

RESEARCH

Open Access



Advancing Alzheimer's disease detection: a novel convolutional neural network based framework leveraging EEG data and segment length analysis

Md Nurul Ahad Tawhid^{1,2,3*}, Siuly Siuly^{2,3*}, Enamul Kabir⁴ and Yan Li⁴

Abstract

Alzheimer's disease (AD) is a progressive neurodegenerative disorder that primarily affects memory, thinking, and behavior, leading to dementia, a severe cognitive decline. While no cure currently exists, recent advancements in preventive drug trials and therapeutic management have increased interest in developing clinical algorithms for early detection and biomarker identification. Electroencephalography (EEG) is noninvasive, cost-effective, and has high temporal resolution, making it a promising tool for automated AD detection. However, conventional machine learning approaches often fall short in accurately detecting AD due to their limited architectures. We also need to investigate the impact of EEG signal segment length on classification accuracy. To address these issues, a deep learning-based framework is proposed to detect AD using EEG data, focusing on determining the optimal segment length for classification. This framework contains EEG data collection, pre-processing for noise removal, temporal segmentation, convolutional neural network (CNN) model training and classification, and finally, evaluation. We have tested different segment lengths to test the impact on AD detection. We have used both 10-fold and leave-one-out cross-validation techniques and obtained accuracy of 97.08% and 96.90%, respectively, on a publicly available dataset from AHEPA General University Hospital of Thessaloniki. We have also tested the generalizability of the proposed model by testing it to detect frontotemporal dementia and obtained better results than existing studies. Furthermore, we have validated our proposed CNN model using several ablation studies and layer-wise extracted feature visualization. This study will establish a pioneering direction for future researchers and technology experts in the field of neurodegenerative diseases.

Keywords Alzheimer's disease (AD), Electroencephalogram (EEG), Frontotemporal dementia (FTD), CNN, Deep learning

*Correspondence:

Md Nurul Ahad Tawhid

tawhid@iit.du.ac.bd

Siuly Siuly

siuly.siuly@vu.edu.au

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

1 Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder that primarily affects cognitive functions, memory, and behavior [1]. It is the most common cause of dementia among older adults [1, 2]. Early symptoms of AD often include difficulty remembering recent events, names, or conversations, and in later stages, individuals may experience confusion, mood swings, language problems, and challenges in performing daily activities [3]. According to the World Health Organisation (WHO), more than 55 million people suffer from dementia worldwide, with 10 million new cases every year; AD is the most common cause of dementia and may contribute to 60–70% of cases. Presently, dementia ranks as the seventh most common cause of death and stands as a significant contributor to disability and dependence among the elderly worldwide [4]. In the year 2020, more than 50 million cases of dementia were reported [5, 6]. Projections suggest that the number of AD patients is expected to increase to 75 million by the year 2030 and further to 131 million by 2050 [5, 7]. As of now, AD lacks a cure, and the existing treatments offer only limited relief from symptoms [8]. Nevertheless, the early detection of the disease may contribute to slowing its progression and enhancing the quality of life for both patients and their caregivers [9].

Early diagnosis of neurodegenerative disorders such as AD or Frontotemporal dementia (FTD) is crucial for effective therapy selection and improving the quality of life for patients and their caregivers. The challenge lies in distinguishing AD symptoms from normal aging [10, 11]. Currently, examining brain tissue through biopsy remains the gold standard for precise diagnosis, while noninvasive techniques are still being explored. Tests like psychological assessments, brain imaging, and neuronal signal recording aid in the diagnostic process [12, 13]. Despite recent proposals to combine tests, challenges such as unclear correlations and high costs persist [14, 15]. Electroencephalography (EEG) stands out as a promising, affordable, non-invasive, and widely available technique [16–21].

EEG records the electrical potential resulting from neurons' physiological activities [22–25]. Scalp electrodes detect waves created by electric currents during cell depolarization, representing neuronal firing. EEG analysis is typically performed visually by a trained neurologist, but it's challenging due to signal artifacts and also time-consuming, subjective, costly, and error-prone [6, 26]. The complex, non-linear dynamics of neural activities demand sophisticated methods for accurate measurement and higher sensitivity than visual analysis provides.

In recent decades, several AD prediction methods leveraging EEG signals have been introduced. Broadly,

those research works can be divided into two categories based on the applied approaches: (i) traditional machine learning (ML)-based classification approaches and (ii) state-of-the-art deep learning (DL)-based classification approaches. Conventional ML techniques have found extensive use in predicting neurological disorders using EEG signals, particularly for AD. Abásolo et al. [27] used spectral entropy (SpecEn) and sample entropy (SampEn) with statistical analysis to differentiate between AD and HC on a dataset of 22 subjects (11 AD patients and 11 age-matched HCs) and obtained an accuracy of 77.27%. Escudero et al. used multiscale entropy (MSE) with statistical analysis on the same dataset and achieved an accuracy of 90.91% [28]. The authors of [29] used fuzzy entropy on the same AD dataset and obtained an accuracy of 86.36%.

Puri et al. [30] used a tunable Q-wavelet transform (TQWT) to decompose the EEG signal into nine different sub-bands. Then they extracted four features from each sub-band named Katz's fractal dimension, Tsallis entropy, Rényi's entropy, and kurtosis and used those extracted features to train and test support vector machine (SVM), k -nearest neighbor (k NN), ensemble bagged tree (EBT), decision tree, and neural network for detecting AD patients from HC subjects. Using 10-fold CV on a dataset of 12 AD and 11 HC subjects, they obtained an accuracy of 96.20%. The same authors in their other study [31] used empirical mode analysis to generate nine intrinsic mode functions (IMF), and then ten different features were extracted from those IMFs. They selected three Hjorth parameters from those ten features using the Kruskal-Wallis test. Finally, a least-square support vector machine (LS-SVM) is used with a 10-fold CV to validate the proposed model, and a maximum of 92.90% accuracy was obtained. In their another study [32], Puri et al. extracted spectral entropy (SpecE) and Kolmogorov complexity (KC) feature sets from the same 23 subject dataset, and using an SVM classifier, they received an accuracy of 95.6% with a 10-fold CV. In the study [33], the authors have used several spectral features to classify AD from healthy control (HC) subjects. They used regularized linear discriminant analysis to classify the extracted features. Using 10-fold cross validation (CV) on a dataset of 228 subjects (114 AD, 114 HC), they achieved an accuracy of 67%.

While ML methods have proven effective in identifying AD-onset signals from EEG data, it's important to recognize their dependence on labor-intensive pre-processing steps [19, 24]. These methods require explicit feature engineering, where human experts select and define the relevant features for the task. In traditional ML, feature engineering is a critical step where domain knowledge is applied to select and transform the most relevant features

for the problem. This step can be time-consuming and requires expertise [19, 24]. This poses a substantial hurdle for early clinical AD screening, where accuracy and adaptability to diverse environments are pivotal considerations.

On the other hand, deep learning is widely acknowledged for its proficiency in extracting intricate insights from diverse datasets [20, 34–37]. In recent times, there has been a growing emphasis on harnessing DL methodologies to handle and analyze EEG signals in the context of neurological disorders like AD and FTD due to its automated feature extraction and representation learning technique [20, 38, 39]. Moreover, DL can capture complex patterns and features directly from raw data, reducing the need for manual feature engineering [19, 20]. Morabito et al. [40] converted the EEG signal into 2D RGB images using a Mexican Hat-based Continuous Wavelet Transform (CWT). Then those images are used to train and classify using a custom CNN model on a dataset of 46 subjects (23 AD, 23 HC). Their proposed model achieved an accuracy of 85% in classifying AD from HC subjects. Ferri et al. [41] proposed artificial neural networks (ANNs) with stacked autoencoders to classify AD from HC using resting state EEG data. Using a dataset of 89 AD and 45 HC subjects, they achieved an accuracy of 80%. Ieracitano et al. [42] used power spectral density (PSD) to represent the EEG signal into 2D grayscale images. Then they used a customized CNN to perform a classification task on those images. Using a dataset of 126 subjects (63 AD, 63 HC), they achieved an accuracy of 92.95%. Alessandrini et al. [43] used a recurrent neural network (RNN) to classify AD from HC. Using a dataset of 35 subjects (20 AD, 15 HC), they achieved an accuracy of 79.3%. Miltiadous et al. [5] proposed a Dual-Input Convolution Encoder Network (DICE-net) for AD EEG data classification. DICE-net consists of convolution, transformer encoder, and feed-forward layers, which are used to classify the band power and coherence features extracted from the denoised signal data. Using a dataset of 88 subjects (36 AD, 23 FTD, and 29 HC), they achieved an accuracy of 83.28% for AD vs. HC classification for Leave-One-Out-validation (LOOV) and 74.96% for FTD vs. HC classification. Chen et al. [38] introduced a two-branch network architecture, comprising CNN and visual transformers (ViTs), to classify EEG data for AD and FTD detection. Using the same dataset as the authors [5], they achieved an accuracy of 87.33% and 82.98% for the classification of AD vs. HC and FTD vs. HC, respectively.

Although some research has been done using DL-based methods for AD classification, it is not enough and has scope to improve in terms of accuracy and

performance. Moreover, most of the DL-based methods extracted handcrafted features to feed into the DL models, which is why those methods have not fully used the power of feature extraction by the DL models themselves. Moreover, we require more precise and effective diagnostic tools based on deep learning, which can harness the wealth of information available in EEG recordings. Moreover, There is no consensus on the optimal epoch duration for segmenting EEG signals, leading to variability across research studies [5, 20, 27–29, 38–40, 44]. Typically, the EEG window length is chosen between 5 to 12 s, either arbitrarily or based on previous literature [44]. The aim of this study is to fill these gaps by using DL-based models on the raw EEG data for AD classification and find an optimal segment length for classification.

In this study, we have developed a DL-based CNN model to perform the classification of AD in HC subjects using EEG data. At first, EEG data is pre-processed for noise removal using different techniques and then the signals are segmented into small and different time frames to check the impact of the segment length. After that, a DL-based CNN is trained, and a classification process is carried out for AD vs. HC. We have used both 10-fold and leave-one-out cross-validation (CV) techniques to validate the proposed model. To check the generalizability of the proposed model, we have also used the proposed model to perform classification between FTD and HC subjects. Finally, the classification performance of the CNN model on different segment lengths is compared to find out the impact of the segment length in the classification process.

The major contributions of this research are listed below:

1. A novel framework is designed to identify the impact of segmentation on the identification of AD and FTD from HC subjects.
2. A novel DL-based CNN model is developed for the classification of both AD and FTD.
3. Explore the performance using different evaluation techniques.
4. Increase classification performance over existing approaches using the same dataset.

2 Methodology and materials

The proposed methods consists of four steps: EEG data collection, data pre-processing for noise removal, data segmentation, CNN model training and classification. Figure 1 gives an overview of the proposed framework. A

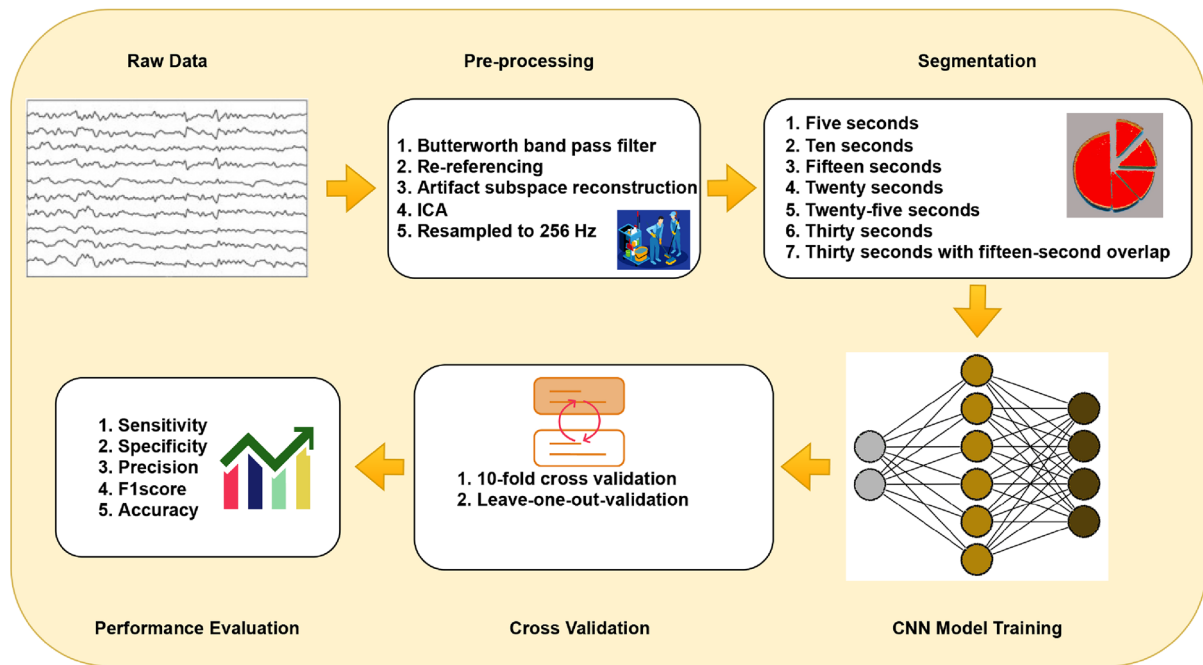


Fig. 1 Schematic illustration of the proposed framework and steps involved in the analysis

detailed discussion of the above mentioned four steps is given in the below subsections.

2.1 EEG data collection

Here, we have used the publicly available EEG dataset of the research study [5]. It contains EEG recordings from 88 participants from the Department of Neurology of AHEPA General University Hospital of Thessaloniki. The participants were divided into three groups: Alzheimer's disease (AD), Frontotemporal Dementia (FTD), and cognitively normal (CN). Cognitive and neuropsychological states were assessed using the Mini-Mental State Examination (MMSE), where lower scores indicated greater cognitive decline. The AD group consisted of 36 participants (13 males, 23 female; mean age 66.4 ± 7.9) with an average MMSE score of 17.75. The FTD group included 23 participants (14 males, 9 female; mean age 63.6 ± 8.2) with an average MMSE score of 22.17. The CN group comprised 29 participants (11 males, 18 female; mean age 67.9 ± 5.4) with a perfect MMSE score of 30. The median duration of the disease was 25 months.

EEG data was recorded from 19 electrode channels (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2) according to the 10–20 international system with two reference electrodes (A1 and A2). Eye closed, resting state EEG data was recorded using a sampling frequency of 500 Hz. Recordings lasted about

13.5 min for AD group (range: 5.1 to 21.3), 12 min for FTD group (range: 7.9 to 16.9), and 13.8 min for CN group (range: 12.5 to 16.5).

The dataset is publicly available online, and each participant freely consented to the publication of the data when it was gathered. Since the published data don't contain any information that may be used to identify the respondents or jeopardise their confidentiality, no ethical approval was required for our study.

2.2 Data pre-processing for noise removal

The EEG signal pre-processing consists of several steps. Initially, a Butterworth band-pass filter is used to allow signals within a specific frequency range (in this case, between 0.5 Hz and 45 Hz) to pass through while attenuating signals outside this range. The Butterworth filter is known for its flat frequency response within the passband. The transfer function $H(s)$ for a Butterworth filter is expressed as:

$$H(s) = \frac{1}{1 + \left(\frac{s}{\omega_c}\right)^{2N}} \quad (1)$$

where N is the order of the filter, ω_c is cut-off frequency and s is complex frequency (Laplace variable). For a band-pass filter, the design combines both low-pass and high-pass Butterworth filters. The band-pass filter's response is derived using:

$$H(s) = \frac{\omega_H \cdot s}{(s^2 + \omega_L \cdot s + \omega_L^2)(s^2 + \omega_H \cdot s + \omega_H^2)} \quad (2)$$

where $\omega_L = 2\pi f_L$ is lower cut-off angular frequency (corresponding to 0.5 Hz), $\omega_H = 2\pi f_H$ is higher cut-off angular frequency (corresponding to 45 Hz). The filter coefficients depend on the filter order (N) and the design specifications. The signals are then re-referenced to A1-A2.

Subsequently, the artifact subspace reconstruction (ASR) [45] routine was employed, a method designed to eliminate transient or high-amplitude artefacts [46], by discarding data periods with a standard deviation greater than 17 within a 0.5-second window. It operates by identifying and reconstructing artifact components in the signal's subspace. The mathematical foundation of ASR involves the following steps:

- **Covariance Matrix Estimation:** A robust covariance matrix (C) is computed from clean EEG data segments. This matrix represents the statistical properties of the clean data.

$$C = \frac{1}{N} \sum_{i=1}^T T \quad (3)$$

where x_i is clean EEG data sample and N is number of samples.

- **Artifact Detection:** The algorithm identifies artifact components by comparing the variance of incoming EEG data against a threshold derived from the clean data covariance matrix. Principal Component Analysis (PCA) is often used to decompose the signal into components.

$$\text{Variance Threshold} = \alpha \cdot \text{Variance of Clean Data} \quad (4)$$

where α is a scaling factor.

- **Reconstruction:** The artifact components are reconstructed using the clean covariance matrix. The reconstruction ensures that the signal remains within the statistical boundaries of clean EEG data.

$$x_{\text{reconstructed}} = C^{-1} \cdot x \quad (5)$$

where x is original EEG data and $x_{\text{reconstructed}}$ is artifact-corrected EEG data.

The Independent Component Analysis (ICA) method (using the RunICA algorithm) was then used to transform the 19 EEG signals into 19 ICA components. ICA components identified as “eye artifacts” or “jaw artifacts” by the automated classification routine “ICLabel” in the EEGLAB platform [47] were automatically removed. Details of those steps are given below:

- **Transforming EEG Signals into ICA Components:** Given the EEG data matrix X, where each row represents an EEG channel and each column represents time samples, ICA decomposes X into:

$$X = A \cdot S \quad (6)$$

where $X \in \mathbb{R}^{n \times T}$ is the observed EEG signals (19 channels in this case), $A \in \mathbb{R}^{n \times n}$ is the mixing matrix, $S \in \mathbb{R}^{n \times T}$ is the source matrix (independent components), n is the number of channels (19 here), and T is the number of time samples. Using the RunICA algorithm, the unmixing matrix W is estimated such that:

$$S = W \cdot X \quad (7)$$

where W is the inverse of A .

- **Identifying Artifact Components:** The ICLabel plugin in the EEGLAB platform is an automated classification tool that labels the ICA components (S) based on their spatial and temporal characteristics. It assigns each component a probability of being: Brain activity, Eye artifacts (e.g., blinks or saccades), Jaw artifacts (e.g., muscle movements), Other noise sources. Components classified as “eye artifacts” or “jaw artifacts” with a probability exceeding a predefined threshold are flagged for removal.
- **Reconstructing Clean EEG Signals:** Once the artifact components are identified, they are excluded from the reconstruction process. Let S_{clean} represent the source matrix after removing artifact components. The clean EEG signals X_{clean} are reconstructed as:

$$X_{\text{clean}} = A \cdot S_{\text{clean}} \quad (8)$$

where S_{clean} only includes the components classified as brain activity or other non-artifact sources. This process effectively removes artifacts while preserving the integrity of the neural signals.

Finally, the signal is resampled to 256 Hz since it is commonly used sampling frequency for EEG data and is computationally less costly than high-frequency bands [48].

2.3 Segmentation of the EEG signals

EEG data segmentation is a fundamental step in EEG signal processing that enables researchers to analyze, interpret, and extract meaningful information from EEG recordings by breaking them into manageable, contextually relevant segments [15, 19, 49]. Furthermore, the lack of data is a significant issue for deep learning-based EEG signal processing systems. To address this issue, researchers frequently use the segmentation technique.

This method increases the data sample size while maintaining an equal ratio by dividing the original EEG data into brief, useful pieces and giving them the same level as the original signal [14, 15, 19, 20, 23, 49, 50]. In this study, we have segmented the filtered signals into seven different time segments to check the effect of the segment length on the classification process. Here, we have tried six non-overlapping segment lengths: five seconds (5 s), ten seconds (10 s), fifteen seconds (15 s), twenty seconds (20 s), twenty-five seconds (25 s), thirty seconds (30 s), and one overlapping segment length: a thirty-second segment with a fifteen-second overlap (30 s+ol). The overlapping segmentation is tested to compare the performance of the proposed model with the study [5] that has published this dataset.

2.4 Proposed model for feature extraction and classification

In this study, we have designed a DL-based CNN model to perform classification of the signal segments. CNN, a widely recognized deep learning model, has proven to be highly effective in classification tasks. It achieves this by autonomously identifying relevant features and categorizing data into multiple classes [51]. Architectures based on CNN models consist of multiple convolution layers, allowing them to acquire both low-level features and high-level features, as well as semantic representations [51]. The acquisition of hierarchical representations aids CNNs in grasping intricate data patterns [52].

The convolutional layer is a key component in CNNs, pivotal for its ability to analyze and extract essential features from input data. This layer plays a fundamental role in tasks such as image recognition, object detection, and various other classification applications [24, 52].

In the context of a CNN, the convolutional layer performs a convolution operation, where a set of learnable filters, also known as kernels, scan across the input data. These filters systematically slide over the input, capturing local patterns and relationships. The result of this convolution is a set of feature maps that represent different aspects of the input data [51]. Each filter in the convolutional layer has learnable parameters that are adjusted during the training process through back-propagation. This enables the network to adapt and recognize hierarchical features in the data, ultimately contributing to its ability to understand and classify complex patterns [51, 52].

The intrinsic functioning of the convolutional layer, characterized by a spatial filter dimension denoted as $M \times N$ and encompassing C channels, can be articulated as follows (equation 9):

$$Y_{i,j,k} = f \left(\sum_{m=1}^M \sum_{n=1}^N \sum_{c=1}^C W_{m,n,c,k} X_{i+m-1,j+n-1,c} + b_k \right) \quad (9)$$

In this context, $Y_{i,j,k}$ represents the value situated at the i^{th} row, j^{th} column, and k^{th} channel of the output feature map. Similarly, $X_{i+m-1,j+n-1,c}$ denotes the value located at the $(i+m-1)^{th}$ row, $(j+n-1)^{th}$ column, and c^{th} channel of the input feature map. The weight $W_{m,n,c,k}$ pertains to the specific weight associated with the m^{th} row, n^{th} column, c^{th} channel of the filter pertaining to the k^{th} channel of the output feature map. The term b_k signifies the bias term corresponding to the k^{th} channel of the output feature map. The activation function $f(\cdot)$ is subsequently applied to the element-wise sum of these components.

The equation 9 calculates the dot product between the filter weights W and the corresponding region of the input feature map X . This result is then aggregated with the bias term b , and the activation function $f(\cdot)$ is introduced to instill non-linearity in the computation.

To do the classification task on the generated signal segments, we have developed a CNN model. The model contains six convolution (Conv) layers, three max pooling layers, four dropout layers, two dense layers, and a classification layer. The model can be broken down into four blocks, among which the first three have two Conv layers, followed by a max pooling layer and a dropout layer. The last block has two dense layers and a classification layer. In the first block, two Conv layers have 16 filters with a 3×3 kernel size, followed by a max pooling layer with a 1×3 kernel and a 20% dropout layer. The second block's Conv layers have 32 filters each, and the dropout layer has 25% dropout rate. The third block is a copy of the first block, except that the dropout layer has a dropout of 25%. The first and second dense layers have 256 and 128 filters, respectively, followed by a 50% dropout layer. The final classification layer is activated by a softmax activation function to perform classification between HC vs. AD or HC vs. FTD. The Adam optimizer and the categorical cross-entropy loss function are used to build the proposed model. Table 1 lists the detailed configuration of those layers.

2.5 Performance evaluation processes and parameters

We have evaluated the performance of the proposed model using a publicly available dataset to perform two different categorization tasks: AD vs. HC and FTD vs. HC. We have used both 10-fold cross validation (CV) and leave-one-out-validation (LOOV) techniques to validate the proposed model. These are the well-known

Table 1 Architecture of the used CNN model

Layers	# Filter	Kernel size	Option
Conv2D	16	3×3	activation=relu
Conv2D	16	3×3	activation=relu
MaxPooling2D		1×3	strides=(2,2)
Dropout			20%
Conv2D	32	3×3	activation=relu
Conv2D	32	3×3	activation=relu
MaxPooling2D		1×3	strides=(2,2)
Dropout			25%
Conv2D	16	3×3	activation=relu
Conv2D	16	3×3	activation=relu
MaxPooling2D		1×3	strides=(2,2)
Dropout			25%
Flatten			
Dense	256		activation=relu
Dense	128		activation=relu
Dropout			50%
Dense (classifier)	2		
Total params: 1,209,058			
Trainable params: 1,209,058			
Non-trainable params: 0			

techniques in machine learning to validate the performance of a model [19].

Five well-known evaluation parameters are used to evaluate the performance of the proposed framework, namely: sensitivity (Sen), specificity (Spec), precision (Prec), F1 score (F1), and accuracy (Acc). Equation ((10)) - ((14)) is used to calculate those parameters:

$$\text{Sensitivity}(\text{Sen}) = \frac{TP}{TP + FN} * 100 \quad (10)$$

$$\text{Specificity}(\text{Spec}) = \frac{TN}{TN + FP} * 100 \quad (11)$$

$$\text{Precision}(\text{Prec}) = \frac{TP}{TP + FP} * 100 \quad (12)$$

$$\text{alignedF1score}(\text{F1}) = \frac{2 TP}{2 TP + FP + FN} \quad (13)$$

$$\text{Accuracy}(\text{Acc}) = \frac{TP + TN}{TP + FP + TN + FN} * 100 \quad (14)$$

Here,

- TP implies the number of correctly identified patients.
- TN implies the number of correctly identified HC.
- FP implies the number of falsely identified HC subjects as patient.
- FN implies the number of falsely identified patients as HC.

We have also used the receiver operating characteristic (ROC) graph, a highly useful tool for visualizing the classifier's reliability, created by graphing sensitivity on the Y-axis and 1-specificity on the X-axis. These criteria allow us to grasp an idea about the classifier's behavior on the test data [6, 14, 20, 23, 50, 53, 54].

3 Results and discussion

In this section, we begin by delving into the specifics of the experimental setup, followed by an in-depth exploration of the results obtained. Ultimately, this section concludes with a thorough discussion of the outcomes.

3.1 Experimental setup

In this study, we have used various segmentation length to check the impact of the segment length in AD and FTD detection from HC subjects. We have checked seven different segment lengths (5 s, 10 s, 15 s, 20 s, 25 s, 30 s and 30 s+ol). For these segment lengths, total number of produced subjects of each category are given in Table 2.

After the segmentation process, the resulting dataset is divided into 10 sub-parts, as we have used the 10-fold CV. On the other hand, for LOOV, all the segments from a particular subject are left out of the training process, and the remaining subjects' segments are used to train the model, while the left-out subjects' segments are used to test the trained model. The experiments are carried out on a computer with an AMD Threadripper Pro processor, 256 GB of RAM, and 48GB of graphics memory. We have used 50 epochs for training the model, as the model starts overfitting after those epochs, and we have used a batch size of 32 to train the model.

Table 2 Total number of segments generated after segmentation process of the signals

	5 s	10 s	15 s	20 s	25 s	30 s	30 s+ol
HC	4808	2397	1592	1192	651	542	1070
AD	5807	2895	1924	1439	1146	953	1888
FTD	3304	1646	1093	815	952	790	1563

3.2 Results

In this research work, we have developed a framework to perform two classification tasks: AD vs. HC and FTD vs. HC, with different segment lengths to check the impact of the segment duration on the identification process. We have used seven different segment lengths with two cross validation techniques: 10-fold CV and LOOV. Details of those two CV results are given in below two subsections.

3.2.1 Results for ten-fold cross validation

In this CV technique, we have divided the generated dataset into ten equal or nearly equal sub-parts and then nine of them are used to train the proposed model and the left over part is used to test the trained model. This process is repeated ten times (10-fold) so that each signal segment belong to the test set exactly once. The final result is calculated by averaging the results over the ten folds. Tables 3 shows the average evaluation parameter values with standard deviation over the 10-fold CV.

From Table 3, we can see that, for both AD vs. HC and FTD vs. HC, the proposed model has produced the best result with a segment size of 5 s. For AD vs. HC, it is 97.08% and for FTD vs. HC, it is 98.14%. On the other hand, 30 s segment produced the lowest accuracy for both classifications which are 86.80% and 89.86% for AD vs. HC and FTD vs. HC, respectively. 30 s+ol has produced the second best accuracy of 95.54% and 96.39%, accordingly. Other evaluation parameters also displayed the same behavior as of accuracy. To further illustrate

those behavior, we have plotted those parameters (sensitivity, specificity, precision and accuracy) against the segment length and given in Figs. 2 and 3 for AD vs. HC and FTD vs. HC, respectively.

From the Fig. 2 and Table 3, we can see that in non-overlapping segments, increase of segment size decreases the accuracy from 97.08% to 86.80%. For 30 s segment with overlap increases the accuracy to 95.54% which is due to the overlapping segmentation process. This decreasing pattern is also observed in F1 and precision values.

In case of sensitivity, the values decreased with the increase of segment length except for 30 s, where the sensitivity increased a little bit (0.84) from 25 s. In case of specificity, 20 s segment length has a increase in the specificity value compared to the 15 s. Other than that, all other segment length have followed the decreasing pattern similar to the other parameters.

On the other hand, for FTD vs. HC, from Fig. 3 and Table 3 we can see that, for all the evaluated parameters, increases of the segment length has a negative effect on the performance of the proposed model except for the overlapping 30 s segmentation. Similar to the AD vs. HC, 30 s with overlap produces the accuracy closer to the accuracy of 5 s.

3.2.2 Results for leave one out cross validation

In this CV process, we have left out all the segments from a subject from the training process and the model

Table 3 Average evaluation parameter values with standard deviation for 10-fold CV for the two classification tasks with different segment lengths

Segment Length	AD vs HC				
	Sensitivity	Specificity	Precision	F1	Accuracy
5 s	97.61	96.40	97.07	0.97	97.08
10 s	95.49	94.10	95.16	0.95	94.88
15 s	94.59	91.92	93.59	0.94	93.43
20 s	91.53	93.30	94.29	0.93	92.32
25 s	90.15	84.45	87.49	0.89	87.56
30 s	90.99	81.47	86.07	0.88	86.80
30 s + overlap	95.94	94.96	95.88	0.96	95.54
Segment Length	FTD vs HC				
	Sensitivity	Specificity	Precision	F1	Accuracy
5 s	98.04	98.21	97.41	0.98	98.14
10 s	96.83	97.29	96.09	0.96	97.10
15 s	93.94	95.85	93.94	0.94	95.05
20 s	92.51	94.87	92.53	0.92	93.87
25 s	92.16	93.94	91.45	0.92	93.27
30 s	86.04	92.54	89.01	0.87	89.86
30 s + overlap	96.01	96.70	95.37	0.96	96.39

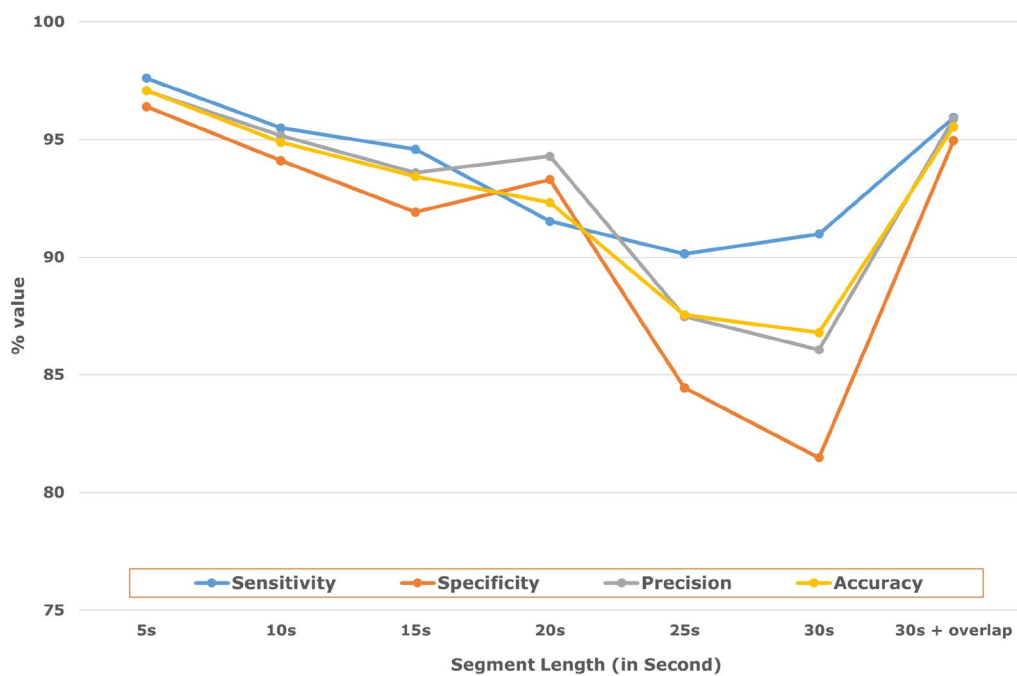


Fig. 2 Comparison of the evaluation parameters against the different segment lengths for AD vs. HC classification

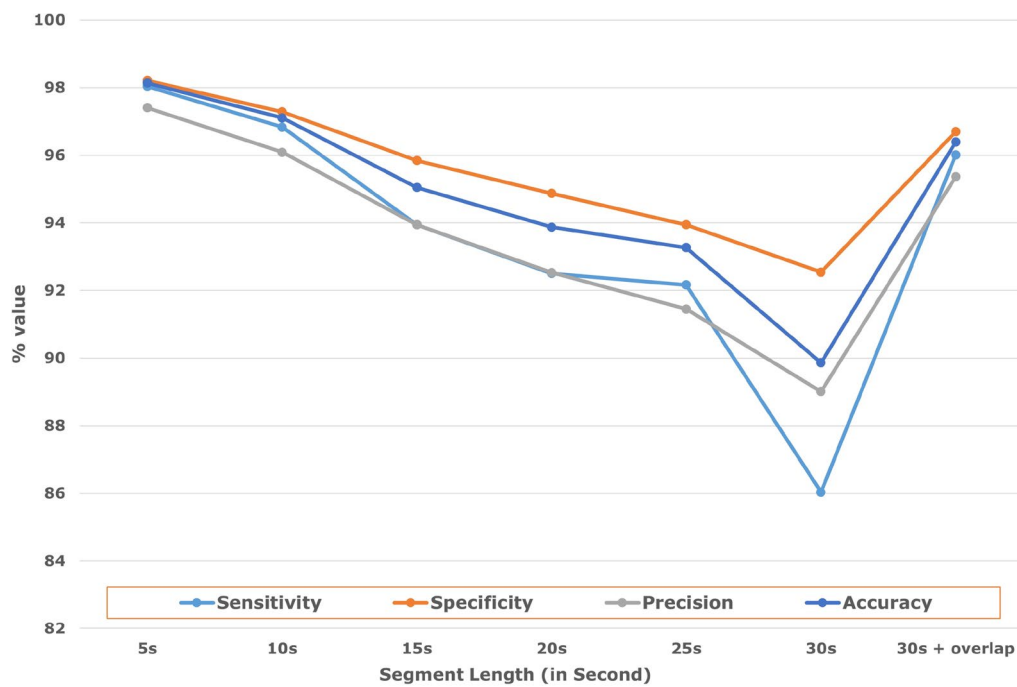


Fig. 3 Comparison of the evaluation parameters against the different segment lengths for FTD vs. HC classification

is trained using the segments of the remaining subjects from the dataset. After that, the left out subject's segments are used to test the trained model. This process is

repeated for all the subjects in the dataset. Final results are calculated using the average of the all subjects accuracy and given in Table 4.

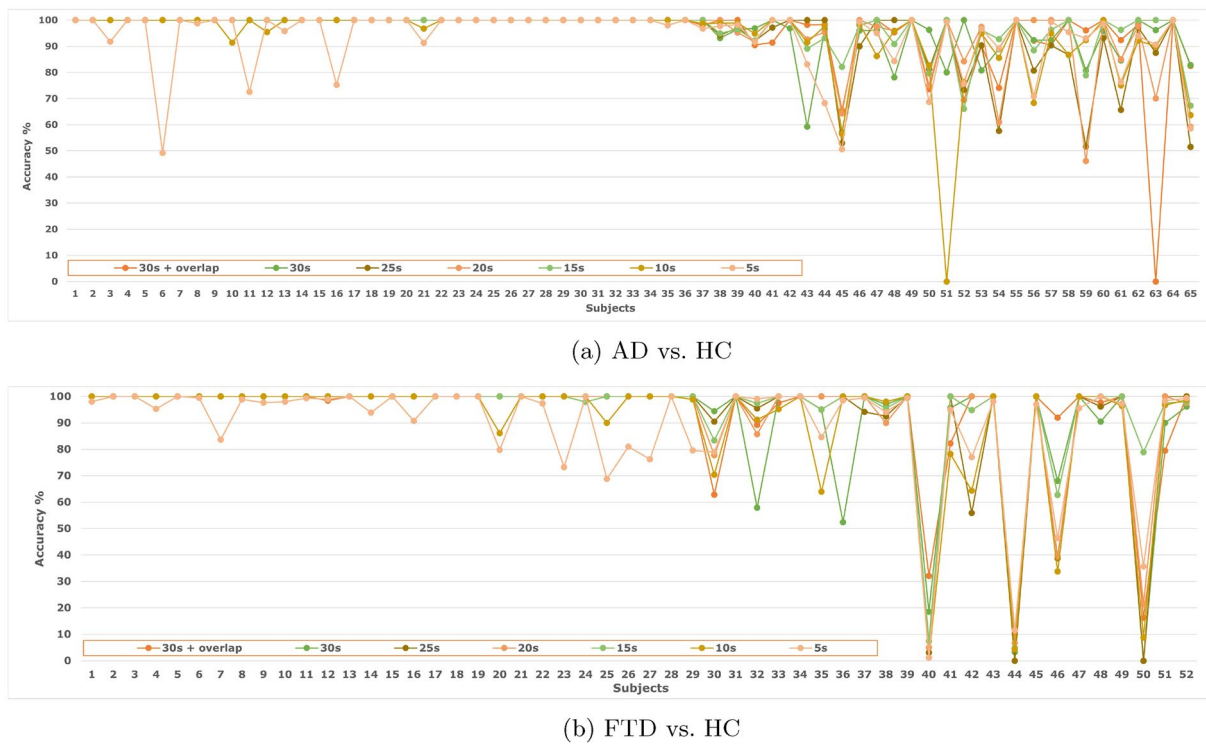


Fig. 4 Comparison of accuracy for the tested different segment lengths in LOOV. Figure 4a and b shows the curves for AD vs. HC and FTD vs. HC, respectively

Table 4 Average accuracy values for the leave-one-out-validation for different segment lengths for the two classification tasks

Accuracy %	5 s	10 s	15 s	20 s	25 s	30 s	30 s+ol
AD vs HC	92.98 ± 12.43	93.90 ± 14.87	96.90 ± 7.20	95.51 ± 11.27	94.31 ± 12.41	95.96 ± 8.77	95.29 ± 14.09
FTD vs HC	88.78 ± 21.15	89.93 ± 24.29	94.50 ± 18.73	92.76 ± 22.21	91.66 ± 24.69	91.52 ± 23.04	93.36 ± 19.49

From Table 4, we can see that for both the classification tasks, 15 s segment has produced the best result. For AD vs. HC, it is 96.90% and for FTD vs. HC, it is 94.50%. To further check the details of the LOOV result, we have plotted the subject-wise accuracy for both the classification tasks and provided in Fig. 4 where Fig. 4a shows the subject vs. accuracy for AD vs. HC and Fig. 4b shows the subject vs. accuracy for FTD vs. HC.

For AD vs. HC classification, there were 65 subjects in the dataset among them 48 subjects have the accuracy of 100% with the segment length of 15 s. 20 subjects have the accuracy between 90% to less than 100% and the remaining 5 subjects have accuracy greater than 60% as shown in Fig. 4a.

On the other hand, for FTD vs. HC, 40 out of 52 subjects have a 100% accuracy with 15 s segment length. Seven subjects have accuracy between 90% to less than 100%, three subjects have between 60% to less than 90%

and the remaining two subjects have accuracy below 50% as depicted in Fig. 4b.

Here, we have systematically evaluated segment lengths ranging from 5 s to 30 s, including a 30 s segment with overlap, across both classification tasks: AD vs HC and FTD vs HC. Our findings show that although shorter segments like 5 s have achieved slightly higher peak performance in 10-fold CV (e.g., 97.08% accuracy for AD vs HC and 98.14% for FTD vs HC), the performance using 30 s segments with overlap remains comparably high (95.54% and 96.90%, respectively), with only a marginal reduction in accuracy. Similar results have obtained for LOOV, 15 s segment has achieved high accuracy while 30 s with overlap has achieved close to it.

Importantly, the use of 30 s segments with overlap offers a practical advantage. It provides a larger context window, which is beneficial for capturing subtle patterns in EEG data relevant to neurodegenerative disorders.

Additionally, the overlap strategy increases the effective number of training samples without requiring additional EEG recordings, thereby enhancing generalizability and model robustness while keeping training costs relatively low. This setup strikes a balance between performance and computational efficiency, making it especially suitable for clinical applications where both accuracy and scalability are important.

3.3 Discussion

In this study, we have developed a CNN model to classify AD and FTD from HC subjects. We have also tested the effect of different segmentation length on the classification performance of the proposed model. Below subsections discuss about the performance of the proposed model from different aspects.

3.3.1 Layer-wise extracted feature visualization

Here, we have developed a small scale CNN model with only six Conv layers with only 1,209,058 parameters. The proposed model is simple, and less memory and time consuming but performs very well on both AD vs. HC and FTD vs. HC classification. Moreover, its performance is verified using both 10-fold and LOOV CV. To further illustrate the classification process of the proposed model, we have used the T-distributed stochastic neighbor embedding (t-SNE) to visualize the layer-wise classification process and given in Fig. 5.

t-SNE is a machine learning algorithm used for dimensionality reduction and visualization of high-dimensional data on a 2D or 3D space. It was introduced by Laurens van der Maaten and Geoffrey Hinton in 2008 [55] and is particularly effective at revealing the underlying structure

and patterns in complex datasets. We employed t-SNE visualization to create two-dimensional (2D) representations of the features extracted from each layer of the proposed model. This approach facilitates the visualization of the model's layer-wise extracted features during the classification process. In the Fig. 5, we have visualized the features extracted by the proposed model for AD vs. HC classification using 5 s segment length for 800 test segments for a single fold. The figure display a 2D map of multidimensional feature vectors, where each symbol represents an individual sample from the test set (red symbol is for AD and blue is for HC).

3.3.2 Ablation study of the proposed model

Generally, ablation study is a type of experiment conducted to analyze the contribution of individual components or modules in a system, such as a neural network model. In the context of CNN, an ablation study helps to understand the impact of different architectural elements, layers, or features on the overall performance of the model. To validate the structure of the proposed model, we have conducted several ablation studies on both AD vs. HC and FTD vs. HC classification tasks and reported in Table 5.

Here we have used the 5 s segment's result as the base for both AD vs. HC and FTD vs. HC as this segment length have produced the best classification result. From the Table 5, we have used ablation techniques like, removing a block, adding a block and changing the number of filters, but none of the techniques have produced better result than the proposed model. All the tested ablation methods prove that the proposed model gives a better result than the other tested models.

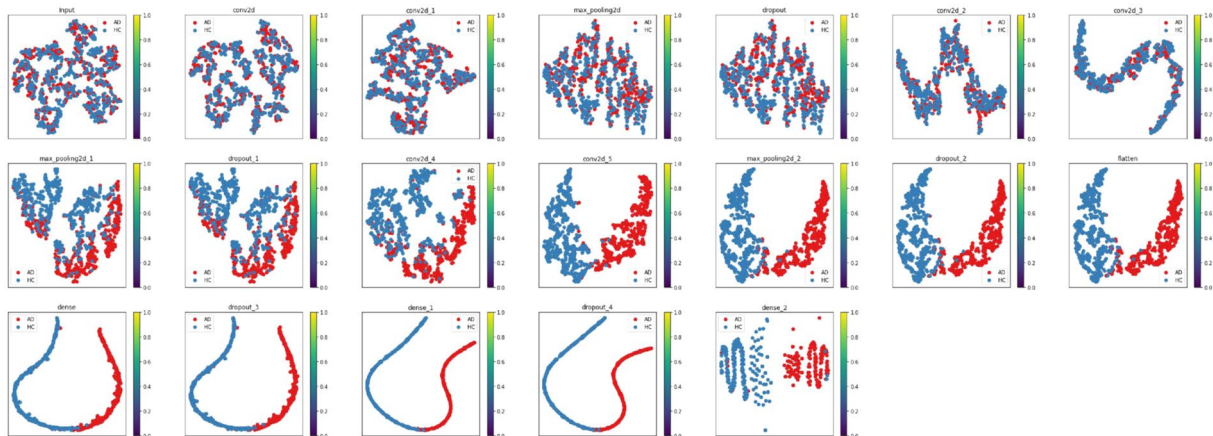


Fig. 5 Visualizing the layer-wise classification process of the proposed CNN model using t-SNE images involved plotting features from 800 test segments across the input layer to the output layer for the AD vs. HC classification using 5 s segment length for a single fold. Initially, at the input layer, no distinct cluster between the two classes (AD vs. HC) was evident. However, as the data advanced through the hidden layers to the output layer, distinct and separable clusters for the two classes emerged

Table 5 Ablation study results (accuracy %) of the proposed model for both AD vs. HC and FTD vs. HC using 5 s segment length

	AD vs. HC	FTD vs. HC
Base (proposed model)	97.08	98.14
Removed the 2nd block	92.79	95.34
Removed the 3rd block	93.44	95.65
Doubled the filters of the conv layer in 1st block	93.50	96.33
Doubled the filters of the conv layer in 3rd block	93.02	96.40
Halved the filters of the conv layer in 2nd block	95.59	97.86
Added a duplicate of 2nd block after the 2nd block	76.56	94.97

Table 6 Comparison with the existing studies of MCI classification

Classification	Study	Segmentation	Validation	Accuracy (%)
AD vs HC	[5]	30 s + ol	LOOV	83.28
	[38]	4 s	10-fold	87.33
	Proposed work	30 s + ol	LOOV	95.29
	Proposed work	15 s	LOOV	96.90
	Proposed work	30 s + ol	10-fold	95.54
	Proposed work	5 s	10-fold	97.08
	[5]	30 s + ol	LOOV	74.96
	[38]	4 s	10-fold	82.98
FTD vs HC	Proposed work	30 s + ol	LOOV	93.36
	Proposed work	15 s	LOOV	94.50
	Proposed work	30 s + ol	10-fold	96.39
	Proposed work	5 s	10-fold	98.14

3.3.3 Comparison with the existing studies

Finally, we have compared the performance of the proposed framework with the existing works that have used the same dataset as ours and reported in Table 6.

From Table 6, we can see that this dataset was only used in two studies to perform the classification tasks (AD vs. HC and FTD vs. HC) [5, 38]. The authors of [5] have used 30 s segment with 15 s overlap to prepare the data, which is why we have also used the same segmentation along with other segmentation lengths to compare the performance with them. Our proposed framework

has performed better than both the studies [5, 38] in all the classification tasks as shown in Table 6.

Although our proposed framework has shown promising results in enhancing classification accuracy for both AD and FTD through EEG signals. However, as with any study, there are some hurdles to acknowledge. One prominent challenge is the limitation posed by our dataset. The population size and the duration of recording in the datasets used are somewhat restrictive. Ideally, a more extensive and diverse dataset could provide a broader perspective on the effectiveness of our framework. Another noteworthy point is the absence of alternative datasets that include both AD and FTD data for cross-validation purposes. Addressing these limitations in future research could open new avenues for refining and expanding our framework.

4 Conclusion

In this research work, we have developed a DL-based CNN model to perform classification of AD and FTD from HC subjects using EEG signal data. We have tested different segment lengths to check the impact of the frame length on the classification process. We have evaluated the framework using both 10-fold CV and LOOV.

For both AD vs. HC and FTD vs. HC classifications, we have achieved the best accuracy in 10-fold CV using the 5 s segment length, which are 97.08% and 98.14%, respectively. In case of LOOV, 15 s segment length has produced the best accuracy of 96.90% and 94.50%, accordingly. In both classification tasks, the proposed method has outperformed the existing studies in terms of classification performance.

In summation, the results of our research paint a compelling picture. This innovative approach is not only adept at classifying AD and FTD but holds promise for a broader spectrum of neurological disorders. Beyond classification, our method's adaptability suggests its potential application in diverse signal processing assignments, marking a significant stride forward in the intersection of neurology and artificial intelligence.

This study, while demonstrating promising results in Alzheimer's disease (AD) detection using a deep learning-based EEG classification framework, has some limitations. The use of a single publicly available dataset from AHEPA General University Hospital may limit the generalizability of the findings across diverse populations and clinical environments. To address these limitations and advance the field, future research should focus on expanding the dataset to include more diverse and real-world data and exploring alternative deep learning architectures beyond CNNs. Also, we can explore integrating data-driven heuristics or optimization strategies to automatically select segment lengths that maximize

classification performance while minimizing training cost.

Author contributions

Md. Nurul Ahad Tawhid contributed to the conceptualization, data curation, formal analysis, investigation, methodology, visualization, and writing of the original draft. Siuly Siuly collaborated on the formal analysis, visualization, and writing of both the original draft and the review and editing process. Enamul Kabir and Yan Li were instrumental in project administration and supervision, with Yan Li also providing resources. All authors, including Md. Nurul Ahad Tawhid, Siuly Siuly, Enamul Kabir, and Yan Li, contributed to the review and editing of the manuscript, ensuring its quality and coherence.

Funding

None.

Availability of data and materials

Here, we have used the publicly available EEG dataset from this link: <https://openneuro.org/datasets/ds004504/versions/1.0.5>.

Declarations

Competing interests

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.

Author details

¹Institute of Information Technology, University of Dhaka, Nilkhet Road, 1000 Dhaka, Bangladesh. ²Centre for Health Research, University of Southern Queensland, 487-535 West St, Darling Heights, Toowoomba, QLD 4350, Australia. ³Institute for Sustainable Industries & Liveable Cities, Victoria University, Ballarat Road, Footscray, Melbourne, VIC 3011, Australia. ⁴School of Mathematics, Physics and Computing, University of Southern Queensland, 487-535 West St, Darling Heights, Toowoomba, QLD 4350, Australia.

Received: 8 February 2025 Accepted: 30 April 2025

Published online: 04 June 2025

References

- Siuly S, Zhang Y (2016) Medical big data: neurological diseases diagnosis through medical data analysis. *Data Sci Eng* 1(2):54–64
- Rodrigues PM, Bispo BC, Garrett C et al (2021) Lacsogram: A new eeg tool to diagnose alzheimer's disease. *IEEE J Biomed Health Inf* 25(9):3384–3395
- Fernández M, Gobartt AL, Balañá M (2010) Behavioural symptoms in patients with alzheimer's disease and their association with cognitive impairment. *BMC Neurol* 10(1):1–9
- WHO (2023) Dementia. <https://www.who.int/news-room/fact-sheets/detail/dementia>, (Date last accessed October 2023)
- Miltiadous A, Gionanidis E, Tzimourta KD, et al (2023) Dice-net: A novel convolution-transformer architecture for alzheimer detection in eeg signals. *IEEE Access*
- Siuly S, Alçin ÖF, Kabir E et al (2020) A new framework for automatic detection of patients with mild cognitive impairment using resting-state eeg signals. *IEEE Trans Neural Syst Rehabil Eng* 28(9):1966–1976
- You Z, Zeng R, Lan X et al (2020) Alzheimer's disease classification with a cascade neural network. *Front Pub Health* 8:584387
- Bracco L, Gallato R, Grigoletto F et al (1994) Factors affecting course and survival in alzheimer's disease: a 9-year longitudinal study. *Arch Neurol* 51(12):1213–1219
- Gauthier S (2006) *Clin Diagn Manag Alzheimer's Dis*. CRC Press
- Feldman H, Woodward M (2005) The staging and assessment of moderate to severe alzheimer disease. *Neurology* 65(6 suppl 3):S10–S17
- Yang S, Bornot JMS, Wong-Lin K et al (2019) M/eeg-based bio-markers to predict the mci and alzheimer's disease: a review from the ml perspective. *IEEE Trans Biomed Eng* 66(10):2924–2935
- Van der Hiele K, Vein A, Van Der Welle A et al (2007) Eeg and mri correlates of mild cognitive impairment and alzheimer's disease. *Neurobiol Aging* 28(9):1322–1329
- Polikar R, Tilley C, Hillis B et al (2010) Multimodal eeg, mri and pet data fusion for alzheimer's disease diagnosis. 2010 Annual international conference of the IEEE engineering in medicine and biology. IEEE, New York, USA; 6058–6061
- Tawhid MNA, Siuly S, Wang K et al (2022a) Brain data mining framework involving entropy topography and deep learning. In: *Australasian Data-base Conference*, Springer, pp 161–168
- Tawhid MNA, Siuly S, Wang H (2020) Diagnosis of autism spectrum disorder from eeg using a time-frequency spectrogram image-based approach. *Electronics Letters*
- Ahmadi N, Pei Y, Carrette E et al (2020) Eeg-based classification of epilepsy and pnes: Eeg microstate and functional brain network features. *Brain Inf* 7:1–22
- Greiner G, Zhang Y (2024) Multi-modal eeg neo-ffi with trained attention layer (mental) for mental disorder prediction. *Brain Inf* 11(1):26
- Şengür A, Guo Y, Akbulut Y (2016) Time-frequency texture descriptors of eeg signals for efficient detection of epileptic seizure. *Brain Inf* 3:101–108
- Tawhid MNA, Siuly S, Wang H et al (2021) A spectrogram image based intelligent technique for automatic detection of autism spectrum disorder from eeg. *Plos one* 16(6):e0253094
- Tawhid MNA, Siuly S, Li T (2022) A convolutional long short-term memory based neural network for epilepsy detection from eeg. *IEEE Trans Instrum Meas* 71:1–11. <https://doi.org/10.1109/TIM.2022.3217515>
- Tsolaki A, Kazis D, Kompatsiaris I et al (2014) Electroencephalogram and alzheimer's disease: clinical and research approaches. *Int J Alzheimer's Dis* 1:349249
- Khoshnevis SA, Sankar R (2021) Classification of the stages of parkinson's disease using novel higher-order statistical features of eeg signals. *Neural Comput Appl* 33:7615–7627
- Tawhid MNA, Siuly S, Wang K et al (2022) Textural feature based intelligent approach for neurological abnormality detection from brain signal data. *Plos one* 17(11):e0277555
- Tawhid MNA, Siuly S, Wang K et al (2023) Automatic and efficient framework for identifying multiple neurological disorders from eeg signals. *IEEE Trans Technol Soc* 4(1):76–86. <https://doi.org/10.1109/TTS.2023.3239526>
- Yao Z, Hu B, Xie Y et al (2015) A review of structural and functional brain networks: small world and atlas. *Brain Inf* 2:45–52
- Singh S, Jadli H, Padma Priya R et al (2024) Kdtl: knowledge-distilled transfer learning framework for diagnosing mental disorders using eeg spectrograms. *Neural Comput Appl* 36:1–16
- Abásolo D, Hornero R, Espino P et al (2006) Entropy analysis of the eeg background activity in alzheimer's disease patients. *Physiol Meas* 27(3):241
- Escudero J, Abásolo D, Hornero R et al (2006) Analysis of electroencephalograms in alzheimer's disease patients with multiscale entropy. *Physiol Meas* 27(11):1091
- Simons S, Espino P, Abásolo D (2018) Fuzzy entropy analysis of the electroencephalogram in patients with alzheimer's disease: is the method superior to sample entropy? *Entropy* 20(1):21
- Puri D, Nalbalwar S, Nandgaonkar A et al (2022) Alzheimer's disease detection from optimal electroencephalogram channels and tunable q-wavelet transform. *Indo J Elec Engg Comp Sci* 25(3):1420–1428
- Puri D, Nalbalwar S, Nandgaonkar A, et al (2022a) Alzheimer's disease detection using empirical mode decomposition and hjorth parameters of eeg signal. In: 2022 International Conference on Decision Aid Sciences and Applications (DASA), IEEE, pp 23–28
- Puri D, Nalbalwar S, Nandgaonkar A, et al (2022c) Eeg-based diagnosis of alzheimer's disease using kolmogorov complexity. In: *Applied Information Processing Systems: Proceedings of ICCET 2021*, Springer, pp 157–165
- Neto E, Biessmann F, Aurlen H et al (2016) Regularized linear discriminant analysis of eeg features in dementia patients. *Front Aging Neurosci* 8:273
- Hajamohideen F, Shaffi N, Mahmud M et al (2023) Four-way classification of alzheimer's disease using deep siamese convolutional neural network with triplet-loss function. *Brain Inf* 10(1):5

35. Noor MBT, Zenia NZ, Kaiser MS et al (2020) Application of deep learning in detecting neurological disorders from magnetic resonance images: a survey on the detection of alzheimer's disease, parkinson's disease and schizophrenia. *Brain Inf* 7:1–21
36. Oh SL, Hagiwara Y, Raghavendra U et al (2020) A deep learning approach for parkinson's disease diagnosis from eeg signals. *Neural Comput Appl* 32(15):10927–10933
37. Tawhid MNA, Siuly S, Wang K et al (2024) Genet: A generic neural network for detecting various neurological disorders from eeg. *IEEE Trans Cognit Dev Syst*. <https://doi.org/10.1109/TCDS.2024.3386364>
38. Chen Y, Wang H, Zhang D et al (2023) Multi-feature fusion learning for alzheimer's disease prediction using eeg signals in resting state. *Front Neurosci* 17:2
39. Tawhid MNA, Siuly S, Kabir E et al (2024) Exploring frequency band-based biomarkers of eeg signals for mild cognitive impairment detection. *IEEE Trans Neural Syst Rehabil Eng* 32:189–199
40. Morabito FC, Campolo M, Ieracitano C, et al (2016) Deep convolutional neural networks for classification of mild cognitive impaired and alzheimer's disease patients from scalp eeg recordings. In: 2016 IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI), IEEE, pp 1–6
41. Ferri R, Babiloni C, Karami V et al (2021) Stacked autoencoders as new models for an accurate alzheimer's disease classification support using resting-state eeg and mri measurements. *Clin Neurophysiol* 132(1):232–245
42. Ieracitano C, Mammone N, Bramanti A et al (2019) A convolutional neural network approach for classification of dementia stages based on 2d-spectral representation of eeg recordings. *Neurocomputing* 323:96–107
43. Alessandrini M, Biagetti G, Crippa P et al (2022) Eeg-based alzheimer's disease recognition using robust-pca and lstm recurrent neural network. *Sensors* 22(10):3696
44. Tzamourta KD, Giannakeas N, Tzallas AT et al (2019) Eeg window length evaluation for the detection of alzheimer's disease over different brain regions. *Brain Sci* 9(4):81
45. Anders P, Müller H, Skjæret-Maroni N et al (2020) The influence of motor tasks and cut-off parameter selection on artifact subspace reconstruction in eeg recordings. *Med Biol Eng Comput* 58:2673–2683
46. Plechawska-Wójcik M, Augustynowicz P, Kaczorowska M et al (2023) The influence assessment of artifact subspace reconstruction on the eeg signal characteristics. *Appl Sci* 13(3):1605
47. Delorme A, Makeig S (2004) Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *J Neurosci Methods* 134(1):9–21
48. Rivera MJ, Teruel MA, Maté A, et al (2021) Diagnosis and prognosis of mental disorders by means of eeg and deep learning: a systematic mapping study. *Artificial Intelligence Review* pp 1–43
49. Aslan Z, Akin M (2020) Automatic detection of schizophrenia by applying deep learning over spectrogram images of eeg signals. *Traitement du Signal* 37(2):235–244
50. Tawhid MNA, Siuly S, Wang K et al (2021a) Data mining based artificial intelligent technique for identifying abnormalities from brain signal data. In: International Conference on Web Information Systems Engineering, Springer, pp 198–206
51. Shin HC, Roth HR, Gao M et al (2016) Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. *IEEE Trans Med Imag* 35(5):1285–1298
52. Goodfellow I, Bengio Y, Courville A (2016) Convolutional networks. *Deep learning*, vol 2016. MIT press Cambridge, MA, USA, pp 330–372
53. Siuly S, Yin X, Hadjiloucas S et al (2016) Classification of thz pulse signals using two-dimensional cross-correlation feature extraction and non-linear classifiers. *Comput Methods Progr Biomed* 127:64–82
54. Siuly S, Khare SK, Bajaj V et al (2020) A computerized method for automatic detection of schizophrenia using eeg signals. *IEEE Trans Neural Syst Rehabil Eng* 28(11):2390–2400
55. Van der Maaten L, Hinton G (2008) Visualizing data using t-sne. *J Mach Learn Res* 9(11):1

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.