Adeli, Z. Zhou, N. Dadmehr, Analysis of EEG records in an epileptic patient using wavelet transform, J. Neurosci. Methods 123 (1) (2003) 69–87.

[9] C.J. Stam, Nonlinear dynamical analysis of EEG and MEG: review of an emerging field, Clinical Neurophysiol. 116 (10) (2005) 2266–2301.

[10] N. Kannathal, M.L. Choo, U.R. Acharya, P. Sadasivan, Entropies for detection of epilepsy in EEG, Comput. Methods Programs Biomed. 80 (3) (2005) 187–194.

[11] Y. Cao, W.-w. Tung, J. Gao, V.A. Protopopescu, L.M. Hively, Detecting dynamical changes in time series using the permutation entropy, Phys. Rev. E 70 (4) (2004) 046217.

[12] O.A. Adegbola, I.A. Adeyemo, F.A. Semire, S.I. Popoola, A.A. Atayero, A principal component analysis-based feature dimensionality reduction scheme for content-based image retrieval system, Telkomnika 18 (4) (2020) 1892–1896.

[13] U.R. Acharya, S.V. Sree, A.P.C. Alvin, J.S. Suri, Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework, Expert Syst. Appl. 39 (10) (2012) 9072–9078.

[14] S. Ullah, Z. Halim, Imagined character recognition through EEG signals using deep convolutional neural network, Med. Biol. Eng. Comput. 59 (5) (2021) 1167–1183.

[15] A.U. Rahman, A. Tubaishat, F. Al-Obeidat, Z. Halim, M. Tahir, F. Qayum, Extended ICA and M-CSP with BiLSTM towards improved classification of EEG signals, Soft Comput. 26 (20) (2022) 10687–10698.

[16] Aadam, A. Tubaishat, F. Al-Obeidat, Z. Halim, M. Waqas, F. Qayum, EmoPercept: EEG-based emotion classification through perceiver, Soft Comput. (2022) 1–8.

[17] T. Hu, M. Khishe, M. Mohammadi, G.-R. Parvizi, S.H.T. Karim, T.A. Rashid, Real-time COVID-19 diagnosis from X-Ray images using deep CNN and extreme learning machines stabilized by chimp optimization algorithm, Biomed. Signal Process. Control 68 (2021) 102764.

[18] C. Cai, B. Gou, M. Khishe, M. Mohammadi, S. Rashidi, R. Moradpour, S. Mirjalili, Improved deep convolutional neural networks using chimp optimization algorithm for covid19 diagnosis from the X-ray images, Expert Syst. Appl. 213 (2023) 119206.

[19] M. Khishe, F. Caraffini, S. Kuhn, Evolving deep learning convolutional neural networks for early COVID-19 detection in chest X-ray images, Mathematics 9 (9) (2021) 1002.

[20] B. Xu, D. Martín, M. Khishe, R. Boostani, COVID-19 diagnosis using chest CT scans and deep convolutional neural networks evolved by IP-based sine-cosine algorithm, Med. Biol. Eng. Comput. 60 (10) (2022) 2931–2949.

[21] A. Saffari, M. Khishe, M. Mohammadi, A.H. Mohammed, S. Rashidi, DCNN-fuzzyWOA: artificial intelligence solution for automatic detection of covid-19 using X-ray images, Comput. Intell. Neurosci. 2022 (2022).

[22] M. Khishe, An automatic COVID-19 diagnosis from chest X-ray images using a deep trigonometric convolutional neural network, J. Imaging Sci. 71 (2) (2023) 128–141.

[23] X. Wang, C. Gong, M. Khishe, M. Mohammadi, T.A. Rashid, Pulmonary diffuse airspace opacities diagnosis from chest X-ray images using deep convolutional neural networks fine-tuned by whale optimizer, Wirel. Pers. Commun. 124 (2) (2022) 1355–1374.

[24] H. Liu, C. Wang, X. Jiang, M. Khishe, A few-shot learning approach for Covid-19 diagnosis using quasi-configured topological spaces, J. Artif. Intell. Soft Comput. Res. 14 (1) (2023) 77–95.

[25] F. Chen, C. Yang, M. Khishe, Diagnose Parkinson's disease and cleft lip and palate using deep convolutional neural networks evolved by IP-based chimp optimization algorithm, Biomed. Signal Process. Control 77 (2022) 103688.

[26] D. Yao, W. Chi, M. Khishe, Parkinson's disease and cleft lip and palate of pathological speech diagnosis using deep convolutional neural networks evolved by IPWOA, Appl. Acoust. 199 (2022) 109003.

[27] C.P. Panayiotopoulos, A clinical guide to epileptic syndromes and their treatment, Springer, 2010.

[28] T. Zhang, W. Chen, M. Li, AR based quadratic feature extraction in the VMD domain for the automated seizure detection of EEG using random forest classifier, Biomed. Signal Process. Control 31 (2017) 550–559.

[29] A. Sharmila, P. Geethanjali, DWT based detection of epileptic seizure from EEG signals using naive Bayes and k-NN classifiers, IEEE Access 4 (2016) 7716–7727.

[30] P. Swami, T.K. Gandhi, B.K. Panigrahi, M. Tripathi, S. Anand, A novel robust diagnostic model to detect seizures in electroencephalography, Expert Syst. Appl. 56 (2016) 116–130.

[31] K. Polat, S. Güneş, Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform, Appl. Math. Comput. 187 (2) (2007) 1017–1026.

[32] Z. Zhang, K.K. Parhi, Low-complexity seizure prediction from iEEG/sEEG using spectral power and ratios of spectral power, IEEE Trans. Biomed. Circuit. Syst. 10 (3) (2015) 693–706.

[33] A.T. Tzallas, M.G. Tsipouras, D.I. Fotiadis, Epileptic seizure detection in EEGs using time–frequency analysis, IEEE Trans. Inf. Technol. Biomed. 13 (5) (2009) 703–710.

[34] K. Samiee, P. Kovács, M. Gabbouj, Epileptic seizure detection in long-term EEG records using sparse rational decomposition and local gabor binary patterns feature extraction, Knowl.-Based Syst. 118 (2017) 228–240.

[35] M. Peker, B. Sen, D. Delen, A novel method for automated diagnosis of epilepsy using complex-valued classifiers, IEEE J. Biomed. Health Inf. 20 (1) (2015) 108–118.

[36] S. Raghu, N. Sriraam, Y. Temel, S.V. Rao, A.S. Hegde, P.L. Kubben, Performance evaluation of DWT based sigmoid entropy in time and frequency domains for automated detection of epileptic seizures using svm classifier, Comput. Biol. Med. 110 (2019) 127–143.
[37] H. Kalbkhani, M.G. Shayesteh, Stockwell transform for epileptic seizure detection from EEG signals, Biomed. Signal Process. Control 38 (2017) 108–118.
[38] T. Zhang, W. Chen, LMD based features for the automatic seizure detection of EEG signals using SVM, IEEE Trans. Neural Syst. Rehabil. Eng. 25 (8) (2016) 1100–1108.
[39] M. Sharma, A.A. Bhurane, U.R. Acharya, MMSFL-OWFB: A novel class of orthogonal wavelet filters for epileptic seizure detection, Knowl.-Based Syst. 160 (2018) 265–277.
[40] Y. Li, W.-G. Cui, H. Huang, Y.-Z. Guo, K. Li, T. Tan, Epileptic seizure detection in EEG signals using sparse multiscale radial basis function networks and the Fisher vector approach, Knowl.-Based Syst. 164 (2019) 96–106.
[41] A.R. Hassan, A. Subasi, Y. Zhang, Epilepsy seizure detection using complete ensemble empirical mode decomposition with adaptive noise, Knowl.-Based Syst. 191 (2020) 105333.
[42] A. Saidi, S.B. Othman, S.B. Saoud, A novel epileptic seizure detection system using scalp EEG signals based on hybrid CNN-svm classifier, in: 2021 IEEE Symposium on Industrial Electronics & Applications, ISIEA, IEEE, 2021, pp. 1–6.
[43] M. Karnati, A. Seal, J. Jaworek-Korjakowska, O. Krejcar, Facial expression recognition in-the-wild using blended feature attention network, IEEE Trans. Instrum. Meas. (2023).
[44] M. Karnati, A. Seal, A. Yazidi, O. Krejcar, LieNet: A deep convolution neural networks framework for detecting deception, IEEE Trans. Cogn. Dev. Syst. (2021).
[45] W. Ding, M. Abdel-Basset, H. Hawash, S. Abdel-Razek, C. Liu, Fed-ESD: Federated learning for efficient epileptic seizure detection in the fog-assisted internet of medical things, Inform. Sci. 630 (2023) 403–419.
[46] M. Karnati, G. Sahu, A. Gupta, A. Seal, O. Krejcar, A pyramidal spatial-based feature attention network for schizophrenia detection using electroencephalography signals, IEEE Trans. Cogn. Dev. Syst. (2023).
[47] G. Sahu, M. Karnati, A. Gupta, A. Seal, Scz-scan: An automated schizophrenia detection system from electroencephalogram signals, Biomed. Signal Process. Control 86 (2023) 105206.
[48] A. Seal, R. Bajpai, M. Karnati, J. Agnihotri, A. Yazidi, E. Herrera-Viedma, O. Krejcar, Benchmarks for machine learning in depression discrimination using electroencephalography signals, Appl. Intell. 53 (10) (2023) 12666–12683.
[49] P. Thodoroff, J. Pineau, A. Lim, Learning robust features using deep learning for automatic seizure detection, in: Machine Learning for Healthcare Conference, PMLR, 2016, pp. 178–190.
[50] U.R. Acharya, S.L. Oh, Y. Hagiwara, J.H. Tan, H. Adeli, D.P. Subha, Automated EEG-based screening of depression using deep convolutional neural network, Comput. Methods Programs Biomed. 161 (2018) 103–113.
[51] I. Ullah, M. Hussain, H. Aboalsamh, et al., An automated system for epilepsy detection using EEG brain signals based on deep learning approach, Expert Syst. Appl. 107 (2018) 61–71.
[52] X. Tian, Z. Deng, W. Ying, K.-S. Choi, D. Wu, B. Qin, J. Wang, H. Shen, S. Wang, Deep multi-view feature learning for EEG-based epileptic seizure detection, IEEE Trans. Neural Syst. Rehabil. Eng. 27 (10) (2019) 1962–1972.
[53] W. Liang, H. Pei, Q. Cai, Y. Wang, Scalp EEG epileptogenic zone recognition and localization based on long-term recurrent convolutional network, Neurocomputing 396 (2020) 569–576.
[54] K. Akyol, Stacking ensemble based deep neural networks modeling for effective epileptic seizure detection, Expert Syst. Appl. 148 (2020) 113239.
[55] H. Daoud, M.A. Bayoumi, Efficient epileptic seizure prediction based on deep learning, IEEE Trans. Biomed. Circuit. Syst. 13 (5) (2019) 804–813.
[56] A. Gogna, A. Majumdar, R. Ward, Semi-supervised stacked label consistent autoencoder for reconstruction and analysis of biomedical signals, IEEE Trans. Biomed. Eng. 64 (9) (2016) 2196–2205.
[57] Ö. Türk, M.S. Özerdem, Epilepsy detection by using scalogram based convolutional neural network from EEG signals, Brain Sci. 9 (5) (2019) 115.
[58] H. Qin, B. Deng, J. Wang, G. Yi, R. Wang, Z. Zhang, Deep multi-scale feature fusion convolutional neural network for automatic epilepsy detection using EEG signals, in: 2020 39th Chinese Control Conference, CCC, IEEE, 2020, pp. 7061–7066.
[59] M. Rashed-Al-Mahfuz, M.A. Moni, S. Uddin, S.A. Alyami, M.A. Summers, V. Eapen, A deep convolutional neural network method to detect seizures and characteristic frequencies using epileptic electroencephalogram (EEG) data, IEEE J. Translat. Eng. Health Med. 9 (2021) 1–12.
[60] Q. Xin, S. Hu, S. Liu, L. Zhao, Y.-D. Zhang, An attention-based wavelet convolution neural network for epilepsy eeg classification, IEEE Trans. Neural Syst. Rehabil. Eng. 30 (2022) 957–966.
[61] O.S. Lih, V. Jahmunah, E.E. Palmer, P.D. Barua, S. Dogan, T. Tuncer, S. García, F. Molinari, U.R. Acharya, EpilepsyNet: Novel automated detection of epilepsy using transformer model with EEG signals from 121 patient population, Comput. Biol. Med. 164 (2023) 107312.
[62] H. Albaqami, G.M. Hassan, A. Datta, MP-SeizNet: A multi-path CNN Bi-LSTM network for seizure-type classification using EEG, Biomed. Signal Process. Control 84 (2023) 104780.
[63] R.G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, C.E. Elger, Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state, Phys. Rev. E 64 (6) (2001) 061907.
[64] P. Swami, B. Panigrahi, S. Nara, M. Bhatia, T. Gandhi, EEG epilepsy datasets, 2016.
[65] P. Handa, M. Mathur, N. Goel, Open and free EEG datasets for epilepsy diagnosis, 2021, arXiv preprint arXiv:2108.01030.
[66] A. Seal, R. Bajpai, J. Agnihotri, A. Yazidi, E. Herrera-Viedma, O. Krejcar, DeprNet: A deep convolution neural network framework for detecting depression using EEG, IEEE Trans. Instrum. Meas. 70 (2021) 1–13.
[67] K. Wang, X. Peng, J. Yang, D. Meng, Y. Qiao, Region attention networks for pose and occlusion robust facial expression recognition, IEEE Trans. Image Process. 29 (2020) 4057–4069.
[68] C. Torrence, G.P. Compo, A practical guide to wavelet analysis, Bull. Am. Meteorol. Soc. 79 (1) (1998) 61–78.
[69] A. Narin, Detection of focal and non-focal epileptic seizure using continuous wavelet transform-based scalogram images and pre-trained deep neural networks, IRBM (2020).
[70] Z. Aslan, M. Akin, A deep learning approach in automated detection of schizophrenia using scalogram images of EEG signals, Phys. Eng. Sci. Med. (2021) 1–14.
[71] E. Bernitsas, The emerging role of scalogram-based convolutional neural network in the diagnosis of epileptic seizures, Brain Sci. 11 (11) (2021).
[72] X. Glorot, A. Bordes, Y. Bengio, Deep sparse rectifier neural networks, in: Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, 2011, pp. 315–323.
[73] A.L. Maas, A.Y. Hannun, A.Y. Ng, et al., Rectifier nonlinearities improve neural network acoustic models, in: Proc. Icml, Vol. 30, No. 1, Citeseer, 2013, p. 3.
[74] J. Cao, J. Zhu, W. Hu, A. Kummert, Epileptic signal classification with deep EEG features by stacked CNNs, IEEE Trans. Cogn. Dev. Syst. 12 (4) (2019) 709–722.
[75] T. Dissanayake, T. Fernando, S. Denman, S. Sridharan, C. Fookes, Deep learning for patient-independent epileptic seizure prediction using scalp EEG signals, IEEE Sens. J. 21 (7) (2021) 9377–9388.
[76] Z. Wang, J. Yang, M. Sawan, A novel multi-scale dilated 3D CNN for epileptic seizure prediction, in: 2021 IEEE 3rd International Conference on Artificial Intelligence Circuits and Systems, AICAS, IEEE, 2021, pp. 1–4.
[77] A. Shoeibi, N. Ghassemi, M. Khodatars, P. Moridian, R. Alizadehsani, A. Zare, A. Khosravi, A. Subasi, U.R. Acharya, J.M. Gorriz, Detection of epileptic seizures on EEG signals using ANFIS classifier, autoencoders and fuzzy entropies, Biomed. Signal Process. Control 73 (2022) 103417.
[78] Y. Gao, X. Chen, A. Liu, D. Liang, L. Wu, R. Qian, H. Xie, Y. Zhang, Pediatric seizure prediction in scalp EEG using a multi-scale neural network with dilated convolutions, IEEE J. Transl. Eng. Health Med. 10 (2022) 1–9.
[79] K.K. Sharma, A. Seal, Clustering analysis using an adaptive fused distance, Eng. Appl. Artif. Intell. 96 (2020) 103928.