Biomedical Signal Processing and Control 89 (2024) 105751

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

A novel hybrid model in the diagnosis and classification of Alzheimer's disease using EEG signals: Deep ensemble learning (DEL) approach

Majid Nour [a], Umit Senturk [b], Kemal Polat [c,*]

[a] Department of Electrical and Computer Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia
[b] Department of Computer Engineering, Bolu Abant Izzet Baysal University, Bolu, Turkey
[c] Department of Electrical and Electronics Engineering, Bolu Abant Izzet Baysal University, Bolu, Turkey

ARTICLE INFO

Keywords:
Alzheimer's disease classification
Deep ensemble learning
EEG
Dementia
Explainability

ABSTRACT

Recent years have witnessed a surge of sophisticated computer-aided diagnosis techniques involving Artificial Intelligence (AI) to accurately diagnose and classify Alzheimer's disease (AD) and other forms of Dementia. Despite these advancements, there is still a lack of reliable and accurate methods for distinguishing between (AD) and Healthy Controls (HC) using Electroencephalography signals (EEG). The main challenge is finding the right features from the intricate spectral-temporal EEG data, which can provide information sufficient for diagnosis. This study proposes a new approach integrating Deep Ensemble Learning (DEL) and 2-dimensional Convolutional Neural Networks (2D-CNN) to address these issues. Combining state-of-the-art supervised deep learning algorithms within an ensemble model architecture aims to accurately diagnose and classify EEG signals of AD and HC subjects. Public EEG-based Alzheimer's datasets have been classified in the DEL model without applying any feature extraction after cleaning from noise and artifacts.

Furthermore, the proposed DEL model used 5 different 2D-CNN models as internal classifiers. As a result, the EEG-based DEL model proposed for the first time provided high accuracy in AD classification. The proposed DEL model reached an average accuracy of 97.9% in AD classification due to 5 cross-fold training.

In conclusion, this work renders that incorporating ensemble learning techniques into automotive health applications create extensible and stable AI models needed for computer-aided diagnostic. However, although the reported results and evaluation are promising, further efforts will need to be made to improve the accuracy of our proposed model. In addition, a fine-grid evaluation will be necessary to accurately understand potential impacts in clinical applications, such as earlier diagnosis or treatment decisions.

1. Introduction

Alzheimer's Disease (AD) is a progressive degenerative disorder that seriously threatens aging populations worldwide. AD is a devastating disorder that slowly deteriorates the human mind, impacting memory, attention, and behavior. It is estimated by the World Health Organization that globally, over 55 million people live with Dementia, and every year, there are nearly 10 million new cases. Especially concerning, more than 60% of people with Dementia reside in low- and middle-income countries [1]. AD is an ongoing and progressive brain disorder that slowly destroys memory and thinking ability and affects the ability to carry out even the simplest tasks. It is the most common cause of Dementia among seniors and is a disease without a known cure. AD is the most common cause of Dementia and may contribute to 60–70% of cases, an umbrella term for a set of cognitive and behavioral symptoms that develop with certain brain diseases [2]. The progressive disorder slowly destroys brain cells, leading to memory and intellectual decline, dysfunction with activities of daily living, behavior changes, and delusions. Early symptoms usually manifest in language, spatial abilities, and memory problems. Over time, linguistic and cognitive impairments may cause a person to struggle with daily tasks, such as completing basic chores and social activities.

AD is divided into three stages: Early, Mild, and Severe [3]. Early-stage AD, sometimes called mild cognitive impairment (MCI), is the beginning of the noticeable cognitive decline, but symptoms are not disabling, and individuals can still live independently. During this stage, individuals may struggle with organizing tasks, multi-tasking, memory, and language. Mild-stage AD is the most common stage and includes

* Corresponding author at: Bolu Abant Izzet Baysal University, Faculty of Engineering, Department of Electrical and Electronics Engineering, Bolu, Turkey.
E-mail addresses: mnour@kau.edu.sa (M. Nour), umit.senturk@ibu.edu.tr (U. Senturk), kpolat@ibu.edu.tr (K. Polat).

https://doi.org/10.1016/j.bspc.2023.105751
Received 1 June 2023; Received in revised form 9 October 2023; Accepted 13 November 2023
Available online 17 November 2023

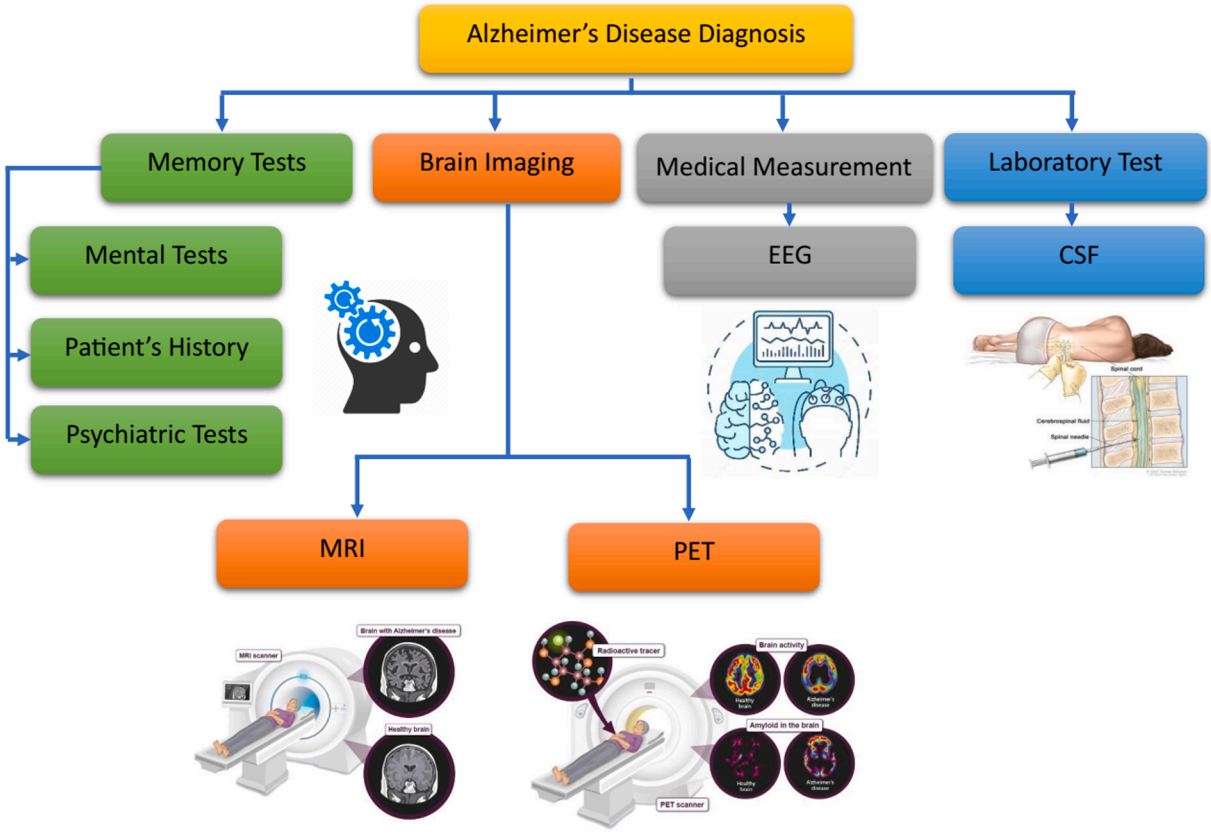

Fig. 1. AD diagnosis methods [16–18].

symptoms such as impaired judgment, difficulty concentrating, and difficulty with everyday tasks. At this stage, individuals may be more dependent on others for help and may need assistance with activities of daily living, such as bathing and dressing. Severe-stage AD includes confusion, agitation, delusions, and incontinence [4]. During this stage, individuals depend entirely on others for their basic needs.

The first sign of AD is usually a gradual decline in cognition or mental functioning [5]. People may notice that they are having difficulty concentrating, making decisions, and recalling events or conversations. They may also find it difficult to follow directions, remember how to do familiar things, or remember details about their daily lives. Memory loss is one of the most common early symptoms, and people often find themselves forgetting dates, appointments, or conversations that occurred a few minutes earlier. In addition to cognitive and memory problems, the person may experience behavioral or mood changes [6]. They might struggle to deal with difficult situations or decisions or become socially withdrawn. Mood swings, irritability, and even depression are common signs of AD [7]. As time progresses, individuals with AD will have progressively more difficulty with activities of daily living (ADLs), like dressing, bathing, or making meals [8]. They may also experience sleep disturbances, confusion, and struggle recognizing familiar faces or places. In addition, individuals with AD may have difficulty with routine tasks, like paying bills and keeping track of medications. AD is unfortunate for many individuals and their families [9]. The good news is that understanding its early symptoms can help people get diagnosed and appropriate treatment. With early intervention, it is possible to minimize the effects of AD and prolong a person's quality of life.

Although there is no definite way to diagnose AD, several tests can aid in diagnosing the disorder. Memory tests, such as the Mini-Mental State Examination, can assess an individual's recollection of recent events and ability to recall past ones [10]. Brain scans, such as a magnetic resonance imaging scan (MRI) or a positron emission tomography (PET) scan, can be used to detect abnormalities in the brain commonly associated with AD [11,12]. Additionally, a lumbar puncture or spinal tap can determine the level of proteins in the cerebrospinal fluid (CSF) [13]. It is recommended that individuals exhibiting signs of AD receive a multidisciplinary evaluation to ensure an accurate and timely diagnosis [14]. Such an evaluation is usually conducted by a team of healthcare professionals and may include a neurologist, a psychologist, a behavioral neurologist, and a geriatric primary care physician. Psychiatric tests, such as the Geriatric Depression Scale, can also assess a patient's mental status [15]. In addition, the patient's caregivers can provide valuable information about the individual's behavior, which can help make a diagnosis. Fig. 1 shows AD diagnosis methods.

Electroencephalography signals (EEG) are a method of detecting electrical activity in the brain. It diagnoses various neurological and medical conditions, including AD [19]. During the early stages of AD, EEGs can be used to detect changes in brain activity related to the disease. These changes can include increased slow brainwaves (delta and theta waves) [20] and decreased faster waves (alpha and beta waves) [21,22]. EEGs can detect these changes even before the onset of physical symptoms. EEGs provide a valuable tool to diagnose and monitor AD. Furthermore, EEGs can be used to monitor changes in brain activity associated with the progression of the disease and help doctors adjust treatments and make better decisions. The advantages of diagnosing AD with EEG are that it provides a unique real-time insight into the activity of the brain [23], offers information that is unobtainable through other means, can detect subtle changes in brain waves due to AD much earlier than traditional diagnostic methods, does not involve radiation or other potentially hazardous procedures, increases the diagnostic accuracy of biomarkers for detecting AD, can detect abnormal activity and measure the progress of an individual's AD over some time and, is highly cost-effective and accessible, making it a potentially helpful tool in the early diagnosis and management of AD [24].

EEG signals have been explored as a potential biomarker for AD

diagnosis and classification [25,26]. Machine learning and deep learning methods can be used to classify EEG signals according to the presence or absence of AD [27]. Machine learning and deep learning approaches begin by extracting relevant features from the EEG signal. These features could include the frequency spectrum of different bands, moments (such as mean, variance, or skewness), entropies, and other measures that provide helpful information about the EEG signal [28]. After features are extracted, machine learning methods can then be used to classify signal segments into AD vs. Healthy Controls (HC)—for example, support vector machines (SVM). The advantages of SVM include its ability to minimize data overfitting, its robustness in large datasets, and its ability to accommodate multiple classes. The disadvantages of SVM include its high computational cost and potential difficulty in tuning parameters [28,29]. Random forests (RF) The advantages of RF include its ability to handle various data types, its flexibility in tuning model parameters, and its ability to predict probabilities accurately. The main disadvantages of RF include its time-consuming nature and lack of interpretability [30–32]. Deep learning models as a one-dimensional convolutional neural network 1D-CNN [33], convolutional neural network (CNN) [34–36]. The advantages of CNNs include their ability to handle large amounts of data, their accuracy in predicting complex patterns, and their power in processing visual data. The disadvantages of CNNs include the computational cost associated with training them, the difficulty verifying their predictions, and reliance on labeled data. Recurrent Neural Networks (RNN) [37,38] and long short-term memory (LSTM) [39,40,41] have been used to classify vital signals, etc. EEG signals. The advantages of RNNs and LSTM include their ability to capture temporal patterns, process variable length data, and extend to higher-level neural network architectures. The disadvantages of RNNs include their high computational cost, long training times, and difficulty using larger datasets. Each method has strengths and weaknesses, and choosing the correct method will depend on the specific problem and desired accuracy.

1.1. Literature gaps

The gaps obtained based on the above literature studies are given below. The papers examined are models frequently used in the detection of AD. We can address the following gaps in the paper:

- Lack of studies on the relationship between EEG signals and Alzheimer's disease: The paper investigated how specific EEG patterns or features are related to Alzheimer's disease and the combination of different datasets [42,43]. This can provide physicians with a decision support system in diagnosing AD.
- Effectiveness of different EEG signal standards for AD diagnosis: The paper compared and evaluated different EEG signal processing techniques or standards to determine their effectiveness in accurately diagnosing AD [28,44]. This may help determine the most reliable and practical approach to using EEG signals as a diagnostic tool for AD.
- Lack of studies using hybrid models [41,45]. The authors emphasized the scarcity of studies that combine multiple EEG frequency bands to diagnose and classify Alzheimer's Disease.
- Limited use of deep ensemble learning in EEG-based AD classification: The paper discusses the limited use of deep ensemble learning approaches in EEG-based AD classification. It is explained how their new model fills this gap by using a deep ensemble learning approach to improve classification performance. When the literature was examined, the use of the deep ensemble learning (DEL) model in AD diagnosis and classification using EEG signals was not found.
- Lack of generalization to different data sets [46,47]: The paper emphasizes the difficulty of generalizing the results from one data set to different data sets. Potential limitations of their study in terms of dataset specificity are discussed.
- Insufficient investigation of feature importance [48,49]: The paper examined the feature importance analysis of different frequency bands in AD diagnosis in more depth. Each frequency band's potential biological significance and relevance to AD pathology is discussed.

The paper can contribute to the existing literature by addressing these gaps, better understanding the relationship between EEG signals and AD, and improving diagnostic and cognitive assessment methods for AD using EEG.

1.2. Our motivation for this study

The medical community faces the daunting challenge of accurately diagnosing and classifying diseases such as AD. Early detection and diagnosis of AD are crucial for the appropriate prognosis and management of the condition. Thus, reliable techniques for diagnosing and classifying AD need to be developed. Addressing this challenge, this study proposes to utilize DEL approaches to accurately diagnose and classify AD using electroencephalography (EEG) signals. Recent years have seen a rise in the application of machine learning techniques to medical diagnosis and classification tasks. Deep learning techniques, such as deep neural networks (DNNs), have been successfully utilized for many problems related to medical image analysis [50,51]. Drawing inspiration from this, our proposed approach uses DNNs to classify EEG signals of AD patients and a separate group of HC patients. By training multiple DNNs and combining their outputs through a dedicated ensemble system, a higher accuracy of the diagnosis and classification process can be achieved. Initially, the EEG signals of multiple subjects will be pre-processed and filtered to extract the vital features. Afterward, multiple DNNs are trained on these features for diagnosis and classification. Multiple DNNs are trained such that the input is multiple EEG features, and the output is a specific diagnosis (e.g., AD vs. HC). Combining multiple DNNs using an ensemble system can achieve superior performance compared to any single DNN [52,53]. Specifically, traditional ensemble techniques such as bagging or boosting can combine the outputs of multiple DNNs for better accuracy [54,55]. By incorporating the ensemble technique, our proposed approach can perform better than any single DNN-based system. Additionally, ensemble techniques are robust to overfitting and can filter out uncertain instances, making them suitable for medical diagnostic tasks. Thus, this DEP system could accurately diagnose and classify AD using EEG signals.

1.3. The main contributions

- We propose a novel hybrid model for diagnosing and classifying AD using EEG signals. The model combines deep learning techniques with ensemble learning to enhance the accuracy and robustness of the classification task. With this method, 98% classification accuracy was achieved.
- Two different data sets were used to evaluate the performance of the hybrid model proposed in the study. Two data sets consisting of EEG recordings taken from AD patients and HCs contain different numbers of AD and HC patients. In the literature, no classification study is made by combining two different datasets.
- The hybrid model proposed in the study was used in AD-HC classification with EEG signals at different frequencies (delta, theta, alpha, beta, and gamma). Thanks to this study, it will be possible to understand which EEG frequencies AD causes changes.

2. Material and method

This article used a public EEG-based AD diagnostic, two different datasets created by researchers at Florida State University. The dataset was passed through a Butterworth band-pass filter to remove noise and

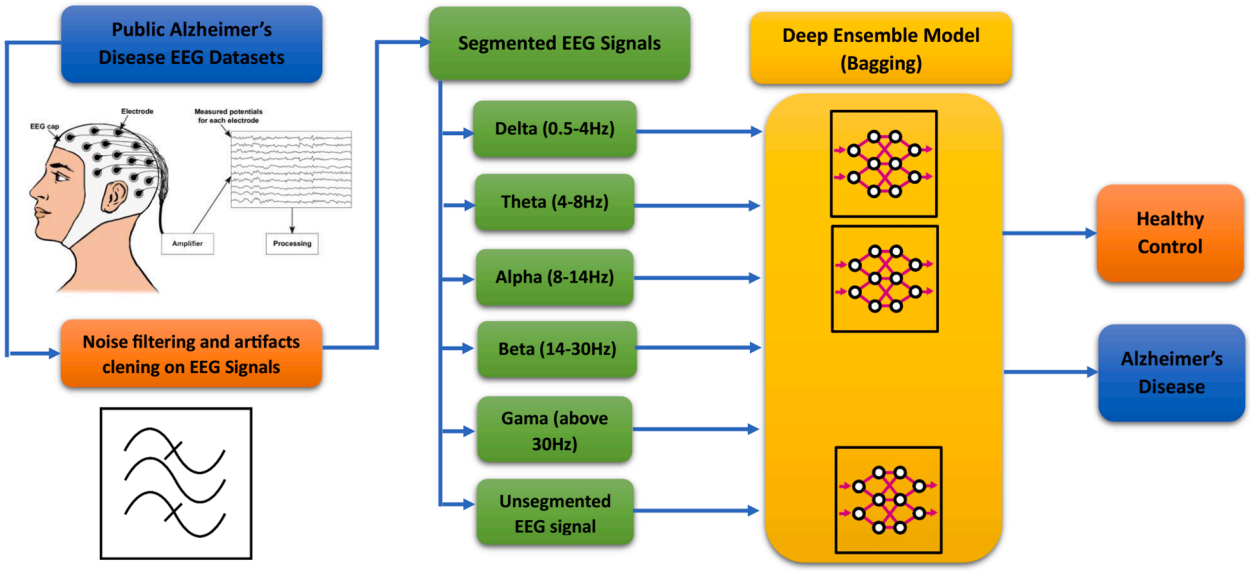

Fig. 2. Block diagram AD diagnosis DEL model system [56].

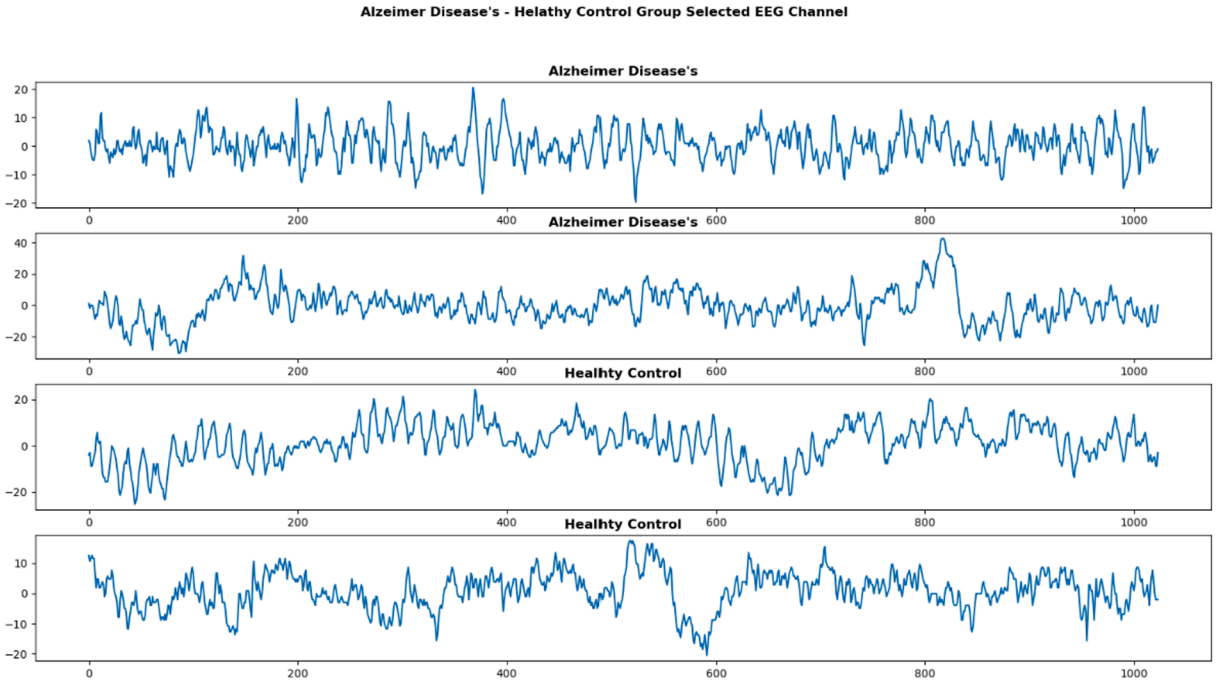

Fig. 3. Alzheimer Disease's (AD) – Healthy Control (HC) Group EEG signals.

artifacts. The AD dataset is separated into 5 frequency bands by filtering with the help of band-pass filters. It has been tried to measure which frequency band is more effective in AD diagnosis. EEG signals in each frequency band are then divided into epochs. EEG signal epochs in five frequency bands are given as input to the DEL model. In the DEL model target, labels are given according to the AD- HC classes in the AD dataset. EEG signals are given to the input of the DEL model in a raw form without using any feature extraction technique. The internal classifier models were randomly generated from the 2D-CNN models in the DEL model. In the DEL model, 5 internal classifier models are used. After the internal classifier models were trained separately, the accuracy of each model was determined as the confidence weight. The metamodel of the DEL model classifies by averaging the confidence weights of the internal classifier models. As a result, the DEL model performance will give the average performance of the internal classifier models. Fig. 2 shows the DEL model used in AD classification.

2.1. Dataset

Dr. Dennis Duke of Florida State University provided this study's two different EEG databases, which are freely accessible via [57,58].

2.1.1. Dataset-A

The dataset by Dr. Dennis Duke consists of 8-second EEG segments, sampled at 128 Hz and band-limited to 1–30 Hz, that have been filtered to remove eye motion, blinking, and myogenic artifacts. The International 10–20 system was used to make recordings of four groups (A, B, C, and D) from 19 scalp loci (Fp1, Fp2, Fz, F3, F4, F7, F8, Cz, C3, C4, T3, T4,

Pz, P3, P4, T5, T6, O1 and O2), of the brain using a Biologic Systems Brain Atlas III Plus workstation. These recordings, made under two conditions (eyes closed and eyes open), were of the frontal, central, parietal, occipital, and temporal lobes (F, C, P, O, and T) and reference the linked mandible on the forehead. Researchers from Florida State University jointly designed this database. The subjects for the database were 24 HC elderly individuals aged 61–83 (average age 72), who were all screened and found to be free of any neurological or psychiatric disorders (Groups A and B). Also included were 24 probable AD patients, aged 53–85 (average age 69), who were diagnosed using NINCDS, ADRDA, and DSM-III-R criteria (Groups C and D). A technician was present during the EEG recordings of all subjects to ensure they remained alert. Additional information about the database can be found in [59,60]. Fig. 3 shows AD – HC Group dataset EEG signals.

As seen in Fig. 3, the 8 s EEG signal generates 1024 data points with a sampling frequency 128. Since EEG signals are complex, they are difficult to distinguish from the human eye.

2.1.2. Dataset-B

Dataset-B by Dr. Dennis Duke, employed in this study, was generously provided by researchers affiliated with Florida State University [58]. It consists of EEG signals divided into four distinct groups: (A) a cohort of 12 healthy elderly subjects who maintained visual fixation with their eyes open, (B) a cohort of 12 healthy elderly subjects with their eyes closed, (C) a group of 80 probable AD patients who visually fixated with their eyes open, and (D) a group of 80 probable AD patients with their eyes closed. The inclusion criteria for the 160 probable AD patients were diagnosed through the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and the Alzheimer's Disease and Related Disorders Association (ADRDA), as well as Diagnostic and Statistical Manual of Mental Disorders (DSM)-III-R criteria. Notably, the patients were not stratified based on disease severity due to the unavailability of relevant information. In this paper, groups A-B and C-D were EEG recordings from the same patients and were recorded with their eyes open and closed. Blindfolded EEG recordings were used in this article to ensure synchronization with the initial dataset. To ensure similarity between these two datasets, EEG records of group B and D patients were selected.

The EEG recordings were obtained using a sampling frequency of 128 Hz and a duration of 8 s per segment. The signals were acquired from 19 scalp electrodes positioned according to the international 10–20 system (Fp1, Fp2, F3, F4, F7, F8, Fz, C3, C4, Cz, P3, P4, Pz, T3, T4, T5, T6, O1, O2). Preprocessing of the signals involved band-limiting within the range of 0.5–30 Hz, and removal of artifacts resulting from subject movements was performed by an experienced EEG technician. Dr. Dennis Duke facilitated the provision of this freely available database [58]. Additional information about the database can be found in [61].

This study combined the two datasets described above and used in AD classification. There are 24 AD and 24 HC patients in the Dataset-A, and 80 AD and 12 HC patients in the Dataset-B. In total, EEG signals of 104 CE patients and 36 HC patients have been used in this study.

2.2. Pre-processing (Filtering)

The Butterworth band-pass filter was used for pre-processing this work. The Butterworth band-pass filter is a linear filter often used in signal-processing applications. It is designed to provide a maximally flat ("Butterworth") response in the pass band, with a steep transition to a stop band. This filter is commonly used in EEG signal processing applications because it allows only frequencies within the defined range to pass while rejecting most unwanted frequencies outside the pass band [62,63]. Advantages: This filter is designed to provide a flat response in the pass band, making it suitable for applications that require frequency-dependent gain within the pass band, like EEG signal processing. Easy to design and implement. Stable over a wide range of frequencies. Limitations: Rejection is not linear, resulting in possibly poor stopband performance. Those seeking narrow roll-off slopes (high-order filters) may need to consider other alternatives. It can be susceptible to breaking down at higher frequencies. This filter also has higher ripple values in both the passband and the stopband than other filters, which can result in slight variations in the frequency response across different frequencies. This filter is used in EEG signal processing applications because it allows only frequencies within the defined range to pass while rejecting most unwanted frequencies outside the pass band. This helps to reduce noise and artifacts from the signal, making the signal more reliable for EEG analysis. The Butterworth band-pass filter chose quality factor (q) = 6 [64]. In this way, signals other than cutoff frequencies are suppressed. Motion artifacts are prevented for individuals with a lower cutoff frequency of 0.5 Hz. The upper cutoff frequency is 45 Hz. In this way, high-frequency noises and city electricity grid noises are suppressed. Formula 1 is a low-pass Butterworth transfer function, and Formula 2 is a high-pass Butterworth transfer function. A band-pass filter is formed when low-pass and high-pass filters are used together.

- $H(j\omega)$ = Transfer function at angular frequency ω.
- ω = Angular frequency and is equal to $2\pi f$.
- ω_c = Cutoff frequency expressed as an angular value and is equal to $2\pi fc$.
- n = Number of elements in the filter.

$$|H(j\omega)| = \frac{1}{\sqrt{1+\left(\frac{\omega}{\omega_c}\right)^{2n}}} \tag{1}$$

$$|H(j\omega)| = \frac{1}{\sqrt{1+\left(\frac{\omega}{\omega_c}\right)^{-2n}}} \tag{2}$$

2.3. Segmentation of EEG signals into frequency bands

EEG signals are usually divided into five frequency bands: Delta, Theta, Alpha, Beta, and Gamma [65]. The Delta band is associated with the deepest level of sleep and coma. It is typically below 4 Hz. The Theta band is typically 4–8 Hz and is associated with teenagers and adults in the first stage of sleeping, daydreaming, and relaxation. The alpha band is typically 8–14 Hz and is associated with light sleep, relaxation in adults, and deep concentration in teenagers. The Beta band is typically 14–30 Hz, associated with alert states and higher cognitive functioning. Finally, the Gamma band is typically 30 + Hz and is associated with the highest level of conscious awareness.

Certain EEG bands may be particularly useful for diagnosing and classifying AD. Specifically, the Theta and Alpha bands are related to cognitive functioning, particularly affected in AD. Decreased Theta activity has been reported for mild and moderate stages of AD [66], whereas Alpha activity may be decreased during mild AD, and further decreases are observed with increasing severity. In addition, an increased Theta/Alpha ratio has also been reported for subjects with mild AD compared to HC subjects [67]. EEG signals may also reveal abnormal neural activity associated with AD. Specifically, some studies have reported increased Delta activity in mild to moderate stages of the disease [68]. This suggests that this EEG band can be used as an early indicator of cognitive impairment due to AD. In addition, increased high-frequency activity in the Beta and Gamma bands has been reported in AD patients compared to HC subjects. This suggests that these EEG bands can give valuable insight into the nature of the neural activity affected by AD. Finally, various EEG measures, such as spectral power, have been used to differentiate AD subjects from HC subjects. In particular, several studies have found increased Delta/Theta power in AD subjects compared to HC subjects [69]. This suggests that EEG spectral power measures may be valuable for AD diagnosis and classification.

Table 1
The Deep Ensemble Model (DEL) and combined internal 2D-CNN models include neural network layers.

2D-CNN Model 1	2D-CNN Model 2
1. Conv2D(32, 3,padding="same", activation="relu"), 2. MaxPooling2D(), 3. Conv2D(64, 3,padding="same", activation="relu"), 4. MaxPooling2D(), 5. Conv2D(128, 3,padding="same", activation="relu"), 6. MaxPooling2D(), 7. Conv2D(256, 3,padding="same", activation="relu"), 8. MaxPooling2D(), 9. Flatten(), 10. Dense(512,activation = 'relu'), 11. Dense(256,activation = 'relu'), 12. Dense (128, activation="relu"), 13. Dense(64, activation="relu"), 14. Dense(32, activation="relu"), 15. Dense(16, activation="relu"), 16. Dense(2, activation="softmax") Optimizer = Adam(learning_rate = 0.0001) loss='categorical_crossentropy' epoch = 12	1. Conv2D(32, 3,padding="same", activation="relu"), 2. MaxPooling2D(), 3. Conv2D(64, 3,padding="same", activation="relu"), 4. MaxPooling2D(), 5. Conv2D(128, 3,padding="same", activation="relu"), 6. MaxPooling2D(), 7. Conv2D(256, 3,padding="same", activation="relu"), 8. MaxPooling2D(), 9. Flatten(), 10. Dense(512,activation = 'relu'), 11. Dense(256,activation = 'relu'), 12. Dense (128, activation="relu"), 13. Dense(64, activation="relu"), 14. Dense(2, activation="softmax") Optimizer = Adam(learning_rate = 0.001) loss='categorical_crossentropy' epoch = 10
2D-CNN Model 3	**2D-CNN Model 4**
1. Conv2D(32, 3,padding="same", activation="relu"), 2. MaxPooling2D(), 3. Conv2D(64, 3,padding="same", activation="relu"), 4. MaxPooling2D(), 5. Conv2D(128, 3,padding="same", activation="relu"), 6. MaxPooling2D(), 7. Flatten(), 8. Dense(512,activation = 'relu'), 9. Dense(256,activation = 'relu'), 10. Dense (128, activation="relu"), 11. Dense(64, activation="relu"), 12. Dense(32, activation="relu"), 13. Dense(16, activation="relu"), 14. Dense(2, activation="softmax") Optimizer = Adam(learning_rate = 0.001) loss='categorical_crossentropy' epoch = 8	1. Conv2D(128, 3,padding="same", activation="relu"), 2. MaxPooling2D(), 3. Conv2D(128, 3,padding="same", activation="relu"), 4. MaxPooling2D(), 5. Conv2D(128, 3,padding="same", activation="relu"), 6. MaxPooling2D(), 7. Flatten(), 8. Dense(512,activation = 'relu'), 9. Dense(256,activation = 'relu'), 10. Dense (128, activation="relu"), 11. Dense(64, activation="relu"), 12. Dense(32, activation="relu"), 13. Dense(16, activation="relu"), 14. Dense(2, activation="softmax") Optimizer = Adam(learning_rate = 0.001) loss='categorical_crossentropy' epoch = 10
2D-CNN Model 5	**DEL model**
1. Conv2D(128, 3,padding="same", activation="relu"), 2. MaxPooling2D(), 3. Conv2D(64, 3,padding="same", activation="relu"), 4. MaxPooling2D(), 5. Conv2D(32, 3,padding="same", activation="relu"), 6. MaxPooling2D(), 7. Flatten(), 8. Dense(512,activation = 'relu'), 9. Dense(256,activation = 'relu'), 10. Dense (128, activation="relu"), 11. Dense(64, activation="relu"), 12. Dense(32, activation="relu"), 13. Dense(16, activation="relu"), 14. Dense(2, activation="softmax") Optimizer = Adam(learning_rate = 0.01) loss='categorical_crossentropy' epoch = 10	$\bar{y} = \frac{\omega_1 * x_1 + \omega_2 * x_2 + \cdots + \omega_n * x_n}{\omega_1 + \omega_2 + \cdots + \omega_n}$ ω_n = *internal 2D −CNN models* classifier's accuracy scorex_n each classifier predicted class $\bar{y}$ predicted class DEL model

The different EEG bands may be used to measure the neural activity associated with AD. Specifically, decreased Theta activity and increased Delta/Theta power have been reported for mild and moderate stages of the disease. In contrast, increased Theta/Alpha ratio, Delta activity, and high-frequency power in the Beta and Gamma bands have been reported for subjects in the advanced stages of the disease. As such, these EEG bands may be used to diagnose and classify AD subjects.

This paper divides EEG signals into 5 frequency bands with a Butterworth band-pass filter. EEG signals recorded as 8 s in the datasets were segmented into 1 s epochs. An epoch was converted into a 2D array of 128x19. A total of 1120 epochs consist of two datasets combined. Then, EEG signals for each frequency band (delta, theta, alpha, beta, and gamma) were segmented into 1 s epochs (128x19 2D array) and given as input to ensemble learning.

2.4. The proposed method: Deep ensemble classification model

DEL is a supervised machine learning technique that combines the predictions of multiple deep learning models. The idea is to combine multiple models to create a more extraordinary generalization ability than a single model and reduce the risk of overfitting [70]. Ensemble learning is preferred over a single classifier model due to noise, bias features, and high variance difficulties [71]. This is achieved by combining the predictions of different models and using a final layer that takes the outputs of the individual models to provide a single output. This output is typically a classification label or a probability distribution. This technique is instrumental in computer vision, natural language processing, and medical diagnosis, where multiple signals can be used to make more accurate predictions.

DEL is a powerful technique for addressing problems in the medical

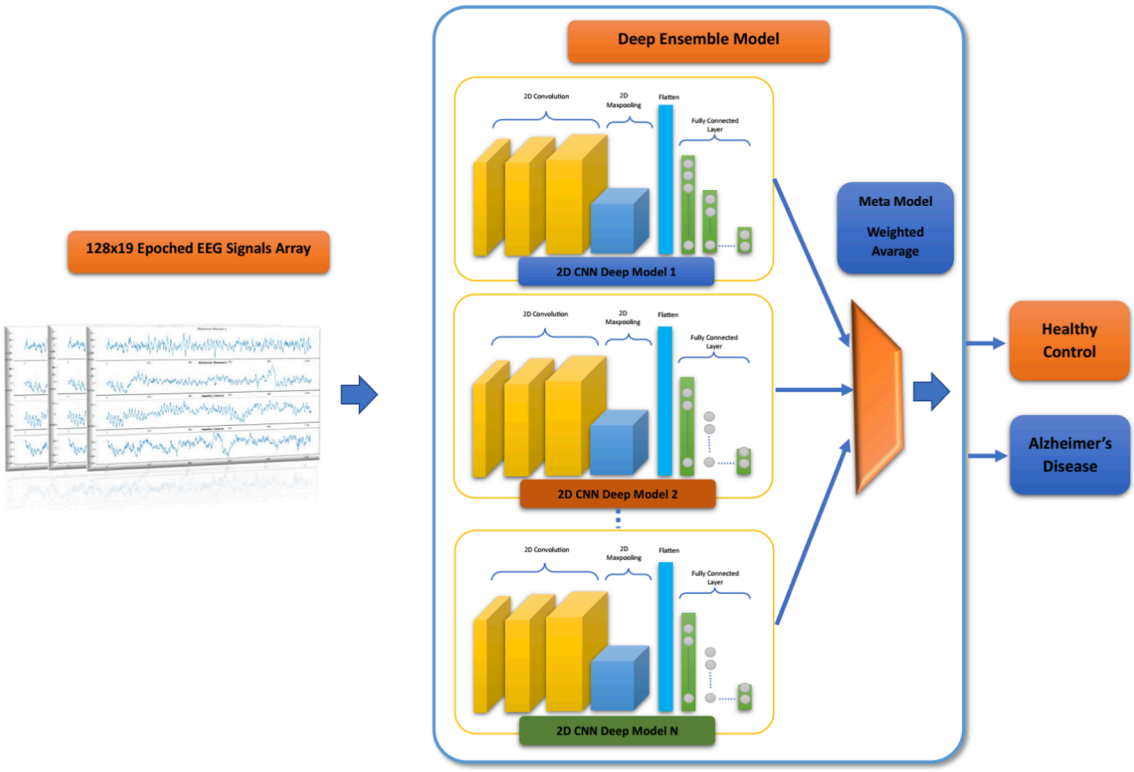

Fig. 4. The proposed Deep Ensemble Learning (DEL) Models for AD-HC Classification.

domain, including those associated with AD. This approach combines several different neural networks into a single prediction model, which can be tuned to provide better performance than any component model. DEL has been applied to AD to analyze brain scans, predict the progression of the disease, and even develop a diagnostic test that can accurately identify individuals with the condition.

The DEL model was applied to its input after the AD-HC EEG datasets was filtered (0.5–45 Hz butterworth bandpass) and segmented (one epoch = 128x19 2D array). DEL model combined internal five different 2D-CNN models. In Table 1 shown DEL model and combined internal 2D-CNN models include neural network layers. In some of the internal 2D-CNN models, the number of Conv2D layers, number of Conv2D filters, number of Dense layers, number of Dense layer neurons, learning rate and number of epochs were randomly changed. This randomness is because the classifiers in ensemble learning must be chosen randomly.

Additionally, DEL has been used to develop treatments that have proven effective against AD. Its power as a technique for addressing complex medical problems has made DEL an increasingly popular approach for medical research. Fig. 4 depicts the DEL model for AD and HC classification. For DEL model, the final layer weighted average calculation is shown in Formula 3.

$$\bar{y} = \frac{\omega_1{}^*x_1 + \omega_2{}^*x_2 + \cdots + \omega_n{}^*x_n}{\omega_1 + \omega_2 + \cdots + \omega_n} \tag{3}$$

$\bar{y}$ predicted class, ω_n DEL model each classifier's accuracy score, x_n each classifier predicted class Fig. 4. In the first, the input data to the epochs was created as a 128 × 19 two-dimensional array from the EEG signals taken from the AD data set. In the DEL model, 5 different 2D-CNN classifier models were randomly selected. Classifier 2D-CNN models are trained with EEG signals segmented in different frequency bands. The accuracy performances of the classifier models are taken as the confidence weight in the Metamodel. The metamodel makes its classification predictions by taking a weighted average. The Meta model has been freed from the overfitting problem thanks to the weighted average.

Table 2
The Deep Ensemble Model (DEL) and combined internal 2D-CNN models performance metrics for AD classification of 5 cross-fold training.

Metrics	DEL Model Mean ± std	2D-CNN Model 1 Mean ± std	2D-CNN Model 2 Mean ± std	2D-CNN Model 3 Mean ± std	2D-CNN Model 4 Mean ± std	2D-CNN Model 5 Mean ± std
Accuracy	**0,97 ± 0,01**	0,98 ± 0,00	0,98 ± 0,00	0,88 ± 0,05	0,85 ± 0,12	0,71 ± 0,02
F1 Score	**0,97 ± 0,01**	0,98 ± 0,00	0,98 ± 0,00	0,86 ± 0,05	0,73 ± 0,28	0,41 ± 0,00
Kappa Score	**0,92 ± 0,03**	0,95 ± 0,02	0,96 ± 0,02	0,69 ± 0,09	0,49 ± 0,45	0,00 ± 0,00
Jaccard Score	**0,94 ± 0,02**	0,97 ± 0,01	0,97 ± 0,01	0,85 ± 0,10	0,53 ± 0,47	0,00 ± 0,00
ROC-AUC Score	**0,98 ± 0,00**	0,98 ± 0,00	0,99 ± 0,00	0,85 ± 0,05	0,74 ± 0,33	0,35 ± 0,01
Recall	**0,96 ± 0,01**	0,97 ± 0,00	0,98 ± 0,01	0,89 ± 0,03	0,75 ± 0,23	0,50 ± 0,00
precision	**0,96 ± 0,01**	0,98 ± 0,00	0,98 ± 0,01	0,89 ± 0,03	0,75 ± 0,23	0,50 ± 0,00

3. Results

This article classified AD-HC subjects with the DEL model using the AD EEG datasets [57,58]. Dataset-A, 24 AD subjects and 24 HC subjects, and Dataset-B, 80 AD subjects and 12 HC subjects, were used in two datasets combined and used in this study. The 19-channel EEG signals were taken for 8 s and filtered with a 0.5–45 Hz Butterworth band-pass filter. The filtered EEG signals were separated into 1 s epochs (128x19 2D array), and the epochs were labeled AD-HC. In total, 1120 epochs were created. The created epochs are divided into an 80% model training set and a 20% model test set. The training dataset was applied to the DEL model, and the DEL model was trained. To increase the

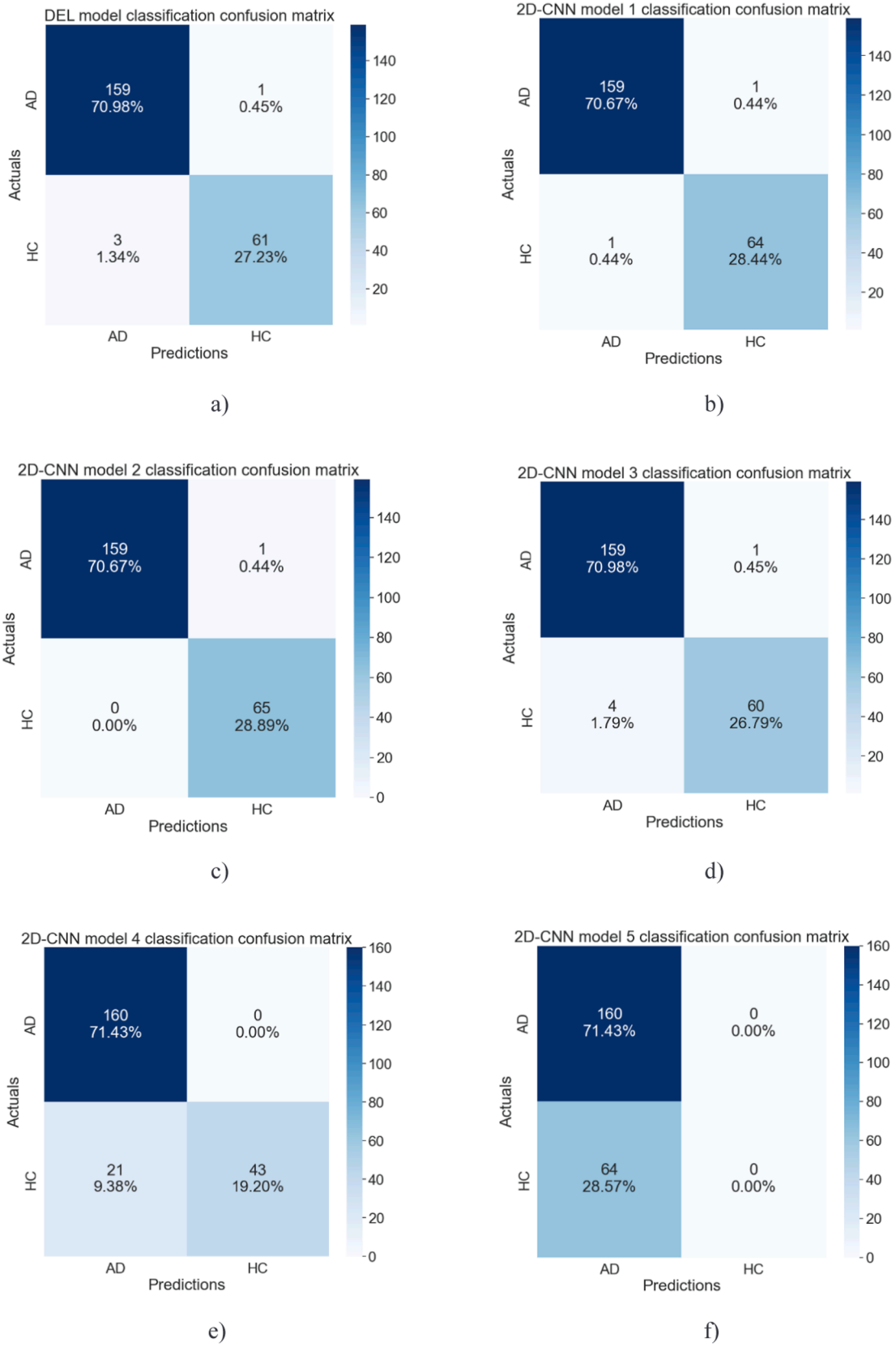

Fig. 5. Alzheimer Disease's (AD) classification confusion matrix a) DEL model, b) 2D-CNN model 1, c) 2D-CNN model 2, d) 2D-CNN model 3, e) 2D-CNN model 4, f) 2D-CNN model 5.

Table 3
The Deep Ensemble Model (DEL) combined internal 2D-CNN models performance metrics for AD datasets 0.5–45 Hz Frequency Band.

	DEL Model Mean ± std	2D-CNN Model 1 Mean ± std	2D-CNN Model 2 Mean ± std	2D-CNN Model 3 Mean ± std	2D-CNN Model 4 Mean ± std	2D-CNN Model 5 Mean ± std
Accuracy	0,97 ± 0,01	0,98 ± 0,00	0,98 ± 0,00	0,88 ± 0,05	0,85 ± 0,12	0,71 ± 0,02
F1 score	0,97 ± 0,01	0,98 ± 0,00	0,98 ± 0,00	0,86 ± 0,05	0,73 ± 0,28	0,41 ± 0,00

Table 4
The Deep Ensemble Model (DEL) combined internal 2D-CNN models performance metrics for AD datasets 0.5–4 Hz Frequency Band (Delta).

	DEL Model Mean ± std	2D-CNN Model 1 Mean ± std	2D-CNN Model 2 Mean ± std	2D-CNN Model 3 Mean ± std	2D-CNN Model 4 Mean ± std	2D-CNN Model 5 Mean ± std
Accuracy	0,92 ± 0,01	0,93 ± 0,02	0,92 ± 0,00	0,86 ± 0,00	0,83 ± 0,03	0,71 ± 0,02
F1 score	0,89 ± 0,01	0,90 ± 0,02	0,90 ± 0,00	0,84 ± 0,00	0,78 ± 0,04	0,41 ± 0,00

reliability of the DEL model, 5 cross-fold validations were applied. Performance criteria were calculated with the DEL model trained on the test dataset.

Classification metrics in deep learning models include accuracy, precision, recall, specificity, F1-score, and AUC-ROC [72]. Accuracy is the fraction of correctly classified examples across all classes. Precision measures how many of the predicted classes were correct. Recall measures how many of the actual classes were correctly identified. Specificity is the fraction of negative examples correctly identified. F1-score combines recall and precision into a single metric. AUC-ROC is the area under the receiver-operating characteristic curve, which measures the proportion of true positives out of all positive predictions.

In Table 2, the AD data set was randomly created with 5 folds. For each fold, the DEL model was trained with the data in the fold, and the performance metrics of the model were calculated with the test data. Cross-fold training involves splitting the training dataset into two parts; one part (called the test set) is used to measure the model's performance, while the other (called the training set) is used for actual training. This process is repeated for each iteration of training, with the validation set rotating such that different parts of the data are evaluated on each iteration. This allows for a more accurate estimate of the model's performance since each part of the data set is evaluated. Crossfold training is also a great way to combat overfitting since it forces the model to learn general features across the dataset rather than just memorizing the data presented in a single iteration. In Table 1, the DEL model's average classification mean ± standart deviation accuracy was calculated as 97.9% ± 0.01. When other performance metrics are examined, a model has been developed to solve the overfitting problem.

The AD datasets are divided into Delta, Theta, Alpha, Beta, and Gamma frequency bands with a butterwort band-pass filter. For each band, the DEL model is trained with 5 cross folds. At the end of the training, the performances of the DEL model were measured using the test data allocated for each fold from the AD dataset.

A confusion matrix is a table that evaluates a classification model's performance. It is a table of numbers that helps to evaluate classification accuracy by comparing the predicted classes to the true classes. It can identify correct and incorrect predictions made by a model and provides an easy-to-read format for evaluating its performance. The confusion

Table 5
The Deep Ensemble Model (DEL) combined internal 2D-CNN models performance metrics for AD datasets 4–8 Hz Frequency Band (Theta).

	DEL Model Mean ± std	2D-CNN Model 1 Mean ± std	2D-CNN Model 2 Mean ± std	2D-CNN Model 3 Mean ± std	2D-CNN Model 4 Mean ± std	2D-CNN Model 5 Mean ± std
Accuracy	0,93 ± 0,00	0,83 ± 0,20	0,95 ± 0,01	0,88 ± 0,02	0,89 ± 0,02	0,71 ± 0,02
F1 score	0,91 ± 0,01	0,82 ± 0,19	0,94 ± 0,01	0,85 ± 0,01	0,86 ± 0,03	0,41 ± 0,00

Table 6
The Deep Ensemble Model (DEL) combined internal 2D-CNN models performance metrics for AD datasets 8–14 Hz Frequency Band (Alpha).

	DEL Model Mean ± std	2D-CNN Model 1 Mean ± std	2D-CNN Model 2 Mean ± std	2D-CNN Model 3 Mean ± std	2D-CNN Model 4 Mean ± std	2D-CNN Model 5 Mean ± std
Accuracy	0,92 ± 0,00	0,84 ± 0,21	0,94 ± 0,02	0,89 ± 0,03	0,89 ± 0,03	0,70 ± 0,03
F1 score	0,90 ± 0,01	0,83 ± 0,18	0,93 ± 0,01	0,84 ± 0,02	0,87 ± 0,02	0,42 ± 0,01

Table 7
The Deep Ensemble Model (DEL) combined internal 2D-CNN models performance metrics for AD datasets 14–30 Hz Frequency Band (Beta).

	DEL Model Mean ± std	2D-CNN Model 1 Mean ± std	2D-CNN Model 2 Mean ± std	2D-CNN Model 3 Mean ± std	2D-CNN Model 4 Mean ± std	2D-CNN Model 5 Mean ± std
Accuracy	0,93 ± 0,02	0,96 ± 0,01	0,96 ± 0,01	0,90 ± 0,06	0,84 ± 0,12	0,71 ± 0,02
F1 score	0,91 ± 0,03	0,95 ± 0,01	0,95 ± 0,01	0,85 ± 0,10	0,72 ± 0,27	0,41 ± 0,00

Table 8
The Deep Ensemble Model (DEL) combined internal 2D-CNN models performance metrics for AD datasets 30 Hz above Frequency Band (Gama).

	DEL Model Mean ± std	2D-CNN Model 1 Mean ± std	2D-CNN Model 2 Mean ± std	2D-CNN Model 3 Mean ± std	2D-CNN Model 4 Mean ± std	2D-CNN Model 5 Mean ± std
Accuracy	0,95 ± 0,00	0,94 ± 0,01	0,95 ± 0,01	0,87 ± 0,05	0,87 ± 0,04	0,71 ± 0,02
F1 score	0,93 ± 0,01	0,93 ± 0,02	0,94 ± 0,01	0,83 ± 0,10	0,81 ± 0,10	0,41 ± 0,00

matrix contains four values: true positive (TP), false positive (FP), true negative (TN), and false negative (FN) [72]. These numbers are used to calculate several accuracy measures, including accuracy, precision, and recall.

Additionally, the matrix rows can be used to find the percentage of each type of prediction the model made. When Fig. 5 is examined, it is seen that the DEL model and DEL model classifier model confusion matrix are like each other. In the DEL model, the TP ratio was calculated as 70.98%. It is seen that the AD-HC classification success of the confusion matrix DEL model in Fig. 5 is very high.

Table 9
The McNamara's test, 2 × 2 contingency table, tables the outcomes of DEL classification of two model's pairs on an AD dataset. a) 2D-CNN Model 1 vs 2D-CNN Model 2 pair, b) 2D-CNN Model 1 vs 2D-CNN Model 3, c) 2D-CNN Model 1 vs 2D-CNN Model 4, d) 2D-CNN Model 1 vs 2D-CNN Model 5.

		2D-CNN Model 2					2D-CNN Model 3		
		True	False	Total			True	False	Total
2D-CNN Model 1	True	220	4	224	2D-CNN Model 1	True	161	63	224
	False	0	1	4		False	0	1	1
	Total	220	5	225		Total	161	64	225

		2D-CNN Model 4					2D-CNN Model 5		
		True	False	Total			True	False	Total
2D-CNN Model 1	True	161	63	224	2D-CNN Model 1	True	161	63	224
	False	0	1	1		False	0	1	1
	Total	161	64	225		Total	161	64	225

Table 10
The p and statistical values of the DEL model pairs for the AD dataset.

2D-CNN Model 1 vs 2D-CNN Model 2	p value	0.13
	statistic	2.25
2D-CNN Model 1 vs 2D-CNN Model 3	p value	5.66
	statistic	61.1
2D-CNN Model 1 vs 2D-CNN Model 4	p value	5.66
	statistic	61.1
2D-CNN Model 1 vs 2D-CNN Model 5	p value	5.66
	statistic	61.1

When Table 3–8 is examined, it is seen that the highest AD classification performance is obtained from the EEG data in the 0.5–45 band. It is seen that the highest accuracy of the EEG signals from the 5 frequency bands and the F1 score is obtained from the Gamma frequency band. When the AD dataset used is examined in terms of AD-HC classification, it is seen that the Gamma frequency band is more distinctive. It is seen that the discrimination of EEG signals in the Beta band in AD-HC classification is less than the Gamma band and higher than the other frequency bands. When the DEL model and internal 2D-CNN model performance metrics are examined, DEL model performance is lower than the 2D-CNN model performance. Overfitting problems may occur from internal 2D-CNN models. A model has been developed to solve the overfitting problem with the DEL model.

McNamara's test is a statistical method used to compare two different classification models. It compares the performance of two models based on the results of a single binary classification task. McNamara's test is used to determine if the difference between the accuracy of the two models is statistically significant. To conduct McNamara's test, a contingency table is constructed that shows the number of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) results that each model has produced. The McNamara's test statistic is then calculated as the differences between the numbers of FP and FN values between the two models. If the test statistic is greater than a certain threshold (specified by the researcher), then the difference between the two models is deemed statistically significant.

The McNamara's test, 2 × 2 contingency table, which tables the outcomes of DEL classification of two model pairs on an AD dataset, can be seen in Table 9. A chi-square table is a table that tells the probability of observing specific values from a sample of data given a hypothesized value. The critical value (p) is the probability value selected before the test to determine if the observed results are statistically meaningful compared to the hypothesized value. 2x2 contingency table and $p = 0.05$ and $s = 3.841$ values were selected from the Chi-Square table. If $p < 0.05$ and $s < 3.841$ values, it indicates a difference between the two classification models used. When Table 10. is examined, it can be concluded that the DEL classifier model 1 used is different from the other DEL classifier models (model 2, model 3, model 4, and model 5) since the p and s values are greater than the p and s values selected in the Chi-Square table.

Table 11
The performance comparison in the Alzheimer's Disease's (AD) diagnosis based EEG signals in the literature.

Patients group	Features	Models	Metrics best	Dataset	Ref.
MAD,HC, MCI, MSAD	Pearson correlation coefficient	linear discriminant analysis (LDA)	Accuracy: 79.31%	EEG-fNIRS dataset, MAD(6),HC (8), MCI (8), MSAD(7) subjects	[73]
AD, HC and MCI	Time–frequency representation (TFR)	Conv-AE, Modified CNN	Accuracy: 89% Accuracy: 92%	Scalp EEGAD (37), HC(16) and MCI (37) subjects	[35]
AD and HC	Robust-principal component analysis (R-PCA)	naive Bayes,SVM Gaussian	Accuracy: 90.91% Accuracy: 93.18%	EEG SignalsAD (7), HC(6) subjects	[74]
HC, Pre-symptomatic AD (aAD) and Prodromal AD (pAD)	event-related spectral perturbation (ERSP)	DeepADNet: A CNN-LSTM	Accuracy: 75.95%	Multichannel EEG 23 HC, 17 Pre-symptomatic AD (aAD) and 23 Prodromal AD (pAD) subjects	[75]
AD, HC, MCI, other Dementia	Hjorth metrics, STFT, powers, sample entropy, and microstate measures	Random forest	Accuracy: 70%	EEG SignalsAD (330), HC (246), MCI (189), other Dementia (125) subjects	[76]
AD, MCI	the integrated multiple signal classification and empirical wavelet transform (MUSIC-EWT)	EPNN	Accuracy: 90.3%.	EEG SignalsAD (37), MCI (37) subjects	[77]
AD, HC	power spectral density (PSD) features	AdaboostM1, Total Boost, Gentle Boost, Logit Boost, Robust Boost, and Bagging ensemble learning	Accuracy: 93.04%	EEG SignalsAD (24), HC (24)	[78]
AD and HC	–	**Deep Ensemble Learning (DEL) Model (our proposed method)**	**Accuracy: 97.9 ± 1.1% (mean ± standard deviation)**	**Dataset 1: AD (24), HC (24) Dataset 2: AD (80), HC (12)**	–

4. Discussion

AD is a neurodegenerative disorder resulting in progressive memory loss and cognitive impairment. It is the most common cause of Dementia, which affects millions of people worldwide and is a significant health problem for the elderly. Accurate diagnosis of AD is paramount to its successful treatment. Yet, the process can be problematic given the complexity of the condition and its early-stage symptoms, which can range from mild cognitive impairments to severe behavioral impairments. The use of DEL for diagnosing and classifying AD is an increasingly attractive option due to its capacity to exploit large datasets and identify complex patterns in the data. DEL is a machine learning approach that combines multiple models to obtain optimal results. DEL creates an ensemble of classifiers using different training data and feature selection methods. This can result in much higher generalization performance than a single classifier. The use of DEL in diagnosing and classifying AD is particularly well-suited because it can identify patterns in many available data, such as EEG signals. EEG signals can provide valuable information regarding a person's neurological health and detect abnormalities indicative of AD. EEG signals can identify specific frequency bands of oscillations associated with AD, such as delta, theta, alpha, and gamma bands.

In this study, the performance of DEL model was evaluated for the diagnosis and classification of AD. Two different AD-HC datasets were combined and used in this article. The model used a 2D-CNN to extract features from EEG signals and classify AD-HC—ensemble methods to combine five different 2D-CNN classifiers. The proposed DEL model was applied to publicly available EEG signals, demographic data, and cognitive scores datasets. The study results showed that the DEL model achieved a high accuracy of 0,97 ± 0,01 (mean ± standard deviation) and excellent performance metrics, such as F1 score, precision, recall, kappa score, and Jaccard score. The model was further evaluated through five-fold cross-validation, and the results showed that the DEL model outperformed a more traditional machine learning approach. To conclude, using DEL for diagnosing and classifying AD has demonstrated the potential for providing highly accurate results. This approach could offer great value in diagnosing and treating AD and should be further explored.

When Table 11 is examined, it is seen that AD EEG data sets are divided into three classes in general. These are AD, HC, and MCI. Considering the complexity of EEG signals, classifier models should be reduced to more straightforward features that can be calculated. In studies in the literature, AD has been classified in artificial intelligence classifier models by extracting different features from EEG signals. It is seen that different numbers of subjects are used in AD classification literature studies. These subjects were determined as AD, HC, and MCI by scoring the memory tests and brain imaging methods. Although the artificial intelligence models used in AD classification reach different accuracy rates, the proposed DEL model has reached a remarkable accuracy value. With 5 cross-fold training, the average accuracy of the DEL model reached 97.9% ± 1.1%. It will enable the DEL model to reach higher accuracy with future studies' best internal classifying model selections.

5. Conclusion

In this paper, we have successfully proposed a compelling novel DEL model for accurately diagnosing and classifying AD using EEG signals. We used 2D-CNN models as classifiers for the DEL model to achieve higher accuracy and robustness. Instead of extracting features from the EEG signal, we applied them directly to 2D-CNN models. The features of the model were created automatically in the convolution layers. We then applied five-fold cross-validation data learning to evaluate classification performance. Experimental results showed that the DEL model correctly classified patients with AD with an overall accuracy of 0,97 ± 0,01. In addition, the proposed model has an F1 score, precision, recall, ROC-AUC score, kappa score, and Jaccard score of 0,97 ± 0,01, 0,96 ± 0,01, 0,96 ± 0,01, 0,98 ± 0,00, 0,92 ± 0,03, and 0,94 ± 0,02, respectively. We achieved excellent average performance measurements.

The proposed DEL model is used to identify AD from EEG signals. Fortunately, there are many areas for improvement in future work. First, the model's accuracy can be further improved using regularization techniques. For example, combining DEL with evolutionary-based optimization techniques can help find optimal hyperparameter settings. Second, the proposed model lacks interpretability, and it is essential to use explainable Artificial intelligence (AI) techniques such as rule extraction to understand the model's decision-making process. Finally, the proposed model can be extended to several other neurological disorders, such as Parkinson's, epilepsy, and stroke, to diagnose and classify various diseases. A combined approach combining multiple modalities (EEG, MRI, PET) can also be used to build a robust deep-learning model for diagnosing neurological diseases.

Data availability

In this paper, we have used this first dataset: (A. S. L. O. Campanharo, F. M. Ramos, A. M. Pineda, and L. E. Betting, "Data from: Quantile Graphs for EEG-Based Diagnosis of Alzheimer's Disease," 2020, doi: 10.17605/OSF.IO/S74QF, online: https://osf.io/pa38y) We can send the datasets at the request of the authors.

Second dataset : (Vicchietti, Mário L., Ramos, Fernando M., Betting, Luiz E., Campanharo, Andriana S. L. O."Data from: Computational methods of EEG signals analysis for Alzheimer's disease classification"2023, 10.17605/OSF.IO/2V5MD, online: https://osf.io/2v5md/?view_only=). We can send the datasets at the request of the authors.

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.

Data availability

Data will be made available on request.

Acknowledgement

This research work was funded by Institutional Fund Projects under grant No. (IFPIP: 1038-135-1443). The authors gratefully acknowledge technical and financial support provided by the Ministry of Education and King Abdulaziz University, DSR, Jeddah, Saudi Arabia.

References

[1] "World Alzheimer Report 2022: Life after diagnosis: Navigating treatment, care and support." Sep. 21, 2022. Accessed: May 26, 2023. [Online]. Available: https://www.alzint.org/resource/world-alzheimer-report-2022/.
[2] B. Kim, G.O. Noh, K. Kim, Behavioural and psychological symptoms of dementia in patients with Alzheimer's disease and family caregiver burden: a path analysis, BMC Geriatr 21 (1) (Dec. 2021) 1–12, https://doi.org/10.1186/S12877-021-02109-W/TABLES/4.
[3] H. Ji, Z. Liu, W. Q. Yan, R. Klette, Early diagnosis of Alzheimer's disease using deep learning, in: ACM International Conference Proceeding Series, pp. 87–91, Jun. 2019, doi: 10.1145/3341016.3341024.
[4] A. Atri, The Alzheimer's Disease Clinical Spectrum: Diagnosis and Management, Medical Clinics 103 (2) (Mar. 2019) 263–293, https://doi.org/10.1016/J.MCNA.2018.10.009.
[5] Y. Agrawal, P. F. Smith, P. B. Rosenberg, Vestibular impairment, cognitive decline and Alzheimer's disease: balancing the evidence, https://doi.org/10.1080/13607863.2019.1566813, vol. 24, no. 5, pp. 705–708, May 2019, doi: 10.1080/13607863.2019.1566813.
[6] H. Jahn, Memory loss in Alzheimer's disease, https://doi.org/10.31887/DCNS.2013.15.4/hjahn, vol. 15, no. 4, pp. 445–454, Dec. 2022, doi: 10.31887/DCNS.2013.15.4/HJAHN.

[7] F. S. Dafsari, F. Jessen, Depression—an underrecognized target for prevention of dementia in Alzheimer's disease, Transl. Psychiatry 10(1) (2020) 1–13, May 2020, doi: 10.1038/s41398-020-0839-1.
[8] F.K. Clemmensen, et al., The role of physical and cognitive function in performance of activities of daily living in patients with mild-to-moderate Alzheimer's disease – a cross-sectional study, BMC Geriatr. 20 (1) (Dec. 2020) 1–9, https://doi.org/10.1186/S12877-020-01926-9/TABLES/2.
[9] H. Ben Abdessalem, Y. Ai, K. S. Marulasidda Swamy, and C. Frasson, Virtual Reality Zoo Therapy for Alzheimer's Disease Using Real-Time Gesture Recognition, Adv. Exp. Med. Biol. vol. 1338, pp. 97–105, 2021, doi: 10.1007/978-3-030-78775-2_12/COVER.
[10] T.C.C. Pinto, et al., Is the Montreal Cognitive Assessment (MoCA) screening superior to the Mini-Mental State Examination (MMSE) in the detection of mild cognitive impairment (MCI) and Alzheimer's Disease (AD) in the elderly? Int. Psychogeriatr. 31 (4) (Apr. 2019) 491–504, https://doi.org/10.1017/S1041610218001370.
[11] G. Chételat, et al., Amyloid-PET and 18F-FDG-PET in the diagnostic investigation of Alzheimer's disease and other dementias, Lancet Neurol. 19 (11) (Nov. 2020) 951–962, https://doi.org/10.1016/S1474-4422(20)30314-8.
[12] S. Basaia, et al., Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks, Neuroimage Clin. 21 (Jan. 2019), 101645, https://doi.org/10.1016/J.NICL.2018.101645.
[13] K. Dhiman, K. Blennow, H. Zetterberg, R. N. Martins, V. B. Gupta, Cerebrospinal fluid biomarkers for understanding multiple aspects of Alzheimer's disease pathogenesis, Cell. Mol. Life Sci. 76(10) (2019)1833–1863, Feb. 2019, doi: 10.1007/S00018-019-03040-5.
[14] Y. Song, et al., The Effect of Estrogen Replacement Therapy on Alzheimer's Disease and Parkinson's Disease in Postmenopausal Women: A Meta-Analysis, Front. Neurosci. 14 (Mar. 2020) 157, https://doi.org/10.3389/FNINS.2020.00157.
[15] S. Invernizzi, I. Simoes Loureiro, K.G. Kandana Arachchige, L. Lefebvre, Late-Life Depression, Cognitive Impairment, and Relationship with Alzheimer's Disease, Dement Geriatr. Cogn. Disord. 50 (5) (Jan. 2022) 414–424, https://doi.org/10.1159/000519453.
[16] "All you need to know about brain scans and dementia - Alzheimer's Research UK." Accessed: May 27, 2023. [Online]. Available: https://www.alzheimersresearchuk.org/blog/all-you-need-to-know-about-brain-scans-and-dementia/.
[17] "Definition of lumbar puncture - NCI Dictionary of Cancer Terms - NCI." Accessed: May 27, 2023. [Online]. Available: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/lumbar-puncture.
[18] D. A. Arafa, H. E. D. Moustafa, A. M. T. Ali-Eldin, H. A. Ali, Early detection of Alzheimer's disease based on the state-of-the-art deep learning approach: a comprehensive survey, Multimedia Tools Appl. 81(17) (2022) 23735–23776, doi: 10.1007/S11042-022-11925-0.
[19] V. Joshi, N. Nanavati, A review of EEG signal analysis for diagnosis of neurological disorders using machine learning, J. Biomed. Photonics Eng. 7 (4) (2021) pp, https://doi.org/10.18287/JBPE21.07.040201.
[20] M. Gawel, E. Zalewska, E. Szmidt-Sałkowska, J. Kowalski, The value of quantitative EEG in differential diagnosis of Alzheimer's disease and subcortical vascular dementia, J. Neurol. Sci. 283 (1–2) (Aug. 2009) 127–133, https://doi.org/10.1016/J.JNS.2009.02.332.
[21] J. Jeong, EEG dynamics in patients with Alzheimer's disease, Clin. Neurophysiol. 115 (7) (Jul. 2004) 1490–1505, https://doi.org/10.1016/J.CLINPH.2004.01.001.
[22] Y.T. Kwak, Quantitative EEG findings in different stages of Alzheimer's disease, J. Clin. Neurophysiol. 23 (5) (Oct. 2006) 457–462, https://doi.org/10.1097/01.WNP.0000223453.47663.63.
[23] P. Saroja, N.J. Nalini, A Systematic Study About EEG Signal Data and Computer Aided Models for the Diagnosis of Alzheimer's Disease, Lecture Notes on Data Eng. Commun. Technol. 139 (2023) 519–531, https://doi.org/10.1007/978-981-19-3015-7_38/COVER.
[24] P.M. Rossini, et al., Early diagnosis of Alzheimer's disease: the role of biomarkers including advanced EEG signal analysis. Report from the IFCN-sponsored panel of experts, Clin. Neurophysiol. 131 (6) (Jun. 2020) 1287–1310, https://doi.org/10.1016/J.CLINPH.2020.03.003.
[25] H. Azami, et al., Beta to theta power ratio in EEG periodic components as a potential biomarker in mild cognitive impairment and Alzheimer's dementia, Alzheimers Res. Ther. 15 (1) (Dec. 2023) 1–12, https://doi.org/10.1186/S13195-023-01280-Z/FIGURES/5.
[26] R.X. Li, Y.H. Ma, L. Tan, J.T. Yu, Prospective biomarkers of Alzheimer's disease: A systematic review and meta-analysis, Ageing Res. Rev. 81 (Nov. 2022), 101699, https://doi.org/10.1016/J.ARR.2022.101699.
[27] P. Khan, et al., Machine Learning and Deep Learning Approaches for Brain Disease Diagnosis: Principles and Recent Advances, IEEE Access 9 (2021) 37622–37655, https://doi.org/10.1109/ACCESS.2021.3062484.
[28] M.S. Safi, S.M.M. Safi, Early detection of Alzheimer's disease from EEG signals using Hjorth parameters, Biomed. Signal Process Control 65 (Mar. 2021), 102338, https://doi.org/10.1016/J.BSPC.2020.102338.
[29] C. Balaji, D. S. Suresh, Multiclass recognition of Alzheimer's and Parkinson's disease using various machine learning techniques: A study, https://doi.org/10.1142/S1793962322500088 13(1) (2021), doi: 10.1142/S1793962322500088.
[30] A. Miltiadous, et al., Alzheimer's Disease and Frontotemporal Dementia: A Robust Classification Method of EEG Signals and a Comparison of Validation Methods, Diagnostics 11(8) (2021) 1437, doi: 10.3390/DIAGNOSTICS11081437.
[31] K. Alsharabi, Y. Bin Salamah, A.M. Abdurraqeeb, M. Aljalal, F.A. Alturki, EEG Signal Processing for Alzheimer's Disorders Using Discrete Wavelet Transform and Machine Learning Approaches, IEEE Access 10 (2022) 89781–89797, https://doi.org/10.1109/ACCESS.2022.3198988.
[32] A. M. Tautan, et al., Preliminary study on the impact of EEG density on TMS-EEG classification in Alzheimer's disease, in: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, vol. 2022-July, pp. 394–397, 2022, doi: 10.1109/EMBC48229.2022.9870920.
[33] D. Klepl, F. He, M. Wu, D. J. Blackburn, P. G. Sarrigiannis, Adaptive Gated Graph Convolutional Network for Explainable Diagnosis of Alzheimer's Disease using EEG Data, Apr. 2023, Accessed: May 26, 2023. [Online]. Available: https://arxiv.org/abs/2304.05874v1.
[34] A. Shikalgar, S. Sonavane, Hybrid Deep Learning Approach for Classifying Alzheimer Disease Based on Multimodal Data, Adv. Intell. Syst. Comput. 1025 (2020) 511–520, https://doi.org/10.1007/978-981-32-9515-5_49/COVER.
[35] S. Fouladi, A. A. Safaei, N. Mammone, F. Ghaderi, M. J. Ebadi, Efficient Deep Neural Networks for Classification of Alzheimer's Disease and Mild Cognitive Impairment from Scalp EEG Recordings, Cognit. Computat. 14(4) (2022) 1247–1268, Jun. 2022, doi: 10.1007/S12559-022-10033-3.
[36] C.J. Huggins, et al., Deep learning of resting-state electroencephalogram signals for three-class classification of Alzheimer's disease, mild cognitive impairment and healthy ageing, J. Neural Eng. 18 (4) (Jun. 2021), 046087, https://doi.org/10.1088/1741-2552/AC05D8.
[37] M. Alessandrini, G. Biagetti, P. Crippa, L. Falaschetti, S. Luzzi, and C. Turchetti, EEG-Based Alzheimer’s Disease Recognition Using Robust-PCA and LSTM Recurrent Neural Network, Sensors 22(10) (2022) 3696, doi: 10.3390/S22103696.
[38] M. Amini, M.M. Pedram, A.R. Moradi, M. Ouchani, Diagnosis of Alzheimer's Disease by Time-Dependent Power Spectrum Descriptors and Convolutional Neural Network Using EEG Signal, Comput. Math. Methods Med. 2021 (2021), https://doi.org/10.1155/2021/5511922.
[39] U. Senturk, K. Polat, I. Yucedag, A non-invasive continuous cuffless blood pressure estimation using dynamic Recurrent Neural Networks, Appl. Acoust. 170 (Dec. 2020), 107534, https://doi.org/10.1016/J.APACOUST.2020.107534.
[40] U. Senturk, I. Yucedag, K. Polat, Repetitive neural network (RNN) based blood pressure estimation using PPG and ECG signals, in: ISMSIT 2018 - 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies, Proceedings, Dec. 2018, doi: 10.1109/ISMSIT.2018.8567071.
[41] G. Sharma, A. Parashar, A.M. Joshi, DepHNN: A novel hybrid neural network for electroencephalogram (EEG)-based screening of depression, Biomed. Signal Process Control 66 (Apr. 2021), 102393, https://doi.org/10.1016/J.BSPC.2020.102393.
[42] M. Imani, Alzheimer's diseases diagnosis using fusion of high informative BiLSTM and CNN features of EEG signal, Biomed. Signal Process Control 86 (Sep. 2023), 105298, https://doi.org/10.1016/J.BSPC.2023.105298.
[43] S. Nobukawa, T. Yamanishi, S. Kasakawa, H. Nishimura, M. Kikuchi, T. Takahashi, Classification Methods Based on Complexity and Synchronization of Electroencephalography Signals in Alzheimer's Disease, Front Psychiatry 11 (Apr. 2020), 511787, https://doi.org/10.3389/FPSYT.2020.00255/BIBTEX.
[44] 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) (Jan. 2020) 60–71, https://doi.org/10.1109/TFUZZ.2019.2903753.
[45] X. Bi, H. Wang, Early Alzheimer's disease diagnosis based on EEG spectral images using deep learning, Neural Netw. 114 (Jun. 2019) 119–135, https://doi.org/10.1016/J.NEUNET.2019.02.005.
[46] I. A. Fouad, F. El-Zahraa, M. Labib, Identification of Alzheimer's disease from central lobe EEG signals utilizing machine learning and residual neural network, Biomed. Signal Process Control 86, p. 105266, Sep. 2023, doi: 10.1016/J.BSPC.2023.105266.
[47] 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 11 (2023) 71840–71858, https://doi.org/10.1109/ACCESS.2023.3294618.
[48] O. A. Dara, J. M. Lopez-Guede, H. I. Raheem, J. Rahebi, E. Zulueta, and U. Fernandez-Gamiz, Alzheimer's Disease Diagnosis Using Machine Learning: A Survey, Appl. Sci. 13 (14) (2023) 8298, doi: 10.3390/APP13148298.
[49] D. Abdolzadegan, M.H. Moattar, M. Ghoshuni, A robust method for early diagnosis of autism spectrum disorder from EEG signals based on feature selection and DBSCAN method, Biocybern Biomed. Eng. 40 (1) (Jan. 2020) 482–493, https://doi.org/10.1016/J.BBE.2020.01.008.
[50] P.K. Mall, et al., A comprehensive review of deep neural networks for medical image processing: Recent developments and future opportunities, Healthcare Analytics 4 (Dec. 2023), 100216, https://doi.org/10.1016/J.HEALTH.2023.100216.
[51] M.A. Abdou, Literature review: efficient deep neural networks techniques for medical image analysis, Neural Comput. Appl. 34 (8) (Apr. 2022) 5791–5812, https://doi.org/10.1007/S00521-022-06960-9/TABLES/5.
[52] J. Moon, S. Jung, J. Rew, S. Rho, E. Hwang, Combination of short-term load forecasting models based on a stacking ensemble approach, Energy Build 216 (Jun. 2020), 109921, https://doi.org/10.1016/J.ENBUILD.2020.109921.
[53] D. K. Nguyen, C. H. Lan, C. L. Chan, Deep Ensemble Learning Approaches in Healthcare to Enhance the Prediction and Diagnosing Performance: The Workflows, Deployments, and Surveys on the Statistical, Image-Based, and Sequential Datasets, Int. J. Environ. Res. Public Health 18(20) (2021) 10811, doi: 10.3390/IJERPH182010811.
[54] A. Mohammed, R. Kora, A comprehensive review on ensemble deep learning: Opportunities and challenges, J. King Saud Univ. – Comput. Inform. Sci. 35 (2) (Feb. 2023) 757–774, https://doi.org/10.1016/J.JKSUCI.2023.01.014.
[55] S. Kumar, P. Kaur, A. Gosain, A Comprehensive Survey on Ensemble Methods, in: 2022 IEEE 7th International conference for Convergence in Technology, I2CT 2022, 2022, doi: 10.1109/I2CT54291.2022.9825269.

[56] "(PDF) Towards a home-use BCI: fast asynchronous control and robust non-control state detection." Accessed: May 27, 2023. [Online]. Available: https://www.researchgate.net/publication/338423585_Towards_a_home-use_BCI_fast_asynchronous_control_and_robust_non-control_state_detection.
[57] A. S. L. O. Campanharo, F. M. Ramos, A. M. Pineda, and L. E. Betting, Data from: Quantile Graphs for EEG-Based Diagnosis of Alzheimer's Disease, 2020, doi: 10.17605/OSF.IO/S74QF.
[58] M. L. Vicchietti, F. M. Ramos, L. E. Betting, A. S. L. O. Campanharo, Data from: Computational methods of EEG signals analysis for Alzheimer's disease classification, 2023, doi: 10.17605/OSF.IO/2V5MD.
[59] W.S. Pritchard, et al., EEG-based, neural-net predictive classification of Alzheimer's disease versus control subjects is augmented by non-linear EEG measures, Electroencephalogr. Clin. Neurophysiol. 91 (2) (Aug. 1994) 118–130, https://doi.org/10.1016/0013-4694(94)90033-7.
[60] A.M. Pineda, F.M. Ramos, L.E. Betting, A.S.L.O. Campanharo, Quantile graphs for EEG-based diagnosis of Alzheimer's disease, PLoS One 15 (6) (Jun. 2020), https://doi.org/10.1371/JOURNAL.PONE.0231169.
[61] M. L. Vicchietti, F. M. Ramos, L. E. Betting, A. S. L. O. Campanharo, Computational methods of EEG signals analysis for Alzheimer's disease classification, Scientific Reports 13(1) (2023) 1–14, May 2023, doi: 10.1038/s41598-023-32664-8.
[62] 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. Top Comput. Intell. (Apr. 2022), https://doi.org/10.1109/TETCI.2022.3186180.
[63] F. A. Alturki, M. Aljalal, A. M. Abdurraqeeb, K. Alsharabi, and A. A. Al-Shamma'a, Common Spatial Pattern Technique with EEG Signals for Diagnosis of Autism and Epilepsy Disorders, IEEE Access, vol. 9, pp. 24334–24349, 2021, doi: 10.1109/ACCESS.2021.3056619.
[64] M. Savadkoohi, T. Oladunni, L. Thompson, A machine learning approach to epileptic seizure prediction using Electroencephalogram (EEG) Signal, Biocybern Biomed. Eng. 40 (3) (Jul. 2020) 1328–1341, https://doi.org/10.1016/J.BBE.2020.07.004.
[65] J.J. Newson, T.C. Thiagarajan, EEG Frequency Bands in Psychiatric Disorders: A Review of Resting State Studies, Front. Hum. Neurosci. 12 (Jan. 2019) 521, https://doi.org/10.3389/FNHUM.2018.00521/BIBTEX.
[66] E. Fide, D. Yerlikaya, D. Öz, İ. Öztura, G. Yener, Normalized Theta but Increased Gamma Activity after Acetylcholinesterase Inhibitor Treatment in Alzheimer's Disease: Preliminary qEEG Study, Clin. EEG Neurosci. (May 2022), https://doi.org/10.1177/15500594221120723/ASSET/IMAGES/LARGE/10.1177_15500594221120723-FIG5.JPEG.
[67] E. Parvin, F. Mohammadian, S. Amani-Shalamzari, M. Bayati, B. Tazesh, Dual-Task Training Affect Cognitive and Physical Performances and Brain Oscillation Ratio of Patients With Alzheimer's Disease: A Randomized Controlled Trial, Front. Aging Neurosci. 12 (Dec. 2020) 456, https://doi.org/10.3389/FNAGI.2020.605317/BIBTEX.
[68] A. D'Atri, et al., EEG alterations during wake and sleep in mild cognitive impairment and Alzheimer's disease, iScience 24 (4) (Apr. 2021), 102386, https://doi.org/10.1016/J.ISCI.2021.102386.
[69] R.A. Wirt, et al., Altered theta rhythm and hippocampal-cortical interactions underlie working memory deficits in a hyperglycemia risk factor model of Alzheimer's disease, Commun. Biol. 2021 4:1, vol. 4, no. 1, pp. 1–16, Sep. 2021, doi: 10.1038/s42003-021-02558-4.
[70] M. Abdar, V. Makarenkov, CWV-BANN-SVM ensemble learning classifier for an accurate diagnosis of breast cancer, Measurement 146 (Nov. 2019) 557–570, https://doi.org/10.1016/J.MEASUREMENT.2019.05.022.
[71] A high-bias, low-variance introduction to Machine Learning for physicists - ScienceDirect. Accessed: Oct. 08, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S037015731930076 6.
[72] U. Senturk, K. Polat, I. Yucedag, F. Alenezi, Arrhythmia diagnosis from ECG signal pulses with one-dimensional convolutional neural networks, in: Diagnostic Biomedical Signal and Image Processing Applications with Deep Learning Methods, pp. 83–101, Jan. 2023, doi: 10.1016/B978-0-323-96129-5.00002-0.
[73] P.A. Cicalese, et al., An EEG-fNIRS hybridization technique in the four-class classification of alzheimer's disease, J. Neurosci. Methods 336 (Apr. 2020), 108618, https://doi.org/10.1016/J.JNEUMETH.2020.108618.
[74] G. Biagetti, P. Crippa, L. Falaschetti, S. Luzzi, C. Turchetti, Classification of Alzheimer's Disease from EEG Signal Using Robust-PCA Feature Extraction, Procedia Comput. Sci. 192 (Jan. 2021) 3114–3122, https://doi.org/10.1016/J.PROCS.2021.09.084.
[75] T.K.K. Ho, et al., DeepADNet: A CNN-LSTM model for the multi-class classification of Alzheimer's disease using multichannel EEG, Alzheimer's Dementia 17 (Dec. 2021) e057573.
[76] B. Jiao, et al., Neural biomarker diagnosis and prediction to mild cognitive impairment and Alzheimer's disease using EEG technology, Alzheimers Res. Ther. 15 (1) (2023) pp, https://doi.org/10.1186/s13195-023-01181-1.
[77] J.P. Amezquita-Sanchez, N. Mammone, F.C. Morabito, S. Marino, H. Adeli, A novel methodology for automated differential diagnosis of mild cognitive impairment and the Alzheimer's disease using EEG signals, J. Neurosci. Methods 322 (2019), https://doi.org/10.1016/j.jneumeth.2019.04.013.
[78] H. Göker, Detection of Alzheimer's disease from electroencephalography (EEG) signals using multitaper and ensemble learning methods, Uludağ Univ. J. Faculty Eng. 28 (1) (Apr. 2023) 141–152, https://doi.org/10.17482/UUMFD.1142345.