Computer Methods and Programs in Biomedicine 262 (2025) 108652

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.sciencedirect.com/journal/computer-methods-and-programs-in-biomedicine

EEGConvNeXt: A novel convolutional neural network model for automated detection of Alzheimer's Disease and Frontotemporal Dementia using EEG signals

Madhav Acharya [a], Ravinesh C Deo [a,*], Prabal Datta Barua [a,b,c,d,e,f,g,h,i,j,k], Aruna Devi [l], Xiaohui Tao [m]

[a] School of Mathematics, Physics and Computing, University of Southern Queensland, Springfield Central, QLD 4300, Australia
[b] Cogninet AI, Sydney, NSW 2010, Australia
[c] Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, NSW 2007, Australia
[d] School of Computing and Information Science, Anglia Ruskin University Cambridge Campus, UK
[e] Australian International Institute of Higher Education, Sydney, NSW 2000, Australia
[f] School of Science Technology, University of New England, Australia
[g] School of Biosciences, Taylor's University, Malaysia
[h] School of Computing, SRM Institute of Science and Technology, India
[i] School of Science and Technology, Kumamoto University, Japan
[j] Sydney School of Education and Social Work, University of Sydney, Australia
[k] School of Mathematics, Physics and Computing, University of Southern Queensland, Toowoomba QLD 4350, Australia
[l] School of Education and Tertiary Access, University of the Sunshine Coast, Petrie, QLD, Australia
[m] School of Mathematics, Physics and Computing, University of Southern Queensland, Toowoomba, QLD 4350, Australia

ARTICLE INFO

Keywords:
EEGConvNeXt
AD detection
Transformer-like CNN
EEG analysis
Signal processing

ABSTRACT

Background and objective: Deep learning models have gained widespread adoption in healthcare for accurate diagnosis through the analysis of brain signals. Neurodegenerative disorders like Alzheimer's Disease (AD) and Frontotemporal Dementia (FD) are increasingly prevalent due to age-related brain volume reduction. Despite advances, existing models often lack comprehensive multi-class classification capabilities and are computationally expensive. This study addresses these gaps by proposing EEGConvNeXt, a novel convolutional neural network (CNN) model for detecting AD and FD using electroencephalogram (EEG) signals with high accuracy.
Materials and method: In this research, we employ an open-access EEG signal public dataset containing three distinct classes: AD, FD, and control subjects. We then constructed a newly proposed EEGConvNeXt model comprised of a 2-dimensional CNN algorithm that firstly converts the EEG signals into power spectrogram-based images. Secondly, these images were used as input for the proposed EEGConvNeXt model for automated classification of AD, FD, and a control outcome. The proposed EEGConvNeXt model is therefore a lightweight model that contributes to a new image classification CNN structure based on the transformer model with four primary stages: a stem, a main model, downsampling, and an output stem.
Results: The EEGConvNeXt model achieved a classification accuracy of ~95.70% for three-class detection (AD, FD, and control), validated using a hold-out strategy. Binary classification cases, such as AD versus FD and FD versus control, achieved accuracies exceeding 98%, demonstrating the model's robustness across scenarios.
Conclusions: The proposed EEGConvNeXt model demonstrates high classification performance with a lightweight architecture suitable for deployment in resource-constrained settings. While the study establishes a novel framework for AD and FD detection, limitations include reliance on a relatively small dataset and the need for further validation on diverse populations. Future research should focus on expanding datasets, optimizing architecture, and exploring additional neurological disorders to enhance the model's utility in clinical applications.

* Corresponding author.

https://doi.org/10.1016/j.cmpb.2025.108652
Received 3 September 2024; Received in revised form 23 January 2025; Accepted 5 February 2025
Available online 8 February 2025

1. Introduction

Alzheimer's disease (AD) and Frontotemporal Dementia (FD) are two of the most common neurodegenerative diseases [1,2]. These two types of dementia appear to be causing progressive loss of the brain cells [3]. AD, in particular, is the most common type of dementia worldwide [4]. Memory loss, thinking and behavioral problems are frequently observed in this disease [5]. It is generally associated with the death of brain cells and shrinkage of brain tissue. It is usually seen in individuals over 65 years old [6]. Although the exact cause is unknown, genetic and environmental factors are thought to be the contributors to AD [7,8]. FD is a less common type of dementia compared to AD. It usually begins between the ages of 45–65 [9]. It affects the frontal and temporal regions of the brain. Hence, it is called FD [1]. This disorder is characterized by personality changes, behavioral disorders, and language problems [10]. Similar to AD, genetic factors play an important role in this disorder. Various methods have been used for the preliminary diagnosis of both these diseases. Although the methods used for the diagnosis of both diseases show similarities, some special tests can be applied [11]. However, the present approaches used in the diagnosis of both diseases are generally based on clinical evaluation, neurological examination, cognitive tests, medical imaging approaches (MRI, CT, PET, etc.), genetic tests, and EEG [12], to name a few. These diagnostic methods are usually employed in combination, and the results of multiple methods are evaluated together to arrive at a conclusion.

Both neurodegenerative disorders are progressive in nature and, therefore, can reach a critical level that significantly affects an individual's daily life activities over time [13]. Therefore, early diagnosis and treatment are very important [14]. The development of appropriate care strategies is especially important in improving the quality of life of the individual. In recent years, deep learning and EEG-based analysis have been widely used in diagnosing brain diseases and neuromodulation applications. In particular, studies such as supervised network-based fuzzy learning methods for analyzing EEG signals in the diagnosis of Alzheimer's disease [15] and the effects of acupuncture stimulation on spectral power and functional connectivity in the human brain [16] demonstrate the potential of innovative approaches in this field. Advances in these studies provide an essential basis for the early diagnosis of neurological diseases and the evaluation of the effects of alternative treatment methods. EEG signals are an important data source that is widely used in the diagnosis of brain diseases and directly reflects time-frequency characteristics. The interpretability of proposed deep learning models is critical for understanding the effectiveness of detection algorithms. In this context, EEG-based spectral power analysis has attracted attention for its capacity to reflect brain dynamics. For example, the effectiveness of acupuncture in modulating brain activity has been evaluated using periodic and non-periodic EEG measurements [17]. Similarly, studies on the spatiotemporal dynamics of brain activity under peripheral nerve stimulation emphasize the role of spectral and connectivity features in EEG analysis [18]. The findings of these studies form the basis for understanding the interpretability of power spectrogram features and their potential for clinical use. In this context, comparing the power of the model proposed in our study with periodic and non-periodic EEG features will be an important step in evaluating the overall effectiveness of the algorithm.

Building on the insights gained from previous research into the interpretability of EEG-based features and their clinical applications, our study seeks to address these gaps by introducing EEGConvNeXt, an innovative model designed to advance the automated detection of neurodegenerative conditions such as AD and FD. To provide a novel contribution for automatically identifying AD and FD from the most commonly available EEG signals, in this research work, we aim to develop an artificial intelligence-based diagnosis model denoted as EEGConvNeXt to test its capability to detect AD, FD, and control groups, respectively. The need to develop the proposed EEGConvNeXt model arises from current trends in artificial intelligence (AI)-based technologies in healthcare, which is a frequently debated topic in health and medical informatics literature [19]. Table 1 summarizes the studies conducted on automated detection of AD using EEG signals. The main focus of any AI-based healthcare technology is to detect unhealthy patients from groups of healthy subjects accurately by using classical machine learning and deep learning methods [20]. In this research work, we contributed a new AI methodology to build the EEGConvNeXt model as an automated system that can accurately detect AD, FD, and control groups using an open-access dataset.

Prior to developing the EEGConvNeXt model, we first review a suite of available literature studies that have focused specifically on AD and FD detection using EEG datasets. In Table 1, we show that various automated approaches have been proposed for AD using EEG signals. In most of these cases, classical machine learning methods have been used, whereas a few such studies have used deep learning methods. Almost all of these studies, however, have used hold-out CV methods. The study of Kachare et al. [25] has used deep learning to classify EEG data and achieved a classification accuracy of 99.5% with its hold-out CV strategy. It is important to note that this table focuses on studies related to Alzheimer's disease detection. However, their study considered two types of dementia, Alzheimer's and FD, as well as a healthy control group. Several other researchers have used established deep learning models to classify the EEG signals, and most of them have reported satisfactory classification performances using well-known techniques [33,34].

A few researchers have applied novel models to analyze EEG signals, but these models have not been aimed at specifically classifying EEG signals for AD and FD detection [35]. It is, therefore, crucial to leverage the existing research on EEG classification while highlighting the specific cases that have either studied or not studied AD and FD to identify gaps and future research directions. Although EEG signal classification has been extensively researched, particularly in the general contexts or in the context of other neurological disorders, there remains a significant gap in such studies that have focused specifically on the nuanced differences between AD and FD using EEG data. Our work is necessary because previous research has largely overlooked the challenges associated with these two neurodegenerative conditions. This approach not only emphasizes the unique contributions of our research study but also identifies the pending areas within the existing body of knowledge, thereby positioning our research as a pivotal step in advancing the application of EEG-based classification for AD and FD. As a result, there is an opportunity to propose new-generation deep models for the neuroscience field using EEG signals [36].

The second problem in previous studies has been the number of classes used. The majority of those studies in the literature have used AD detection only and very few of those studies have attempted to classify the EEG signals using multiclass classification models [37,38].

The motivation for this research stems from the identified gaps in existing literature, particularly the need for more efficient and accurate models for EEG-based classification of neurodegenerative diseases. Our study therefore builds on the success of previous models, such as the ConvNeXt [39] and the transformer architecture [40], known for their high classification capabilities using deep learning techniques for healthcare applications. Considering this, the primary novelties of the present research is as follows:

- This study developed a new CNN model, EEGConvNeXt, for classifying AD and FD using EEG signals. It combines ideas from ConvNeXt and Transformer models to improve accuracy and performance.
- EEG signals are converted into power spectrogram images using Continuous Wavelet Transform (CWT). This process captures time and frequency information, making the model more accurate.
- The EEGConvNeXt model is lightweight, with only 7.5 million parameters. This makes it suitable for mobile and wearable devices in healthcare.

Table 1
Summary of recent studies on AD detection using machine learning methods and EEG signal datasets. (Note that Acc. = Accuracy, Sen. = Sensitivity, Spe. = Specificity, Pre. = Precision, F1scr. = F1 score)

Reference	Dataset	Methods	Classification	Validation (CV)	Result(s)
Rodrigues et al., 2021 [21]	EEG, 38 subjects (8 MCI, 11 Mild and Moderate AD, 8 Advanced AD and 11 HC), 256 Hz, 19 channel, 5-second segmented	DWT, Cepstral and Lacstral analysis, Distance measures, Statistical analysis	ANN	Leave-one-out	Acc.= 95.55 Sen.= 90.83 Spe.= 97.73
Puri et al., 2022 [22]	EEG, 23 subjects (12 AD and 11 HC), 256 Hz, 16 channel, 5-second segmented	IMFs; Hjorth parameters, EMD, PCA	SVM	10-fold	Acc.= 92.90 Sen.= 94.32 Spe.= 94.34 Pre.= 94.33 F1Scr.= 94.32
Araujo et al., 2022 [23]	EEG, 38 subjects (8 MCI, 11 Mild and Moderate AD, 8 Advanced AD and 11 HC), 256 Hz, 19 channel, 5-second segmented	WPD; Entropy- based feature extraction, Fractal analysis, Higuchi exponent, Kruskal-Wallis test for wavelet selection process	SVM	Leave-one-out	Acc.= 56.80
Pallathadka et al., 2024 [24]	EEG, 131 subjects (66 Dementia and 65 HC), 512 Hz, 19 channel	Katz fractal dimension and sample entropy, Statistical analysis with ANOVA	kNN	Hold-out (75:25)	Acc.= 98.05 Sen.= 97.03 Spe.= 99.16
Kachare et al., 2024 [25]	EEG, 23 subjects (12 AD and 11 NC), 256 Hz, 16 channel, 5-second segmented	Preprocessing and custom-designed CNN	Softmax	Hold-out (60:20:20)	Acc.= 98.50 Sen.= 100 Spe.= 97.55
Gopu et al., 2024 [26]	EEG, 86 subjects (23 HC and 63 Mild Cognitive Impairment (MCI) or AD, 256 Hz, 19 channel	HHT, PCA, EMD, IMFs	kNN	10-fold CV	Acc.= 98.60
Faghfouri et al., 2024 [27]	Two EEG datasets, 61 subjects for DB1 (29 MCI and 32 HC), 43 subjects for DB2 (22 MCI and 21 HC), 19 channel and 256 Hz for DB1, 14 channel and 128 Hz for DB2	PARAFAC; Spectral and entropy features, Statistical complexity	SVM	Hold-out CV (70:30)	For DB1 Acc.= 93.96 Sen.= 96.07 Spe.= 96.98 For DB2 Acc.=78.65
Nour et al., 2024 [28]	Two EEG datasets, 48 subjects for DB1 (24 AD and 24 HC), 92 subjects for DB2 (80 AD and 12 HC), 19 channel and 128 Hz for DB1, 19 channel and 128 Hz for DB2	5 frequency band segmented, Deep ensemble classification model	2D-CNN	5-fold CV	Acc.= 97.9 Rec.= 96.0 Pre.= 96.0 F1Scr.= 97.0
Joshi et al., 2025 [29]	EEG, 168 subjects (102 HC, 59 AD and 7 MCI), 19 channel	Power spectral density, Hjorth parameters, Multilayer perceptron, LSTM	Softmax	5-fold CV	Acc.= 97.27 Pre.= 96.43 Spe.= 96.36 Sen.= 98.18
Sharma and Meena, 2025 [30]	Two EEG datasets, 48 subjects for DB1 (24 AD and 24 HC), 92 subjects for DB2 (80 AD and 12 HC), 19 channel and 128 Hz for DB1, 19 channel and 128 Hz for DB2	Handling missing value, Artifacts removal, Graph fourier transform, Graph wavelet transform, Statistical moment,	Softmax	Hold-out CV (80:20)	For DB1 Acc.= 99.48 Pre.=99.2 Rec.= 99.0 F1Scr.= 99.4 For DB2 Acc.= 99.22 Pre.=99.0 Rec.= 99.0 F1Scr.= 99.1
Abadal et al., 2025 [31]	EEG, 189 subjects (63 HC, 63 AD and 63 MCI), 19 channel and 256 Hz	Graph Convolutional Networks, Graph Attention Networks (GAT), Self-supervised GAT, Graph Transformer Networks	GNN	Leave one out	Acc.= 91.77
Rezaee and Zhu, 2025 [32]	EEG, 189 subjects (110 HC, 59 AD and 56 MCI), 21 channel and 256 Hz	Short time fourier transform, Deep feature extraction, Handcrafted feature extraction,	optCascadeNet	5-fold CV	Acc.= 98.84 Pre.= 98.67 Rec.= 98.76 F1Scr.=98.71

*Acc: Accuracy, Sen: Sensitivity, Spe: Specificity, F1Scr: F1-Score, ANN: Artificial Neural Network, SVM: Support Vector Machine, kNN: k-Nearest Neighbor, CV: Cross Validation, MCI: Mild Cognitive Impairment; DWT: Discrete Wavelet Transform; IMFs: Intrinsic Mode Functions; EMD: Empirical Mode Decomposition; PCA: Principal Component Analysis; WPD: Wavelet Packet Decomposition; PARAFAC: Parallel Factor Analysis Tensor decomposition; HHT: Hilbert-Huang Transform; LSTM: Long short term memory

- The model uses Grad-CAM to create heatmaps that show which parts of the EEG signal are important for classification. This helps explain the model's decisions, making it more trustworthy for clinicians.
- The EEGConvNeXt model achieves high classification accuracy, reaching 95.70% for detecting AD, FD, and healthy controls. It is efficient and works well with minimal computational resources.
- The proposed EEGConvNeXt model is lightweight with only 7.5 million learnable parameters, making it practically suitable to deploy in future practical healthcare applications, including mobile and wearable electronic devices [41].

The remainder of this paper is structured as follows: Section 2 details the materials and methods used in this study, including the EEG dataset, the proposed EEGConvNeXt model architecture, and the procedures for data preprocessing and model training. Section 3 presents the results of the study, highlighting the model's classification accuracy and performance across various test cases. Section 4 discusses the implications of these results, compares them with previous studies, and explores the potential applications of the model in clinical settings. Finally, Section 5 concludes the paper by summarizing the key findings, acknowledging the limitations, and suggesting directions for future research.

2. Materials and method

2.1. EEG dataset

To construct the proposed EEGConvNeXt model, in this research work, we utilized an open-access dataset with EEG recordings. These recordings were taken from several individuals with AD, FD, and healthy subjects [42,43]. The dataset included EEG recordings collected from 88 subjects categorized as 36 AD patients, 23 FD patients, and 29 control individuals. To acquire the dataset, the EEG signals were recorded at a sampling frequency of 500 Hz by means of a 19-channel Nihon Kohden EEG device. The placement of electrodes followed the internationallyaccepted 10–20 system, with the channels positioned at Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2. The electrodes A1 and A2 were designated as the reference electrodes. The EEG signals were segmented into 5-second epochs with these segments later transformed into spectrogram images using the CWT algorithm. The transformed images were then fed as potential input to the proposed EEGConvNeXt model for the classification of data showing AD, FD, and control subjects. Table 2 shows the distribution of the datasets used to construct the proposed EEGConvNeXt model.

2.2. Development of the EEGConvNeXt model

In this study, a novel EEGConvNeXt model is proposed with considerations to achieve a high classification performance with a model architecture that has fewer learnable parameters. The proposed EEGConvNeXt model is described as lightweight due to its relatively small number of learnable parameters, totaling only 7.5 million. This number is significantly lower compared to other state-of-the-art deep learning models, which often exceed 10 million parameters. Such a reduction in parameters enhances computational efficiency, making the model more suitable for deployment in resource-constrained environments such as mobile or wearable devices [41]. With respect to its design, this model is primarily inspired by the ConvNeXt [39] and the transformers [40] model design structures.

Table 2 also shows that the training and the testing data split ratios that were allocated into approximately 80%:20% ratios.

The proposed EEGConvNeXt model draws inspiration from ConvNeXt and transformer approaches but is specifically designed for EEG signal classification. Unlike ConvNeXt, which focuses on general image tasks, EEGConvNeXt processes EEG signals by converting them into spectrogram images, making it suitable for detecting AD and FD. Compared to transformers, which rely on attention mechanisms to analyze global relationships, EEGConvNeXt emphasizes local feature extraction using convolutional layers. This makes it less complex and more efficient, especially for smaller datasets. The model combines the strengths of both approaches, using a lightweight architecture and unique preprocessing steps like CWT to optimize performance for clinical applications. By tailoring these features, EEGConvNeXt achieves high accuracy while remaining practical for deployment in real-world healthcare settings. Moreover, we have used a convolution-based residual block in this work, and the number of filters is different from ConvNeXt.

Using these ratios, the CNN model was then trained by using the training dataset whereas the testing results were computed using the trained CNN and the testing images. Fig. 1 shows the graphical representation of the proposed EEGConvNeXt model.

Table 2
Details of the dataset used to develop the proposed EEGConvNeXt model.

No	Class	Train	Test	Total
1	AD	4,692	1,172	5,864
2	FD	2,672	667	3,339
3	Control	3,899	972	4,871
Total		11,263	2,811	14,074

In accordance with the schematic in Fig. 1, the proposed EEGConvNeXt model has four distinct steps, which are explained as follows.

Step 1: Reading each of the EEG signals.
Step 2: Converting each EEG signal to an image using the CWT method to generate a visual of the patterns embedded within the signal. The CWT was selected for its ability to provide both temporal and frequency-domain representations of non-stationary EEG signals. Unlike the Fourier Transform, CWT captures signal features across different scales and time intervals, making it ideal for complex and transient EEG patterns. By converting EEG signals into spectrogram images, CWT preserves critical temporal and spectral information, enabling the EEGConvNeXt model to leverage these features for accurate classification. The selected features, such as transient temporal events and disrupted frequency bands (e.g., theta, alpha), are aligned with prior studies on EEG-based neurological disorder detection, ensuring methodological soundness, robustness, and interpretability.
Step 3: Resize each of the images to 224 × 224 × 3 pixels, which makes it convenient to use in the proposed EEGConvNeXt model.
Step 4: Feed the generated images to the proposed EEGConvNeXt model for classifying AD, FD, and healthy/control observations.

In what follows next, we describe the various phases of the proposed EEGConvNeXt model. These crucial phases of the EEGConvNeXt model include the ‘stem’, ‘main’, ‘downsampling’ and the ‘output’ phase, described as follows:

Stem phase: The initial phase of the EEGConvNeXt model uses a patchify block that aims to transform the 224 × 224 × 3 sized images into a 56 × 56 × 96 sized tensor. The mathematical definition of this phase is given below.

$$Sout = G\left(BN\left(C^{96}_{4\times 4,\ Stride=4}(Im)\right)\right) \tag{1}$$

where *Sout*: Out of the stem block with a size of 56 × 56 × 96,*Im*: image with a size of 224 × 224 × 3, $G(.)$: GELU function, $BN(.)$: batch normalization and $C(.)$: convolution. The attributes of the convolution are given as follows. Filter size: 4 × 4, number of filters: 96, stride: 4. Therefore, this convolution is a patchify block.

Main phase: This is the microstructure of the proposed EEGConvNeXt model. This part is responsible for creating the primary feature map to detect AD, FD, and healthy controls. As a major contribution of this study, we have proposed a new block inspired by the original ConvNeXt algorithm and the transformer block. In this phase, we have used an inverted bottleneck with *three* convolutions, with filter sizes 3 × 3, 1 × 1, and 1 × 1. The mathematical definition of this phase is given below.

$$x_t = G\left(C^F_{3\times 3}(BN(x_{t-1}))\right) + x_{t-1} \tag{2}$$

$$x_{t+1} = C^F_{1\times 1}\left(G\left(C^{4F}_{1\times 1}(BN(x_t))\right)\right) + x_t \tag{L2}$$

where, x_{t-1}: the input tensor of the main stage, x_{t+1}: the output tensor of the main stage, and*F*: number of filters.

As part of the main phase, these operations are designed to capture complex features and patterns in the data using depth-based convolutions and residual connections. Following this, the channel-based relationships and spatial features are extracted using 3 × 3 and 1 × 1 convolutions. 3 × 3 convolutions to help extract local spatial features and provide more distinct features that can be extracted using an inverted bottleneck with 1 × 1 convolutions.

To overcome the vanishing gradient problem and facilitating the training of deeper networks, residual connections are integrated into the main stage. It is through these connections that the model learns identity mappings, ensuring that important features are preserved and passed through the layers. In combination with these techniques, the proposed

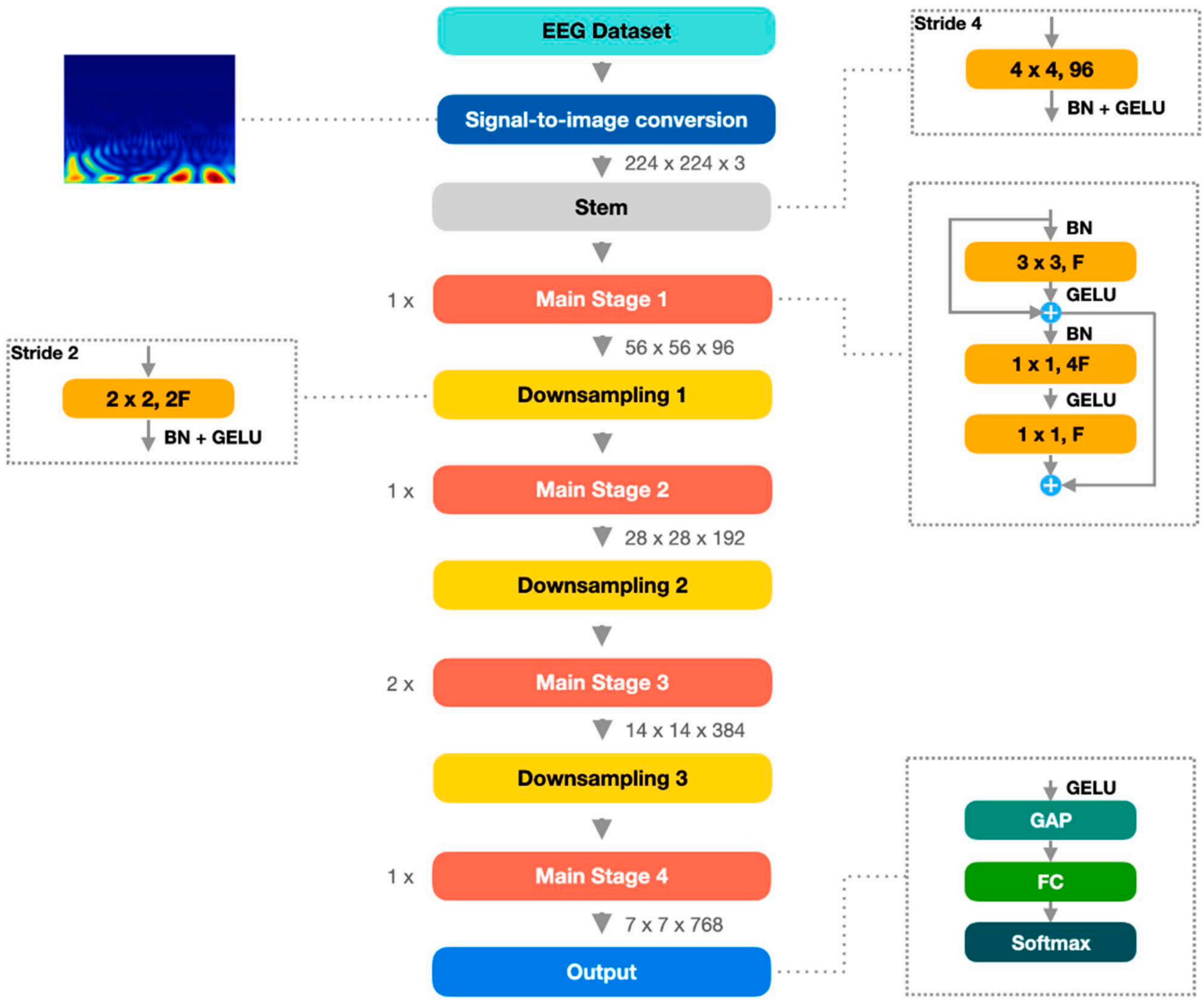

Fig. 1. A graphical representation of the proposed EEGConvNeXt model used to detect AD, FD, and healthy subjects.=

EEGConvNeXt model is then able to efficiently learn complex structures in EEG signals with fewer parameters.

Downsampling: To decrease the width and the height of the tensor and to increase the depth of the tensor, we have used a patchify-based downsampling block. In this phase, the pooling operator is used. However, pooling operators have routing problems. Therefore, new generation models have used convolution instead of pooling. In this work, we have used the downsampling block of ConvNeXt [39], and this patchify-based downsampling block is mathematically defined below.

$$DT = G\left(BN\left(C^{2F}_{2\times 2,Stride:2}(IT)\right)\right) \quad (3)$$

where, DT: the output tensor and IT: input tensor.

It is important to note that the patchify-based downsampling block used in Eq. (3) reduces the spatial dimensions of the input tensor while increasing its depth. It uses convolutional operations combined with batch normalization and GELU activation to enhance the learning capacity of the network.

Output: We have used GELU activation, global average pooling, fully connected layers, and Softmax to obtain classification results. The mathematical definition of this phase is:

$$Out = Softmax(FC(GAP(G(In)))) \quad (4)$$

where Out: the output of the proposed model, FC(.): fully connected layer, GAP(.): global average pooling and In: input tensor.

The complete structure of the proposed EEGConvNeXt model, along

Table 3
The transition table of the proposed EEGConvNeXt model demonstrating the layers, input, operation, and output matrix.

Layer	Input	Operation	Output
Stem	224 × 224 × 3	4 × 4, 96, BN + GELU, stride: 4	56 × 56 × 96
Main stage 1	56 × 56 × 96	$\begin{bmatrix} 3\times 3,\ 96 \\ 1\times 1,\ 384 \\ 1\times 1,\ 96 \end{bmatrix}$	56 × 56 × 96
DS 1	56 × 56 × 96	2 × 2, 192, BN + GELU, stride: 2	28 × 28 × 192
Main stage 2	28 × 28 × 192	$\begin{bmatrix} 3\times 3,\ 192 \\ 1\times 1,\ 768 \\ 1\times 1,\ 192 \end{bmatrix}$	28 × 28 × 192
DS 2	28 × 28 × 192	2 × 2, 384, BN + GELU, stride: 2	14 × 14 × 384
Main stage 3	14 × 14 × 384	$\begin{bmatrix} 3\times 3,\ 384 \\ 1\times 1,\ 1536 \\ 1\times 1,\ 384 \end{bmatrix} \times 2$	14 × 14 × 384
DS 3	14 × 14 × 384	2 × 2, 768, BN + GELU, stride: 2	7 × 7 × 768
Main stage 4	7 × 7 × 768	$\begin{bmatrix} 3\times 3,\ 768 \\ 1\times 1,\ 3072 \\ 1\times 1,\ 768 \end{bmatrix}$	7 × 7 × 768
Output	7 × 7 × 768	GELU, GAP, fully connected layer, Softmax	Number of classes
Total learnable parameters			7.5 million

with its transition table, is presented in Table 3.

In Table 3, we provide a detailed overview of the proposed EEG-ConvNeXt model architecture. Here, we show details how the input tensor's dimensions are transformed at each stage as well as the specific operations performed, and the role of each layer in the network. The repetitions ($\times$ 2) in Main Stage 3 indicate that the operations within that block are repeated twice, enhancing the model's capacity to capture complex patterns in the data. Moreover, this model has 7.5 million learnable parameters (a much smaller value than 10 million) demonstrating that the proposed EEGConvNeXt is a lightweight model suitable for practical healthcare applications in battery-powered gadgets like smartphones, wearables, and IoT devices.

According to Table 3, the general mathematical definition of the proposed EEGConvNeXt is given below.

$$F = \{96, 192, 384, 768\}, \ R = \{1, 1, 2, 1\} \tag{5}$$

where, F is the number of the filters and R signifies the number of repetitions.

3. Results and discussions

This section provides the experimental results of the presented research. It includes the experimental setup, classification performance metrics, classification results for the four defined cases, comparative analysis, ablation studies and explainable results analysis.

3.1. Experimental setup

In this research, we have presented a new-generation deep learning model called EEGConvNeXt. To implement this model, we used a simple-configured computer equipped with an NVIDIA GeForce 4090 graphical processing unit (GPU). By utilizing this GPU, the model was successfully implemented. To design the CNN structure, a total of 74 operators (or blocks) and 85 connections were used. To create a dataset, each EEG segment was transformed into power spectrogram images using a CWT transformation process. The results with more than 14,000 images were then divided into two sets (i.e., the training set and the testing set).

To create the EEGConvNeXt model, we utilized MATLAB (version 2023b) Deep Network Designer, a scratch-based tool. Layers (operators) such as convolutions, pooling, normalizations, and activations were selected using this tool, and these components were connected to form the deep learning model.

In the training phase, we used the following parameters:
Solver: Stochastic Gradient Descent Momentum,
Number of epochs: 30,
Initial learning rate: 0.01,
Mini-batch size: 128,
L2 Regularization: 0.0001,
Training and validation split ratio: 80:20 (randomized),
Augmentation: None.

To evaluate the classification capability of the proposed EEGConvNeXt model, we performed the classification of EEG signals by dividing them into various groups (or cases) as follows:

Case 1: This case included three classes: (1) AD, (2) FD, and (3) Control.
Case 2: This case included two classes: (1) AD and (2) FD.
Case 3: This case included two classes: (1) AD and (2) Control. The objective was to detect AD classes most accurately.
Case 4: This case included two classes: (1) FD and (2) Control. The objective was to detect FD classes accurately.

By utilizing these settings and the defined cases, we obtained the classification results of the proposed EEGConvNeXt model for the specified cases.

3.2. Model performance evaluation criteria

The performance of the proposed EEGConvNeXt model was evaluated using visual and statistical metrics as key indicators of the capability to detect AD, FD, and healthy individuals from EEG signals. In doing so, we have used accuracy, sensitivity, specificity, recall, precision and F1 score [44] which aims to measure the overall capability to detect the three classes from EEG signals. The mathematical formulation of these metrics are as follows:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

$$Spe = \frac{TN}{TN + FP} \tag{7}$$

$$Pre = \frac{TP}{TP + FP} \tag{8}$$

$$Rec = \frac{TP}{TP + FN} \tag{9}$$

In principle, these formulas are used to evaluate the performance of classification models. *Accuracy (Acc)* represents the proportion of correct predictions out of the total predictions. *Specificity (Spe)* measures how well the model identifies true negatives. *Precision (Pre)* indicates the proportion of true positives among the predicted positives. *Recall (Rec)* reflects how well the model identifies actual positive cases.

It should be noted that the classification accuracy measures the overall correctness of the proposed EEGConvNeXt model by indicating the percentage of correctly classified instances. Recall, also known as sensitivity, assesses the EEGConvNeXt model's ability to identify true positives, reflecting its effectiveness in detecting the actual cases. Precision, on the other hand, evaluates the proportion of true positive predictions among all positive predictions made, while the F1-score balances precision and recall, providing a single metric that considers both false positives and false negatives detected by the proposed EEG-ConvNeXt model.

3.3. Classification results

To demonstrate the proposed model's classification capability using these four cases, we show in Fig. 2 the training and validation curves obtained for these specific cases.

Fig. 2 shows that the proposed EEGConvNeXt model has reached nearly 100% training accuracy and over 99% validation accuracy for all four cases by approximately the 25th epoch. This indicates that the proposed model converges effectively within this range. This level of performance across multiple cases demonstrates the proposed model's robust learning and generalization capacity.

In Fig. 3, the confusion matrices for the tested results of the various cases are presented. These matrices illustrate how well the proposed model distinguished AD, FD, and control subjects. For instance, the confusion matrix for Case 1 (i.e., three-class classification) shows minimal misclassifications between the classes, highlighting the model's precision in handling complex, multi-class scenarios. Similarly, in Cases 2 and 3 (i.e., binary classifications), the proposed model exhibited a strong ability to differentiate between AD and FD, as well as between AD and control subjects, with very few errors. Case 4, which involves distinguishing between FD and control subjects, also showed high accuracy, reinforcing the model's reliability in binary classification tasks.

These results collectively demonstrate the proposed EEGConvNeXt model's capability to classify EEG signals accurately and efficiently across various scenarios.

To demonstrate the capability of the proposed EEGConvNeXt model, Table 4 lists the performance matrices, namely the classification

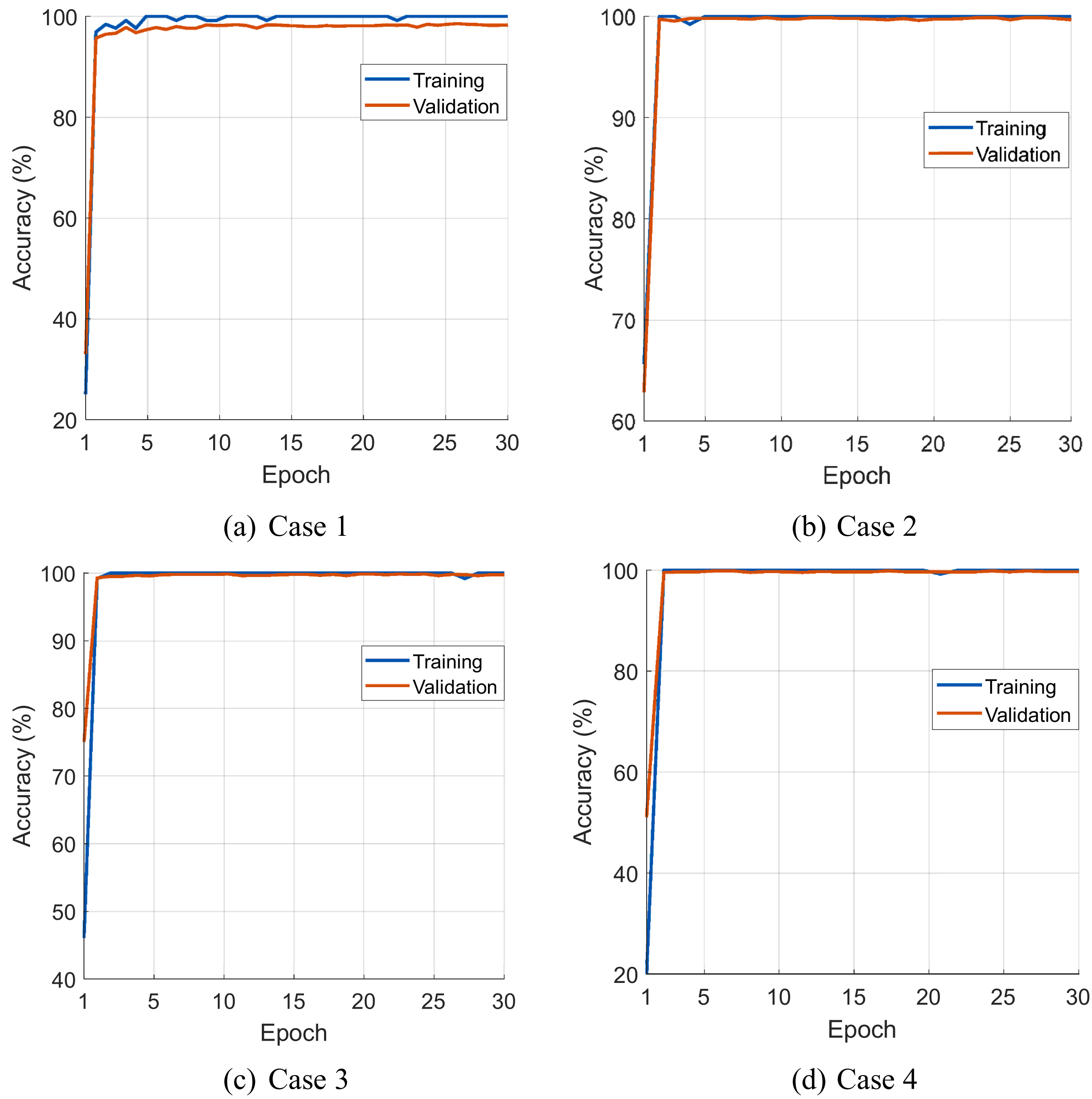

Fig. 2. The training and validation curves obtained for the four different cases in detecting AD, FD, and control individuals from EEG signals.

accuracy, recall, precision, and F1-score computed from the independently tested results for each of the four classification cases to detect AD, FD, and control individuals from EEG signals. Following Table 4, the proposed model has obtained more than 95% classification performance for all four test cases. This is exemplified by the maximum accuracy of 98.67% that was obtained for Case 4 (i.e., the FD detection model), whereas for Cases 1, 2, and 3, the EEGConvNeXt model produced an accuracy between 95.7–96.32%. In agreement with the accuracy metric, the Recall, Precision, and F1-score were also superior by about [96.08-98.17], [95.66-97.96], and [95.87-96.32], respectively. These impressive performance metrics suggest that the EEGConvNeXt model could be highly effective in real-world clinical settings for early detection and accurate diagnosis of AD and FD especially noting the high accuracy, recall, precision, and F1-scores showing robust predictive capabilities, potentially leading to better patient outcomes through timely and precise interventions. Moreover, the proposed model's reliability could support healthcare professionals in making more informed decisions, ultimately enhancing the overall quality of care for neurological conditions. As an additional note, we can ascertain that our proposed model uses about 7.5 million learnable parameters, making it a lightweight model.

3.4. Comparative results

In this part of the results section, we demonstrate the efficacy of the proposed lightweight CNN-based EEGConvNeXt model to unveil the hidden attributes of the EEG signals used to detect AD, FD, and control subjects. To pursue this, we first converted the EEG signals into 2D images using CWT transformation so that the features within these images are captured in a way that the proposed EEGConvNeXt can classify with high accuracy. Therefore, we defined four prior studies that were most relevant to justify the proposed EEGConvNeXt model's superiority. Based on this comparison, our proposed EEGConvNeXt model has attained 95.70% testing classification accuracy for this three-class classification problem and more than 96% classification accuracy for these two-class cases. Table 5 compares our proposed EEGConvNeXt

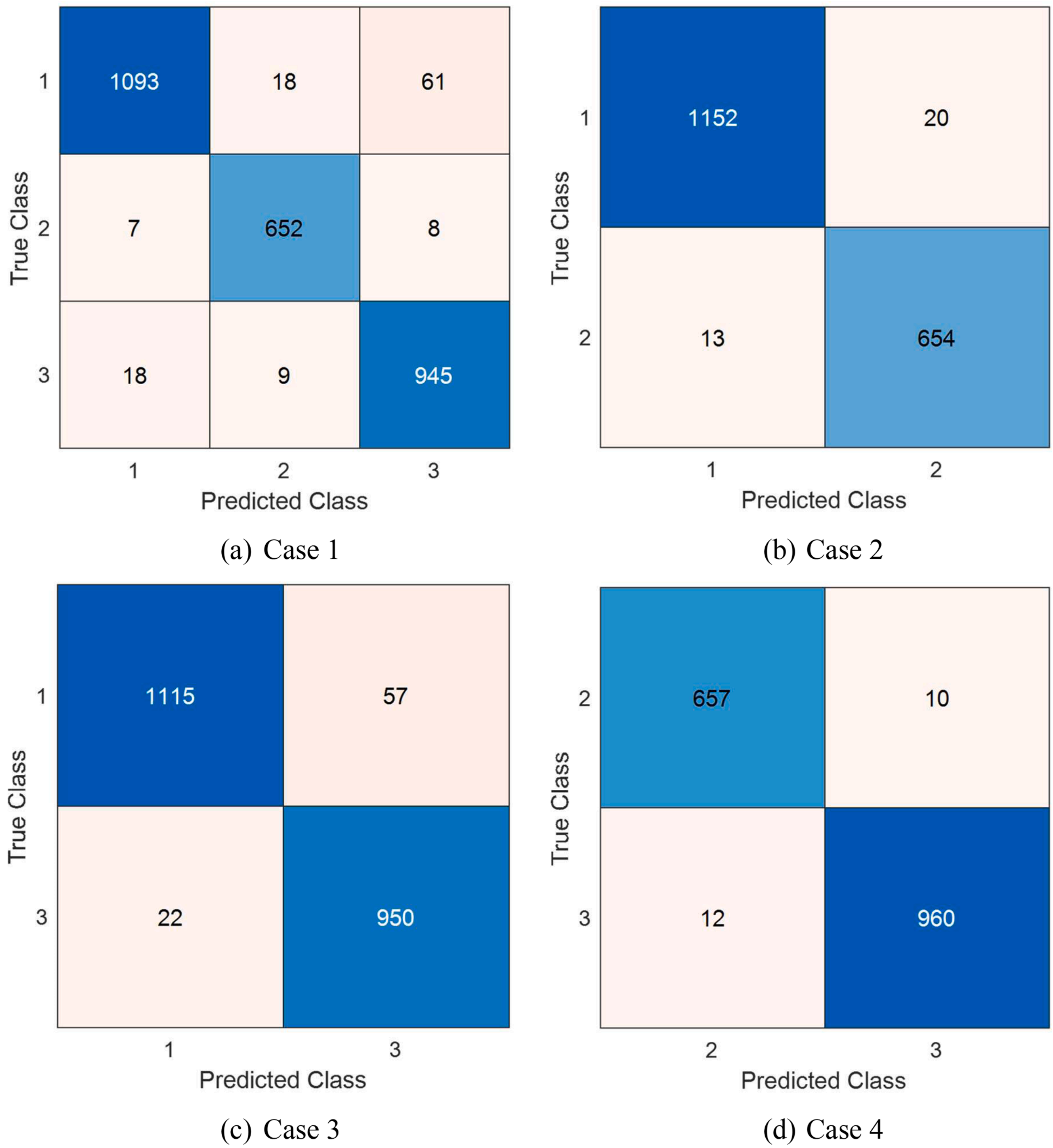

Fig. 3. The test phase confusion matrices obtained for the four different cases in detecting AD, FD, and control individuals from EEG signals using the proposed EEGConvNeXt model.
** 1: AD; 2: FD and 3: Control.

Table 4
Summary of performance metrics (%) obtained using the proposed EEGConvNeXt model for various cases.

Performance evaluation metric	Case 1	Case 2	Case 3	Case 4
Accuracy	95.70	98.21	96.32	98.67
Recall	96.08	98.17	96.44	98.64
Precision	95.66	97.96	96.20	98.60
F1-score	95.87	98.07	96.32	98.62

model with several state-of-the-art methods developed for the automated detection of AD and FD using the same dataset.

It is imperative to note that prior studies have generated either a two-class or a single-class classification outcome. In contrast, our study has produced a three-class outcome, with the four distinct performance metrics being much superior to those in existing literature. Unlike the other models, which primarily focused on binary classifications and yielded moderate accuracies and F1-scores, the proposed EEGConvNeXt model demonstrates its robustness by achieving significantly higher accuracy, recall, precision, and F1-scores across both binary and multi-class classification tasks.

The comparisons in Table 5 clearly illustrate the superior

Table 5
Comparison of the proposed EEGConvNeXt model with four other research works developed for automated detection of AD and FD using EEG signals with the same dataset.

Author(s) and Year	Classification	Method	Validation (CV)	Result(s)
Our Method	AD vs. FD vs. HC AD vs. FD AD vs. HC FD vs. HC	EEGConvNeXt	Hold-out (80:20)	For AD vs. FD vs. HC Acc.= 95.70 Rec.= 96.08 Pre. = 95.66 F1Scr.= 95.87 For AD vs. FD Acc.= 98.21 Rec.= 98.17 Pre. = 97.96 F1Scr.= 98.07 For AD vs. HC Acc.= 96.32 Rec.= 96.44 Pre.= 96.20 F1Scr.= 96.32 For FD vs. HC Acc.= 98.67 Rec.= 98.64 Pre.= 98.60 F1Scr.= 98.62 For AD vs. FD vs. HC with 10-fold CV Acc.= 98.43 Rec.= 98.13 Pre. = 98.79 F1Scr.= 98.46
Miltiadous et al., 2023 [42]	AD vs. HC FD vs. HC	Relative Band Power (RBP), Spectral Coherence Connectivity (SCC), Dual-Input Convolutional Encoder Network (DICE-net)	LOSO	For AD vs. HC Acc.= 83.28 Sen.=78.81 Spe.= 87.94 Pre.= 88.94 F1Scr.= 84.12 For FD vs. HC Acc.=74.96 Sen.= 60.62 Spe.= 78.63 Pre.= 64.01 F1Scr.= 62.27
Puri et al., 2024 [45]	AD vs. HC	Noise removal, Custom-designed CNN (LEADNet)	Hold-out (70:10:20)	Acc.=99.24 Sen.=100 Spe.=98.18 F1Scr.=99.35
Chen et al., 2023 [46]	AD vs. FD vs. HC	Custom-designed CNN for frequency channels, Vision transformer for time channels	10-fold	Acc.= 79.12 Sen.= 77.09 Spe.= 78.30 Auc.= 80.23
Kachare et al., 2024 [47]	AD vs. FD vs. HC	Custom-designed CNN (STEADYNet)	Hold-out (60:20:20)	Acc.= 84.59 Sen.= 84.58 Spe.= 82.56 F1Scr.= 84.63 Auc.= 88.02
Rostamikia et al., 2024 [48]	AD vs. FD (AD+FD) vs. HC	Time and frequency domain feature extraction, Kolmogorov-Smirnov test-based statistical analysis, SVM	Hold-out (70:30)	For AD vs. HC Acc.= 87.8 Sen.= 85.1 Spe.= 90.0 For AD±FD vs. HC Acc.= 93.5 Sen.= 90.0 Spe.= 93.0

**LOSO: Leave on subject out; CV: Cross-validation; Acc.: Accuracy; Rec.: Recall; Pre.: Precision; F1Scr.: F1-score.

performance of the proposed EEGConvNeXt model over several other existing methods. While prior studies focused on two-class or single-class classifications, the EEGConvNeXt model not only handles the more complex three-class and four-class classification tasks but also achieves significantly higher accuracy, recall, precision, and F1-scores across all cases. This demonstrates the effectiveness of the EEGConvNeXt model in accurately distinguishing between AD, FD, and control subjects, highlighting its potential for practical clinical application in the diagnosis of neurodegenerative diseases. Moreover, a comparison with well-known classical deep learning architectures in the literature is also made to prove the performance of the developed model. The results of this comparison process are given in Fig. 4.

As shown in Fig. 4, EEGConvNeXt has achieved significantly higher classification performance than well-known deep learning methods in the literature. In Fig. 4, a 10-fold CV was applied as a validation strategy and a three-class comparison (AD vs. FD vs. HC) was performed.

3.5. Ablation studies

To further ascertain the predictive capability of the proposed EEGConvNeXt model, we have conducted an ablation study to systematically analyze the model's performance. In this context, ablation refers to selectively removing or modifying specific components of the EEGConvNeXt model to better understand their individual contributions to the overall system. This approach helps identify which elements are most critical to the model's predictive accuracy. In doing so, we have presented three additional CNN models, as variants of the final model, with the fourth CNN being our proposed EEGConvNeXt model. To compare these CNN models, their validation accuracies have been used. The details of these CNN models are given below.

CNN Model 1: In this model, we have removed the repeated main stage 3 of the proposed EEGConvNeXt (see Fig. 1) and repeated all the main stages only once.
CNN Model 2: In this CNN model, only the bottleneck block is utilized as the main block. We have, therefore, removed the 3 × 3 sized convolution block in the main block of the CNN algorithm.
CNN Model 3: We have created a main stage by utilizing 3 × 3 sized convolutions.
CNN Model 4: This is our proposed final EEGConvNeXt model.

The validation accuracies obtained for these variants of CNN models using the present dataset are shown in Fig. 5.

As shown in Fig. 5, the proposed EEGConvNeXt model has achieved

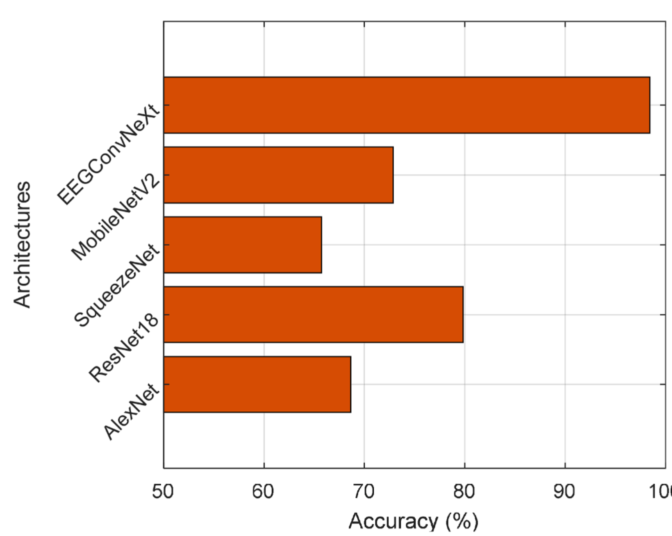

Fig. 4. Comparison of deep learning architectures

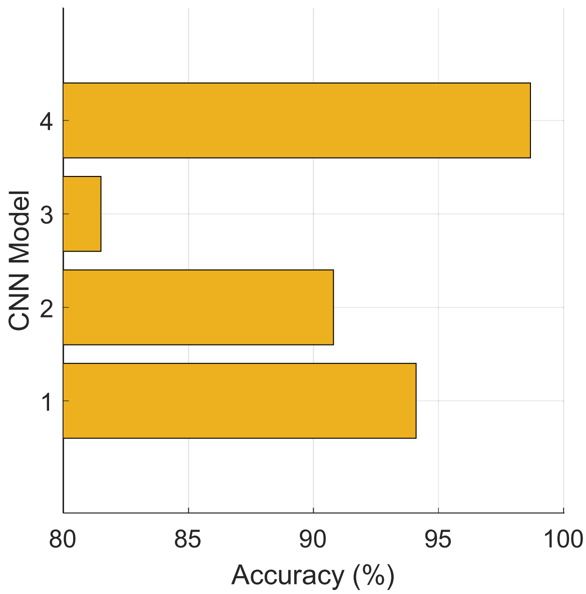

Fig. 5. A demonstration of the validation accuracies obtained for various CNN modes used in ablation studies to study the behavior of the final EEGConvNeXt model.

the highest validation accuracy compared to the ablation variants of the model (i.e., CNN Models 1, 2, and 3). This result also underscores the importance of the specific architectural choices made in constructing the proposed EEGConvNeXt model. It is imperative to note that the inferior performance of CNN Models 1, 2, and 3 can be attributed to the removal or modification of key components that are essential for capturing complex patterns in the EEG data. For instance, removing the repeated main stage in CNN Model 1 likely reduced the model's ability to learn hierarchical features. At the same time, the absence of the 3 × 3 convolution block in CNN Model 2 may have limited the model's capacity to extract spatial features. CNN Model 3, which relied solely on 3 × 3 convolutions, likely lacked the flexibility provided by the hybrid architecture of the final model. These findings align with existing literature, emphasizing preserving key architectural components to maintain model performance (see [Reference]). The ablation study not only highlights the proposed model's superiority but also provides insights into why specific architectural choices contribute to its effectiveness in detecting AD, FD, and control subjects from EEG datasets.

3.6. Explainable results

Given the potential for the proposed EEGConvNeXt model to be used in practical healthcare settings for identifying AD, FD, and control subjects, it is crucial to ensure that the model's decision-making process is transparent and understandable to clinicians. To achieve this, we have integrated an explainable artificial intelligence (XAI) approach into our study. Adopting XAI is essential because, in clinical environments, the interpretability of AI models is as important as their accuracy. Clinicians need to understand how a model arrives at its decisions, especially when these decisions can significantly impact patient care.

To provide this level of transparency, we employed the Grad-CAM technique. In principle, the Grad-CAM algorithm generates heatmaps highlighting the key areas of input images (in this case, EEG spectrograms) that the model has focused on when making its predictions. This approach allows us to visualize which regions of the EEG data were most influential in classifying a subject as having AD, FD, or being a healthy control. By providing these visual explanations, the Grad-CAM algorithms have not only increased the trustworthiness of the proposed EEGConvNeXt model but also is likely to aid clinicians in validating the model's outputs against their medical expertise.

Using the Grad-CAM algorithm in this context is particularly meaningful with respect to creating greater explainability of the proposed EEGConvNeXt model because this technique can help bridge the gap between complex neural network outputs and clinical interpretability. These heatmaps serve as a crucial tool for understanding the decision-making process of the proposed EEGConvNeXt model, allowing us to assess the accuracy and reliability of its classifications. Moreover, the ability to visualize the activation patterns associated with different disease states supports the clinical applicability of the model, making it a valuable asset in real-world healthcare scenarios. Fig. 6 shows the results of the Grad-CAM [49] activation function, which has been used to demonstrate the heatmaps on the important regions of the sample images from where AD, FD, and control subjects have been classified.

As evidenced in the heatmaps, we note that the model has successfully recognized the disease-specific symptoms and features with distinct activation regions. Notably, these activations are concentrated in the regions reflecting the disease-specific changes in the CWT images obtained from the EEG signals. One may perceive that in the CWT images of the EEG signals obtained from healthy individuals, low-density and widespread activations suggest that the proposed model does not identify distinct pathological changes in these cases. However, this interpretation is tentative and reflects our current understanding. Additionally, the heatmap images for each class indicate that the proposed model focuses on specific regions to differentiate between AD, FD, and control subjects.

The images above display the Grad-CAM results for classifying AD, FD, and healthy control subjects. Each row corresponds to a different class: (a) AD, (b) FD, and (c) Healthy.

The Grad-CAM heatmaps for AD cases show concentrated activation in specific regions of the spectrogram images. This suggests that the model identified characteristic patterns associated with AD in these areas. These activated regions likely correspond to abnormal EEG signal features indicative of AD.

Similarly, the heatmaps for FD cases exhibit distinct activation zones that differ from those observed in AD cases. This indicates that the model has learned to focus on different EEG signal features more relevant to FD, distinguishing it from other conditions.

For healthy subjects, the central region of the image, dominated by red and yellow hues, represents the areas where the model has the highest level of activation. These are the regions the model deemed most critical for making its classification decision. This high activation suggests that the model identified these regions as containing features indicative of a specific class, such as AD, FD, or healthy controls.

This focus potentially explains the classification behavior of the proposed EEGConvNeXt model, although further analysis may be needed to confirm these observations.

Briefly, Grad-CAM contributes to the explainability of the EEGConvNeXt model by highlighting regions of the EEG spectrograms most relevant for classification. The heatmaps show distinct activation patterns for each class, helping to identify disease-specific features.

For AD, Grad-CAM highlights concentrated regions in the spectrogram images, likely representing abnormal EEG features linked to AD. For FD, the highlighted zones differ, showing that the model focuses on features unique to FD. In healthy controls, widespread and low-density activations indicate the absence of distinct pathological changes.

These heatmaps demonstrate how the model distinguishes between AD, FD, and healthy subjects. They provide valuable insights into the decision-making process and support clinicians in interpreting the model's outputs. This explainability builds trust and makes the model more suitable for clinical use.

3.7. Highlights

The most critical points of this research are discussed in this section. Findings:

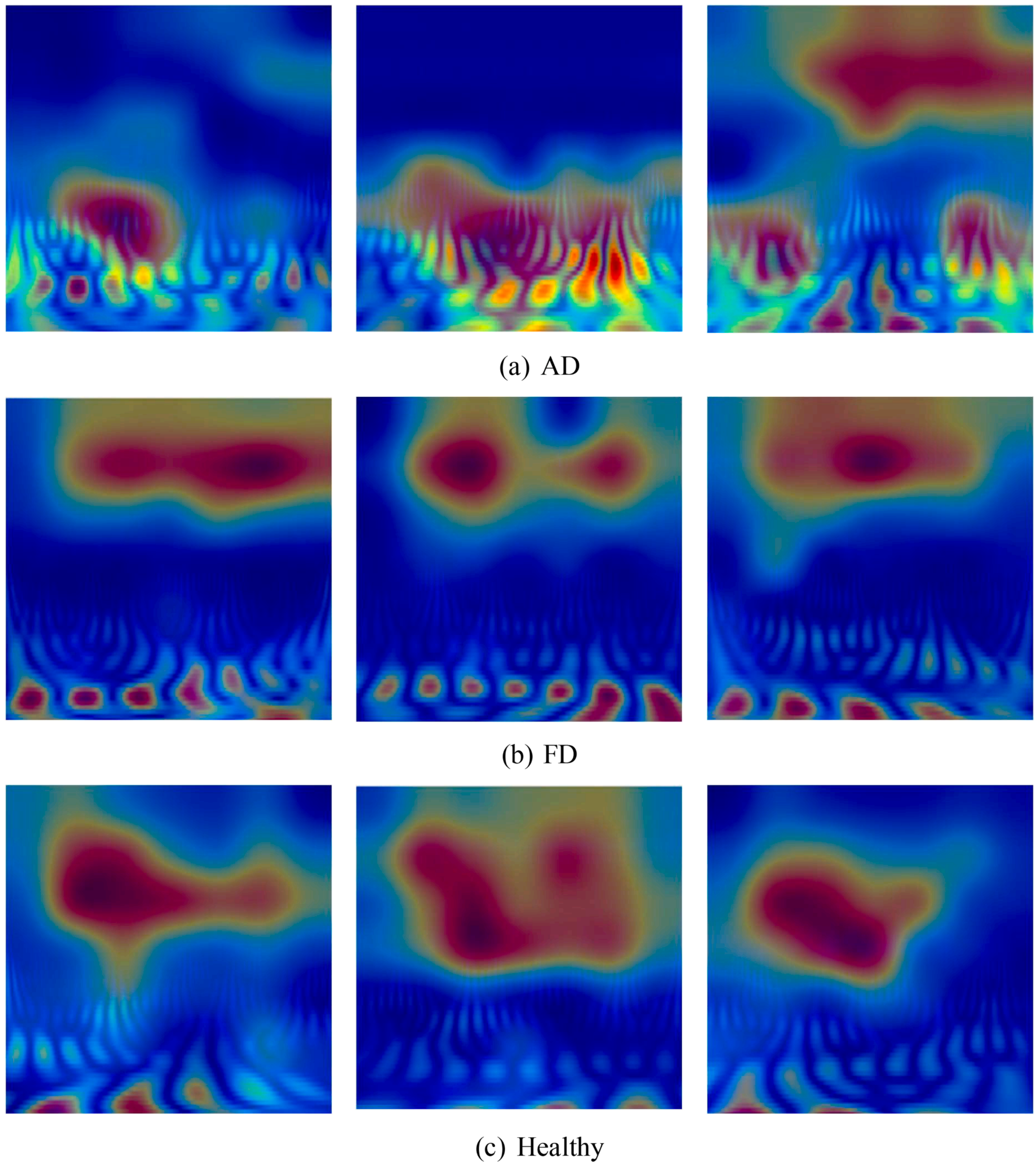

Fig. 6. Heatmaps obtained for the various classes using the Grad-CAM to demonstrate the explainability of the proposed EEGConvNeXt model.

- AD cases have reduced alpha power (0.35 μV^2) and lower temporal stability (0.72 correlation).
- FD cases exhibit increased theta power (0.62 μV^2) and higher kurtosis (3.8) in theta band.
- Grad-CAM activations align with statistical analyses, confirming its effectiveness in highlighting class-specific features.
- All results are statistically significant, with p-values < 0.01 across metrics.AD cases exhibited reduced alpha power (mean $= 0.35$ μV^2, SD $= 0.05$) compared to FD (mean $= 0.42$ μV^2, SD $= 0.07$) and normal (mean $= 0.48$ μV^2, SD $= 0.04$).
- A new CNN model has been presented in this work and this model is termed as EEGConvNeXt. Combines features of ConvNeXt and transformer architectures for increasing feature extraction.
- The proposed EEGConvNeXt model achieved 95.70% accuracy for three-class classification (AD vs. FD vs. Control), 98.21% accuracy for AD vs. FD, 96.32% accuracy for AD vs. Control and 98.67% accuracy for FD vs. Control.
- The EEGConvNeXt is a lightweight model.
- Outperformed existing models like STEADYNet, DICE-Net, and custom CNNs.

Advantages:

- The proposed EEGConvNeXt model combines ConvNeXt and Transformer-inspired approaches. This unique architecture leverages both spatial and temporal information effectively, unlike traditional CNN-based methods.

- Unlike many studies focusing only on binary classifications, the proposed model demonstrates superior performance in multi-class classification (AD, FD, and control), achieving a high accuracy of 95.70% for three-class detection. Moreover, this validation technique is similar to subject-wise techniques since tests and training contain observations of the different subjects.
- The recommended EEGConvNeXt yielded high classification performances on the defined four cases.
- The introduced EEGConvNeXt model showcases superior performance in terms of classification accuracy, recall, precision, and F1-score across both binary and multi-class classification tasks.
- The presented model is a lightweight CNN since it has about 7.5 million learnable parameters.
- Its high accuracy rates in distinguishing AD, FD, and control groups suggest that the model could support early diagnosis and intervention, improving patient outcomes.
- The presented model generates explainable results. Therefore, intelligence assistants can be developed.
- The introduced EEGConvNeXt attained can be implemented for image classification problems. In this aspect, this model can be implemented in imaging devices to obtain intelligent assistants.

Clinical potentials:

- EEGConvNeXt enables clinicians to detect AD and FD early.
- Clinicians use early diagnosis to implement timely interventions, improving patient outcomes.
- Clinicians trust the model more when they see visual explanations (by deploying Grad-CAM like XAI models) of its decisions.
- Community clinics and rural health centers benefit from its lightweight design and portability.
- Hospitals use EEGConvNeXt to automate EEG analysis, reducing reliance on expert neurologists.
- Researchers adapt EEGConvNeXt to detect other neurological disorders like epilepsy and Parkinson's disease.
- Providers use EEGConvNeXt as a cost-effective alternative to expensive imaging tools like MRI and PET scans.
- Medical educators use Grad-CAM visualizations to train clinicians on disease-specific EEG patterns.

4. Conclusions, limitations and recommendations for future work

In this research work, we have proposed a novel EEGConvNeXt model specifically designed for classifying AD, FD, and healthy control classes using EEG signals. To construct this model, the EEG signals were first transformed into images using the CWT method and then fed into the EEGConvNeXt model for classification. The proposed model is lightweight, comprising only 7.5 million parameters. It achieved a classification accuracy of 95.70% for the three-class classification task (AD vs. FD vs. control), 98.21% for the two-class task of distinguishing AD from FD, and 98.67% for distinguishing FD from control subjects.

It is essential to clarify that while the model showed high performance, these results are specific to the dataset used in this study. The ablation study did not compare the EEGConvNeXt model directly against other models in the literature; instead, it focused on understanding the impact of different CNN layers on the model's performance. Specifically, the ablation study explored how variations in the CNN architecture—such as removing or modifying certain layers—affected the classification accuracy. The proposed EEGConvNeXt architecture provided the best balance of performance and computational efficiency, confirming the importance of each layer in the final model's structure.

To further investigate the interpretability of the EEGConvNeXt model, we applied the Grad-CAM algorithm, which provided valuable insights into the regions of the input images (i.e., EEG spectrograms) that the model focused on when making its classifications. This interpretability is crucial for clinical settings, as it enhances the model's transparency and reliability, helping clinicians understand and trust the model's decisions.

In summary, the EEGConvNeXt system, inspired by the ConvNeXt and transformer models, efficiently and accurately identified AD, FD, and control subjects. However, future work should aim to validate these results on more extensive and more diverse datasets and explore further optimization of the model's architecture and interpretability techniques.

The proposed EEGConvNeXt is lightweight, utilizing only 7.5 million parameters. This efficiency is particularly significant in deploying deep learning models in resource-constrained environments, such as mobile or wearable devices, where computational resources are limited. The EEGConvNeXt model incorporates a novel architecture that leverages a patchify-based downsampling block and a hybrid stage inspired by the ConvNeXt and transformer models. This unique design allows for effective feature extraction and classification from EEG signals, achieving high performance across various tasks.

In terms of performance, the EEGConvNeXt model achieved an impressive 95.70% classification accuracy in the three-class classification task (AD, FD, and healthy controls). For binary classifications, the model exceeded 96% accuracy. These results underscore the model's robustness and efficacy in detecting complex neurodegenerative conditions from EEG data. Furthermore, the model's integration of Grad-CAM for generating heatmaps adds a layer of interpretability, which is crucial in clinical settings. By highlighting the regions of EEG spectrograms that contributed most to the model's decisions, these heatmaps help clinicians understand and trust the model's outputs, facilitating its potential adoption in real-world applications.

However, this study is not without its limitations. A primary limitation is the relatively small sample size used to train and evaluate the model—only 88 subjects. While the model performed well on this dataset, the limited diversity and size of the dataset may restrict the generalizability of the findings. Future research should address this by incorporating larger and more diverse datasets, including EEG recordings from different populations and clinical settings. This would help ensure the model can generalize well to a broader range of subjects and conditions.

Another limitation relates to the computational demands of training the EEGConvNeXt model. Despite being lightweight in terms of the number of parameters, the model still requires substantial computational resources for training, which may pose challenges for institutions with limited access to high-performance computing infrastructure. To mitigate this, future work could explore optimization techniques that reduce the computational complexity of the model without sacrificing performance. This could involve experimenting with more efficient model architectures, pruning techniques, or quantization methods that make the model more feasible to train and deploy in less resource-intensive environments.

Additionally, while using Grad-CAM for model interpretability is a significant advantage, it is important to recognize that the interpretability provided by these visualizations is still somewhat limited. Grad-CAM offers insights into which regions of the input data influenced the model's decisions. Still, it does not fully explain the underlying reasoning or the specific features the model relies on. Future research could explore more advanced explainability methods that provide deeper insights into the decision-making process of the EEGConvNeXt model, potentially improving the model's transparency and trustworthiness.

While the EEGConvNeXt model demonstrates promising results in classifying EEG data for detecting neurodegenerative diseases, there are several areas where further research and development are needed. Expanding the dataset, optimizing computational efficiency, and enhancing model interpretability are key steps that should be taken to advance this work and bring the model closer to practical clinical application.

To advance the practical adoption of the proposed EEGConvNeXt model, future researchers may extend this work by using various EEG datasets and further optimizing its computational efficiency. A future study may also explore additional signal-to-image conversion techniques such as time-frequency analysis, wavelet transform, or empirical mode decomposition to further enhance its performance and application to other neurological diseases [50–52].

Funding

The first author acknowledges funding from University of Southern Queensland through providing a Domestic PhD Stipend Scholarship and Australian Government Research and Training (RTP) fee support (2023–2026).

Ethical statement

The authors confirm that only public datasets were acquired and used in accordance with relevant laws and institutional guidelines and have been approved by institutional committee under Ethics Approval ETH2024–0811. No primary data were collected from the human subjects and no personal identifiers were available to the project.

CRediT authorship contribution statement

Madhav Acharya: Writing – original draft, Visualization, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Ravinesh C Deo:** Supervision, Resources, Project administration. **Prabal Datta Barua:** Writing – review & editing, Validation, Supervision, Resources, Project administration. **Aruna Devi:** Writing – review & editing, Supervision, Project administration. **Xiaohui Tao:** Writing – review & editing, Validation, Supervision, Resources, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A. Bhalla, N. Motos, Psychiatric manifestations of Neurodegenerative Disorders: Focus on Frontotemporal Dementia, Arch. Clin. Psych. 50 (5) (2023).
[2] A.S.L. Ng, et al., Distinct network topology in Alzheimer's disease and behavioral variant frontotemporal dementia, Alzheimer's Res. Therapy 13 (2021) 1–16.
[3] M.B. Sandilyan, T. Dening, Brain function, disease and dementia, Nurs. Stand. 29 (39) (2015) 36.
[4] C. Wang, P. Song, Y. Niu, The management of dementia worldwide: a review on policy practices, clinical guidelines, end-of-life care, and challenge along with aging population, Biosci. Trends 16 (2) (2022) 119–129.
[5] A.E. Budson, P.R. Solomon, MemoryLoss, Alzheimer's Disease, and Dementia-E-Book: A Practical Guide for Clinicians, Elsevier Health Sciences, 2021.
[6] A.A. Tahami Monfared, M.J. Byrnes, L.A. White, Q. Zhang, Alzheimer's disease: epidemiology and clinical progression, Neurol. Therapy 11 (2) (2022) 553–569.
[7] S.M. Neuner, T.C.W. Julia, A.M. Goate, Genetic architecture of Alzheimer's disease, Neurobiol. Disease 143 (2020) 104976.
[8] M.A. Rahman, M.S. Rahman, M.J. Uddin, A.N.M. Mamum-Or-Rashid, M.-G. Pang, H. Rhim, Emerging risk of environmental factors: insight mechanisms of Alzheimer's diseases, Environ. Sci. Pollut. Res. 27 (2020) 44659–44672.
[9] A. Antonioni, et al., Frontotemporal dementia, where do we stand? A narrative review, Int. J. Mol. Sci. 24 (14) (2023) 11732.
[10] L.B. Gambogi, H.C. Guimarães, L.C. de Souza, P. Caramelli, Treatment of the behavioral variant of frontotemporal dementia: a narrative review, Dement. Neuropsychol. 15 (3) (2021) 331.
[11] G. Musa, et al., Alzheimer's disease or behavioral variant frontotemporal dementia? Review of key points toward an accurate clinical and neuropsychological diagnosis, J. Alzheimer's Disease 73 (3) (2020) 833–848.
[12] T. Koenig, U. Smailovic, V. Jelic, Past, present and future EEG in the clinical workup of dementias, Psychiatry Res. Neuroimaging 306 (2020) 111182.
[13] P. Maresova, J. Hruska, B. Klimova, S. Barakovic, O. Krejcar, Activities of daily living and associated costs in the most widespread neurodegenerative diseases: A systematic review, Clin. Intervent. Aging (2020) 1841–1862.
[14] M. Pais, et al., Early diagnosis and treatment of Alzheimer's disease: new definitions and challenges, Brazil. J. Psych. 42 (2020) 431–441.
[15] H. Yu, X. Lei, Z. Song, C. Liu, J. Wang, Supervised network-based fuzzy learning of EEG signals for Alzheimer's disease identification, IEEE Trans. Fuzzy Syst. 28 (1) (2019) 60–71.
[16] H. Yu, X. Wu, L. Cai, B. Deng, J. Wang, Modulation of spectral power and functional connectivity in human brain by acupuncture stimulation, IEEE Trans. Neural Syst. Rehabil. Eng. 26 (5) (2018) 977–986.
[17] H. Yu, et al., Evaluation of acupuncture efficacy in modulating brain activity with periodic-aperiodic EEG measurements, IEEE Trans.Neural Syst. Rehabil. Eng. (2024).
[18] H. Yu, F. Li, J. Liu, C. Liu, G. Li, J. Wang, Spatiotemporal dynamics of periodic and aperiodic brain activity under peripheral nerve stimulation with acupuncture, IEEE Trans. Neural Syst. Rehabil. Eng. (2024).
[19] M. Baygin, T. Tuncer, S. Dogan, R.-S. Tan, U.R. Acharya, Automated arrhythmia detection with homeomorphically irreducible tree technique using more than 10,000 individual subject ECG records, Inf. Sci. 575 (2021) 323–337.
[20] I. Castiglioni, et al., AI applications to medical images: From machine learning to deep learning, Phys. Med. 83 (2021) 9–24.
[21] P.M. Rodrigues, B.C. Bispo, C. Garrett, D. Alves, J.P. Teixeira, D. Freitas, Lacsogram: A new EEG tool to diagnose Alzheimer's disease, IEEE J. Biomed. Health Inf. 25 (9) (2021) 3384–3395.
[22] D. Puri, S. Nalbalwar, A. Nandgaonkar, P. Kachare, J. Rajput, A. Wagh, Alzheimer's disease detection using empirical mode decomposition and Hjorth parameters of EEG signal, IEEE (2022) 23–28.
[23] T. Araújo, J.P. Teixeira, P.M. Rodrigues, Smart-data-driven system for alzheimer disease detection through electroencephalographic signals, Bioengineering 9 (4) (2022) 141.
[24] H. Pallathadka, Z. Gardanova, A.R. Al-Tameemi, A.M.B. Al-Dhalimy, E.H. Kadhum, A. huseen Redhee, Investigating cortical complexity in mixed dementia through nonlinear dynamic analyses: a resting-state EEG study, Iran. J. Psych. (2024) 1–10.
[25] P. Kachare, et al., LCADNet: a novel light CNN architecture for EEG-based Alzheimer disease detection, Phys. Eng. Sci. Med. (2024) 1–14.
[26] G. Gopu, S. Sathish, T.S. Shanmathan, G. Suthakar, N. Sivakumaran, M.S. G. Devasena, Multi modal EEG-based classification of Alzheimer's Disease and mild cognitive impairment using hilbert huang transform, IEEE (2024) 1–8.
[27] A. Faghfouri, V. Shalchyan, H.G. Toor, I. Amjad, I.K. Niazi, A tensor decomposition scheme for EEG-based diagnosis of mild cognitive impairment, Heliyon 10 (4) (2024).
[28] M. Nour, U. Senturk, K. Polat, A novel hybrid model in the diagnosis and classification of Alzheimer's disease using EEG signals: Deep ensemble learning (DEL) approach, Biomed. Signal Process. Control 89 (2024) 105751.
[29] V.M. Joshi, P.P. Dandavate, R. Rashmi, G.R. Shinde, D.D. Kulkarni, R. Mirajkar, DEMNET NeuroDeep: Alzheimer detection using electroencephalogram and deep learning, Bullet. Electr. Eng. Inf. 14 (1) (2025) 457–465.
[30] R. Sharma, H.K. Meena, Graph based novel features for detection of Alzheimer's disease using EEG signals, Biomed. Signal Process. Control 103 (2025) 107380.
[31] S. Abadal, et al., Graph neural networks for electroencephalogram analysis: Alzheimer's disease and epilepsy use cases, Neural Netw. 181 (2025) 106792.
[32] K. Rezaee, M. Zhu, Diagnose Alzheimer's disease and mild cognitive impairment using deep CascadeNet and handcrafted features from EEG signals, Biomed. Signal Process. Control 99 (2025) 106895.
[33] H. Altaheri, et al., Deep learning techniques for classification of electroencephalogram (EEG) motor imagery (MI) signals: A review, Neural Comput. Appl. 35 (20) (2023) 14681–14722.
[34] M.S. Nafea, Z.H. Ismail, Supervised machine learning and deep learning techniques for epileptic seizure recognition using EEG signals—A systematic literature review, Bioengineering 9 (12) (2022) 781.
[35] L.-M. Sánchez-Reyes, J. Rodríguez-Reséndiz, G.N. Avecilla-Ramírez, M.-L. García-Gomar, J.-B. Robles-Ocampo, Impact of eeg parameters detecting dementia diseases: A systematic review, IEEE Access 9 (2021) 78060–78074.
[36] G. Xu, et al., A deep transfer convolutional neural network framework for EEG signal classification, IEEE Access 7 (2019) 112767–112776.
[37] E. Perez-Valero, M.Á. Lopez-Gordo, C.M. Gutiérrez, I. Carrera-Muñoz, R. M. Vílchez-Carrillo, A self-driven approach for multi-class discrimination in Alzheimer's disease based on wearable EEG, Comput. Methods Programs Biomed. 220 (2022) 106841.
[38] A.M. Alvi, S. Siuly, H. Wang, A long short-term memory based framework for early detection of mild cognitive impairment from EEG signals, IEEE Trans. Emerg. Topics Comput. Intell. 7 (2) (2022) 375–388.
[39] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A convnet for the 2020s," 2022, pp. 11976–11986.
[40] A. Vaswani, et al., Attention is all you need, Adv. Neural Inf. Process. Syst. 30 (2017).
[41] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, S. Xie, A convnet for the 2020s, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022, pp. 11976–11986.
[42] A. Miltiadous, E. Gionanidis, K.D. Tzimourta, N. Giannakeas, A.T. Tzallas, DICE-net: a novel convolution-transformer architecture for Alzheimer detection in EEG signals, IEEE Access (2023).
[43] A. Miltiadous, et al., A dataset of scalp EEG recordings of Alzheimer's disease, frontotemporal dementia and healthy subjects from routine EEG, Data 8 (6) (2023) 95.

[44] M.J. Warrens, On the equivalence of Cohen's kappa and the Hubert-Arabie adjusted Rand index, J. Class. 25 (2) (2008) 177–183.
[45] D.V. Puri, et al., LEADNet: Detection of Alzheimer's disease using spatiotemporal EEG analysis and low-complexity CNN, IEEE Access (2024).
[46] Y. Chen, H. Wang, D. Zhang, L. Zhang, L. Tao, Multi-feature fusion learning for Alzheimer's disease prediction using EEG signals in resting state, Front. Neurosci. 17 (2023) 1272834.
[47] P.H. Kachare, S.B. Sangle, D.V. Puri, M.M. Khubrani, I. Al-Shourbaji, STEADYNet: Spatiotemporal EEG analysis for dementia detection using convolutional neural network, Cognit. Neurodyn. (2024) 1–14.
[48] M. Rostamikia, Y. Sarbaz, S. Makouei, EEG-based classification of Alzheimer's disease and frontotemporal dementia: a comprehensive analysis of discriminative features, Cognitive Neurodyn. (2024) 1–16.
[49] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," 2016, pp. 2921–2929.
[50] L. Cohen, Time-frequency Analysis, Prentice Hall PTR New Jersey, 1995.
[51] F.A. Alturki, K. AlSharabi, A.M. Abdurraqeeb, M. Aljalal, EEG signal analysis for diagnosing neurological disorders using discrete wavelet transform and intelligent techniques, Sensors 20 (9) (2020) 2505.
[52] P. Ren, et al., Gait rhythm fluctuation analysis for neurodegenerative diseases by empirical mode decomposition, IEEE Trans. Biomed. Eng. 64 (1) (2016) 52–60.