Al-Nahrain Journal for Engineering Sciences **NJES** 28(4)545-554, 2025
http://doi.org/10.29194/NJES.28040545

Deep Learning-Based Classification of Alzheimer's Disease Using EEG Signals: A CNN Approach for Early Detection

Najlaa S. Mezher[1*], Ahmed F. Hussein[2], Sufian M. Salih[3]

Authors affiliations:
1*) Department of Biomedical Engineering, Al-Nahrain University, Baghdad-Iraq.
najlaa.suhail.msc2023@ced.nahrainuniv.edu.iq.

2)Department of Artificial Intelligence and Robotics, Al-Nahrain University, Baghdad-Iraq.
ahmed.f.hussein@nahrainuniv.edu.iq.

3) Department of Banking Management Economics, Al-Nahrain University, Baghdad-Iraq.
sufian.m.salih@nahrainuniv.edu.iq

Paper History:

Received: 23rd Dec. 2024

Revised: 12th Mar. 2025

Accepted: 19th Apr. 2025

Abstract

Alzheimer's disease (AD) is a progressive neurodegenerative disorder that severely impacts cognitive functions such as memory, attention, and reasoning, ultimately affecting daily life. Early and accurate detection is crucial for timely intervention and management. Traditional diagnostic methods, including neuroimaging and cognitive assessments, can be expensive and time-consuming, necessitating more accessible and efficient alternatives. This study aims to develop an automated and efficient deep learning-based detection system that uses Electroencephalogram (EEG) signals to accurately classify AD and healthy individuals. A Convolutional Neural Network (CNN) model was designed to extract meaningful features from preprocessed EEG data. The architecture consists of convolutional layers with max pooling, dropout regularization, and fully connected layers to improve classification accuracy. The model was trained and evaluated on a comprehensive EEG dataset, using key performance metrics such as accuracy, recall, precision, and F1-score. The proposed CNN model achieved a high classification accuracy of 94.56%, a low loss of 0.2162, and an AUC value of 0.93828, demonstrating superior classification capability. The results indicate that the model effectively distinguishes between AD and healthy individuals, outperforming several state-of-the-art approaches. The findings highlight the potential of deep learning-based EEG analysis for AD detection, providing an accessible and cost-effective tool for early diagnosis. The high accuracy of the proposed CNN model suggests that it can assist medical professionals in making well-informed decisions, ultimately improving patient outcomes.

Keywords: Alzheimer's Disease, Electroencephalogram (EEG), Convolutional Neural Network (CNN), Deep Learning (DL), Classification.

التصنيف القائم على التعلم العميق لمرض الزهايمر باستخدام إشارات تخطيط كهربية الدماغ: نهج CNN للكشف المبكر

نجلاء سهيل مزهر، احمد فائق حسين، سفيان منذر صالح

الخلاصة:

مرض الزهايمر اضطراب عصبي تنكسي تدريجي يؤثر بشدة على الوظائف الإدراكية كالذاكرة والانتباه والتفكير، مما يؤثر في نهاية المطاف على الحياة اليومية. يُعدّ الكشف المبكر والدقيق أمرًا بالغ الأهمية للتدخل والإدارة في الوقت المناسب. قد تكون طرق التشخيص التقليدية، بما في ذلك التصوير العصبي والتقييمات الإدراكية، مكلفة وتستغرق وقتًا طويلاً، مما يستلزم بدائل أكثر سهولة وفعالية. تهدف هذه الدراسة إلى تطوير نظام كشف آلي وفعال يعتمد على التعلم العميق ويستخدم إشارات تخطيط كهربية الدماغ (EEG) لتصنيف الأفراد المصابين بمرض الزهايمر والأفراد الأصحاء بدقة. صُمم نموذج شبكة عصبية تلافيفية (CNN) لاستخراج الميزات المهمة من بيانات تخطيط كهربية الدماغ المُعالجة مسبقًا. تتكون البنية من طبقات التفافية مع طبقات تجميع قصوى، بالإضافة إلى استخدام تقنية الإسقاط (Dropout) والطبقات المتصلة بالكامل لتحسين دقة التصنيف. تم تدريب النموذج وتقييمه على مجموعة بيانات تخطيط كهربية الدماغ الشاملة، باستخدام مقاييس الأداء الرئيسية مثل الدقة، والتذكر، والدقة، ودرجة F1. حقق النموذج المقترح دقة تصنيف عالية بلغت ٩٤,٥٦٪، وخسارة منخفضة بلغت ٠,٢١٦٢، وقيمة AUC بلغت ٠,٩٣٨٢٨، مما يثبت قدرته الفائقة على التصنيف. أظهرت

NJES is an **open access Journal** with **ISSN 2521-9154** and **eISSN 2521-9162**

النتائج أن النموذج قادر على التمييز بفعالية بين مرضى الزهايمر والأشخاص الأصحاء، متفوقًا على العديد من الأساليب الحديثة في هذا المجال. تُبرز النتائج إمكانات تحليل تخطيط كهربية الدماغ القائم على التعلم العميق في الكشف عن مرض الزهايمر، مما يوفر أداةً سهلة المنال وفعّالة من حيث التكلفة للتشخيص المبكر. وتشير الدقة العالية لنموذج CNN المقترح إلى أنه يمكن أن يُساعد الأطباء على اتخاذ قرارات مدروسة، مما يُحسّن في نهاية المطاف نتائج المرضى.

1. Introduction

The human brain is one of the most remarkable organs in the body, exerting significant control and influence over our daily lives. Therefore, disorders related to it represent an important area of study in the medical field [1]. Medical issues have emerged as a major concern worldwide, leading to continuous efforts by practitioners and scholars who strive tirelessly for better diagnostic techniques, evaluations, and treatments to save lives and promote health. Among the significant challenges in healthcare is Alzheimer's disease (AD) [2, 3].

AD is an incurable neurodegenerative illness that gradually impairs cognitive and memory abilities, making it crucial to take action in the early stages of the disease, when symptoms are just beginning to manifest [4]. However, the transition from the preclinical to advanced stages of AD can span several decades. Throughout this time, various biomarkers representing underlying pathologies exhibit dynamic changes at specific disease stages [5, 6]. Therefore, prompt and accurate identification of AD is essential for proactively managing patients and driving the future development of therapeutic strategies [7].

The risk factors for AD are multifactorial and include the abnormal accumulation of proteins, such as amyloid-beta and tau, within and around neurons, increasing age, genetic predisposition, and vascular disorders such as hypertension, stroke, and heart disease. Metabolic disorders, including diabetes and obesity, also contribute to the disease's risk profile. Increased awareness and early detection of AD are crucial for implementing preventive strategies and slowing the disease's progression [8]. Currently, no definitive treatment exists to cure AD. Available pharmacological interventions only slow down or mitigate the progression of symptoms, thus improving the quality of life for impacted individuals, caregivers, and their family members [9].

In clinical practice, there is no single test that effectively diagnoses AD. Consequently, several medical modalities are available for the early diagnosis of AD and for monitoring treatment efficacy, including techniques such as neuropsychological tests, brain imaging, electroencephalography (EEG), and cerebrospinal fluid (CSF) analysis, which evaluate cognitive functions and identify secondary causes of cognitive decline. A safer approach involves examining EEG to identify brain abnormalities associated with AD [10]. EEG signals can't capture the activity of an individual neuron; rather, they often reflect the summation of synchronous firing among a group of cortical cells.

EEG has gathered significant scientific interest because of its inexpensive, high temporal resolution, wide availability, non-invasive, and portable nature. It is considered a powerful tool that captures the electrical activity generated by neurons in the brain [11]. The electric currents produced through the cell membranes' depolarization can create distinctive waves identified by scalp electrodes, reflecting neurons' firing activity. The interpretation of EEG signals is typically performed by a trained neurologist, as analyzing these recordings can be quite complex because of the artifacts and noise included in the signal [12]. Furthermore, the diverse neural activities reveal intricate and nonlinear dynamics, necessitating sophisticated analytical techniques. Such techniques are designed to enhance sensitivity and provide deeper insights compared to traditional visual analysis [13, 14]. There is growing evidence that EEG signals may be useful in the differential diagnosis of neural characteristics and cognitive deficits in early AD. AD patients exhibit more decreased synchrony, slowing, and diminished complexity of EEG signals [15].

The EEG is generally divided into four neurophysiological subbands, namely: δ (0.5–4 Hz), θ (4–8 Hz), α (8–12 Hz), β (12–32 Hz). It has been demonstrated, that, local abnormalities in EEG recordings, including loss of complexity, slowing of the rhythms, and altered synchronization between channels, may suggest cerebral degeneration attributable to AD, although not specific to the disease [16]. Different EEG bands can be employed for quantifying the neural activity related to AD. Several studies have already exhibited significant alterations in EEG patterns among AD patients compared to healthy controls. These changes represented increased theta and delta power, diminished alpha and beta power, and disruptions in connectivity measures across multiple brain regions [17, 18].

Many deep AD detection methods have recently emerged with the increasing use of deep learning (DL) techniques. DL is a part of AI that has gained popularity in recent years as a result of developments in Graphics Processing Units (GPUs) in computing. It aims to develop algorithms that are similar to the human brain to extract important features of data that would have been impossible to observe utilizing conventional statistical analysis methods. Classification of EEG signals as AD and non-AD can be performed using ML and DL methods [19]. Machin learning(ML) and DL approaches extract pertinent features from EEG signals, which may encompass measures related to the frequency spectrum of various bands, entropies, statistical moments (such as variance, mean, or skewness), and other metrics that provide valuable insights into the characteristics of the EEG signal. [20, 21]. CNN, one of the most commonly employed DL frameworks, is particularly effective in handling large datasets while

achieving high classification performance [22]. A standard CNN architecture comprises multiple layers, including convolution, pooling, activation, and classification layers. The disadvantages of CNNs include their computationally expensive training, the challenge of verifying their predictions, and their dependence on labeled data [23]. Moreover, one-dimensional CNN (1D CNN) architecture has demonstrated noteworthy efficacy in advancing the accuracy and discriminative capabilities of deep neural network architectures, especially in participating in the detailed examination of AD disease [8].

The major contributions of our paper include:

- Categorize AD into two distinct stages of its progression using EEG signals to improve diagnostic and cognitive assessment methods for AD
- Perform an empirical analysis of AD classification utilizing a CNN architecture combined with the SoftMax classifier based on the EEG database supplied by Dr. Dennis Duke and made publicly accessible by Vicchietti et al. [24].

-Determine the efficacy of these classification algorithms through various assessment measures, encompassing accuracy, recall, precision, and F1-score. Moreover, the Receiver Operator Characteristic (ROC) curve is also plotted.

The rest of this paper is structured as follows: Section 2 presents a review of related works and highlights key advancements in EEG-based AD detection. Section 3 details the proposed CNN methodology, including dataset preprocessing, model architecture, and training parameters. Section 4 discusses the experimental results and comparative analysis with existing approaches. Finally, Section 5 concludes the study, summarizing key findings and outlining potential future research directions.

2. Related Works

In recent years, there has been an increase in the application of DL techniques to classify AD through EEG data. EEG has gathered significant scientific interest because of its cost-effectiveness and non-invasive diagnostic nature. Several researchers have focused on analyzing EEG records from AD patients to follow the changes, such as those related to the complexity of EEG activity.

Ieracitano et al. (2019) conducted a study involving 63 individuals with AD, 63 healthy controls (HC), and 63 individuals with mild cognitive impairment (MCI). They used two-dimensional grayscale images of EEG data as features for a three-class classification model based on a CNN. While their CNN model outperformed traditional ML methods like Multilayer Perceptron (MLP) and SVM, it achieved an accuracy of 83.33%.[16].

Another study by Ieracitano et al. (2020) developed an automatic classification system using EEG signals from AD/MCI/CS subjects. This system relies on 1D CNN, and it directly utilizes time-series signals without any feature extraction. Their proposed architecture has achieved an accuracy of 85.78% for AD/CS classification, which is the highest among the reported results [25].

You et al. (2020) suggested a cascade neural network that utilized the sequential characteristics of EEG and human gait features to categorize AD, HC, and MCI from 35 HC, 35 MCI, and 17 AD patients. 91.70% classification accuracy was attained by this approach [26].

Rad et al. (2021) acquired brain signals from three channels, Pz, Fz, and Pz, of 40 individuals across four conditions: open eyes, closed eyes, recall, and stimulation EEG, to diagnose AD. According to the data, the Pz channel exhibited higher accuracy than the others. Utilizing the characteristics of the Pz channel, the accuracy of LDA was 59.4% and 66.4% in the recall and excitation modes, the accuracy of the Elman neural network was 92.3% and 94.1%, and CNN was 97.5% and 99%, correspondingly [27].

Amini et al. (2021) utilized the time-dependent power spectrum descriptor for the extraction of EEG features of each channel as CNN input. A dataset including the resting-state EEG of 64 AD subjects, 64 HC subjects, and 64 MCI subjects. The CNN method archives an accuracy percentage of 82.3%. While 89.1% of AD cases and 75% of healthy individuals receive a proper diagnosis, only 85% of cases of MCI are accurately identified in-depth [28].

In related work, Huggins et al. (2021), utilized a data collection of resting-state EEG samples from age-matched groups of 52 AD subjects, 52 HA subjects, and 37 MCI subjects. Continuous wavelet transform was employed to convert EEG into a time-frequency graph which was then fed into an AlexNet DL model. The model has an accuracy rating of up to 98.90% for three-class classification [23].

Gkenios et al. (2022) state that EEG signals are used to classify MCI subjects, AD subjects, and CSs. Because of the time-series nature of EEG signals, an RNN classifier is employed for the classification process. The corresponding dataset, comprising 54 subjects, is supplied by the "Greek Association of AD and Related Disorders." It includes 18 CS, 18 patients with MCI, and 18 patients with AD. In the context of AD and CS classification, the resulting performance of the DL architecture is 75% [29].

Kim et al. (2023) propose an automated EEG detection method to distinguish between demented and non-demented subjects using a dataset that includes 230 patients with AD. The evaluation compares CS versus AD versus MCI, yielding classification accuracies of 64.00% for 1D VGG-19, 68.75% for 1D ResNet-18, 67.00% for ResNet-50, and 68.54% for ResNeXt-50 [30].

Nour et al. (2024) introduced a new approach to differentiate AD and HC using EEG. The method combines Deep Ensemble Learning (DEL) and 2-dimensional Convolutional Neural Networks (2D-CNN). The DEL model yielded an average 97.9% accuracy in AD classification because of 5 cross-fold training [10].

A paper published by Kachare et al. (2024) introduced an innovative and effective AD detection model, termed LCADNet, which employs a four-layer convolutional network that achieves an accuracy of 98.50% and an F1-score of 99.68% through hand-crafted EEG signal features [31].

Despite significant advancements in DL for AD classification, existing models continue to face challenges in balancing computational efficiency with high classification accuracy, limited evaluation metrics, and data size. Addressing these limitations, this study introduces an optimized DL framework that enhances feature extraction while maintaining efficiency to improve AD diagnosis using EEG signals. Unlike studies that use a large number of electrodes, this research focused on a subset of electrodes (T4, P3, P4, O1, O2, and Cz) that are known to be related to the brain regions that are most affected by AD. This allows for a more focused and efficient analysis.

3.The Proposed Methodology

The methodology employed in this study is illustrated in Figure 1. It encompasses several key stages: gathering the dataset, preparing the data, dividing it into training and testing sets, training the model, and evaluating its performance. Each stage is critical for ensuring the effectiveness and accuracy of the developed model.

3.1 Dataset

Dr. Dennis Duke and other researchers from Florida State University provided the EEG database utilized in this study [24]. The EEG signals are divided into four groups: A) 12 HC with eyes open by visually fixating, B) 12 HC with eyes closed, C) 80 probable AD patients with eyes open by visually fixating, and D) 80 probable AD patients with eyes closed. The 160 probable AD patients were diagnosed according to the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria, as well as the Diagnostic and Statistical Manual of Mental Disorders-III-R (DSM-III-R) criteria. EEG segments of 8-second duration were recorded at a sampling frequency of 128 Hz using 19 scalp electrodes (Fp1, Fp2, F3, F4, F7, F8, Fz, C3, C4, Cz, P3, P4, Pz, T3, T4, T5, T6, O1, O2), following the international 10-20 system of electrode placement. The EEG signals were band-pass filtered between 0.5 and 30 Hz, and an EEG technician removed movement artifacts [32]. See Figure (2), which shows a sample of EEG signals with AD and HC subjects.

3.2 Preprocessing Stage

Data preprocessing is a critical step that prepares the data for subsequent processes, such as extracting signal features and classification, essential for training the model. Different techniques have been applied during preprocessing to enhance the quality of the captured features. In this study, we utilized EEG signals recorded with eyes closed to reduce artifacts related to visual stimuli and eye movements. We initially loaded EEG datasets, which included recordings from participants with AD and HC. From the complete set of electrodes, we focused on specific EEG electrodes (indices T4, P3, P4, O1, O2, and Cz). This targeted approach enables a detailed analysis of brain regions that are particularly relevant to AD pathology, while reducing the complexity of the data analysis and computational demands, especially in real-time applications. Also, it can reduce the amount of noise that is collected.

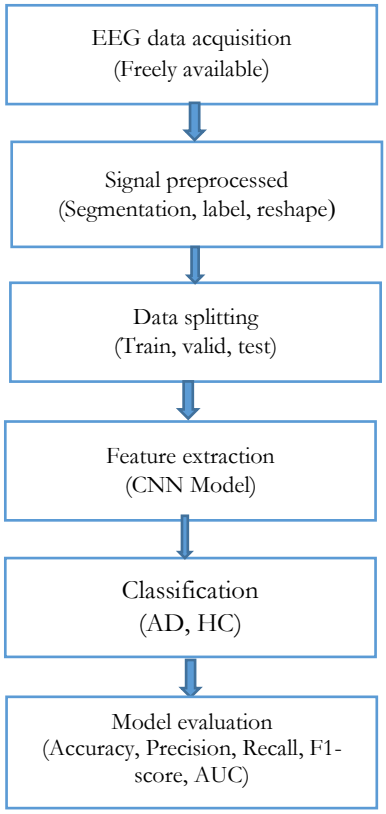

Figure (1): The pipeline of the proposed strategy

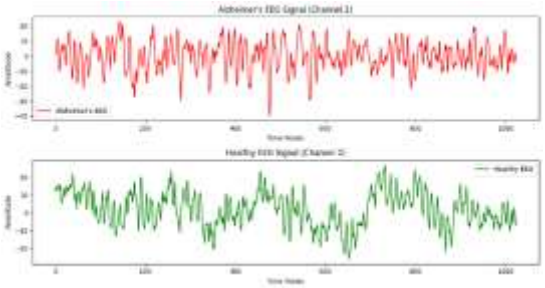

Figure (2): An illustration of an EEG signal with AD and HC subjects.

For a sampling frequency of 128 Hz, each channel contains 1024 samples, equivalent to 8 seconds of EEG data. We limited the data to the first 1000 time points of each chosen channel. This aims to condense the data or emphasize the most critical part of the signal. Subsequently, we divided the EEG data into smaller segments to facilitate better training and to expand the dataset's size. Each channel's data is divided into n_divisions segments along with their corresponding labels. Next, the labels, initially in string format, are converted into integer labels, ranging from 0 to 1, where 1 indicates AD and 0 signifies HC, to enable binary classification. Before inputting the data into a CNN, we must reshape it to align with the expected format by adding a singleton dimension (axis) at the end, and the labels need to be one-hot encoded for correct class representation. Figure (3) illustrates a sample of the EEG signal after segmentation.

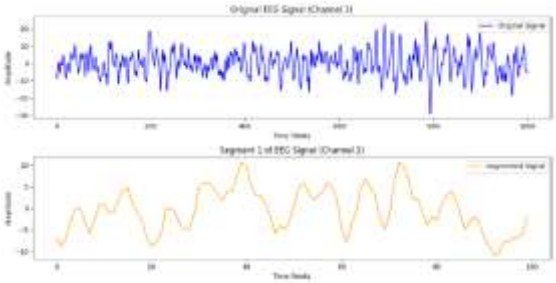

Figure (3): An example of an EEG segment.

3.3 Data Splitting

The 80 AD and 12 HC patients with closed eyes utilized in DL are generally divided into three groups: training, testing, and validation. Each patient has 6 channels, and each channel is divided into 10 divisions. Consequently, there are 800 AD patients and 120 HC patients. The combination of 800 individuals with AD and 120 HC yielded 920 outcomes. The classifier uses a training set to learn what each class looks like by predicting the input data. The validation set serves dual functions: to fine-tune the model and assess its effectiveness, while a test set is used to evaluate the classifier's output after training. This section ensures that the model is trained and evaluated robustly and reliably. A random allocation method was employed, with 80% of the data designated for training and 20% for testing. Furthermore, the training set was further subdivided into two parts, with 80% allocated for actual training and 20% for validation to enhance the model's performance, ensuring equal label distribution through stratification.

4. Proposed CNN Architecture

After pre-processing and segmenting the dataset, the input data is analyzed by a CNN that extracts relevant and distinct features while maintaining computational efficiency to aid in classifying affected areas in AD cases. The CNN model was rigorously trained on a preprocessed signal with dimensions of (Channels, Segment Length, 1). The architecture includes two convolutional layers, with each having a kernel size of (2×2), utilizing the ReLU activation function and 'same' padding to maintain the spatial dimensions of the feature maps. Each convolutional layer is followed by a 2D max-pooling layer with a pool size of (2×1) to reduce the spatial dimensions by 50%, thereby lowering computational complexity while retaining critical features. A flattening layer is then used to convert the two-dimensional feature maps into a one-dimensional vector, preparing the data for fully connected layers. A dropout layer with a rate of 0.25 is included to mitigate overfitting, and a hyperparameter chosen after fine-tuning experiments showed it provided the best balance between model complexity and generalization. A fully connected layer containing 128 neurons collects overarching patterns and insights derived from the generated feature maps. The final layer of the model is a fully connected output layer that employs a sigmoid activation function to categorize the input data into two distinct categories, as shown in the structure of the CNN architecture in Figure 4.

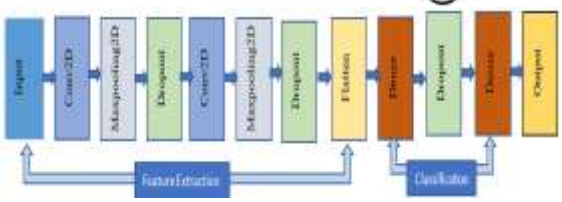

Figure (4): The structure of the CNN architecture.

During training, we used specific parameters, including 25 epochs and a batch size of 32. The Adam optimizer is configured to enhance the model's performance. The main benefits of this approach include the elimination of unnecessary features, reduced computational burden, and improved understanding of the model's behavior due to dense connections. These advantages lead to enhanced accuracy and efficiency in deep learning tasks. The settings parameters were chosen based on the model's performance on the validated data. Tables (1) provide more information on these parameters. The model's output provides detailed performance metrics, including loss, precision, accuracy, recall, and F-score, complemented by graphical visualizations.

Table (1): Training parameters for the CNN

Training parameter	The setup values for parameters
Optimizer	Adam
Learning rate	0.001
Batch-size	32
Loss function	categorical_crossentropy
Epochs	25
The activation function of the output layer	softmax

5. Performance Evaluation Metric

A test set was generated by dividing the original dataset before training the algorithm to evaluate its efficacy [33]. The model's reliability and robustness have been confirmed through the investigation of several evaluation measures. A comprehensive interpretation of these metrics creates the basis for the model's training efficacy. Recent works have highlighted the significance of using the confusion matrix to analyze classification performance, as it offers a comprehensive representation of the relationships and distribution of classified data. The confusion matrix has been one of the main and crucial instruments for evaluating the effectiveness of categorization techniques. It provides substantial insights by an exhaustive investigation of the predictions of each model. The confusion matrix is a unique table that spreads out results based on the classifier model in use and categorizes them into four metrics: true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values. Then those values are calculated using four interesting outcome metrics - precision (PPV), accuracy (ACC), sensitivity (Recall), and F1-score [34]. Moreover, we use the area under the receiver operating characteristic (ROC) curve (AUC) to assess the classifier's performance. The AUC value, which ranges from 0 to

1, signifies the model's effectiveness in accurately categorizing instances. A number of 1 indicates perfect classification capability, but a value of 0.5 proposes that the model's predicted performance is equivalent to random probability [35].

Accuracy: The proportion of accurate predictions to total predictions is the measure of a predictive model's accuracy. In most cases, an accuracy score of 80% or higher is satisfactory, while a value of 90% or higher is outstanding [36].

$$\text{Accuracy}=\frac{\mathbf{TP+TN}}{\mathbf{TP+TN+FP+FN}} \quad (1)$$

Precision is referred to as the percentage of optimistic predictions that are accurate among all optimistic predictions. Precision levels of more than 80% are usually considered acceptable [36].

$$\text{Precision}=\frac{\boldsymbol{TP}}{\boldsymbol{TP+FP}} \quad (2)$$

Recall is an assessment of a model's efficacy in identifying actual positive cases; it is also referred to as sensitivity or the true positive rate. It is the ratio of the expected positive cases to the overall number of positive predictions. Recall values between 70% and 90% are considered acceptable [37].

$$\text{Recall}=\frac{\mathbf{TP}}{\mathbf{TP+FN}} \quad (3)$$

The F1 score is a remarkable statistic since it ensures that every class label is assigned a unique and accurate value, achieving an optimal compromise between recall and precision [38].

$$F1-Score=2*\frac{\boldsymbol{Precision\times Recall}}{\boldsymbol{Precision+Recall}} \quad (4)$$

6. Model Development and Training

The Training and development of the model required substantial computational resources, including high RAM and memory capacity. Most of the experimental work was conducted on Google Colaboratory, leveraging its cloud-based infrastructure. The dataset was uploaded to Google Drive for convenient access and integration into the experiments. Several essential packages were installed to facilitate specific tasks in the model development process, including TensorFlow, Keras, NumPy, Matplotlib, sklearn, and Seaborn. TensorFlow was used to analyze the EEG dataset and implement the neural network models. NumPy for data handling. sklearn for data splitting and performance visualization models. All experiments were performed on a laptop computer featuring the 12th Gen Intel(R) Core (TM) i7-12700H CPU @ 2.30GHz, RAM 16 GB, and 8GB NVIDIA GeForce RTX 3050. The operating system used is Windows 10 64-bit.

7. Experimental Results and Discussion

The CNN model was evaluated and used to test the dataset. Accuracy and loss curves were plotted through the training process for both the testing and training phases, as Figure 5 and Figure 6 clarify. The variations in loss and accuracy values over the training epochs are shown in these graphs. The blue line shows the model's accuracy, which shows the level of accurate learning attained with each epoch. On the other hand, the testing accuracy, shown by the orange line, sheds light on the model's capacity for generalization. The y-axis in both graphs shows the relevant accuracy or loss values, while the x-axis indicates the total number of epochs. The plots visually depict the model's performance during the training and validation stages. The plots show that at epoch 23, the model reached its maximum training accuracy of 0.9769 and validation accuracy of 0.9584 suggesting that the model has effectively learned relevant features from the dataset. The stability of the loss values further confirms this, as the training loss reached 0.0358, while validation loss remained low at 0.1239 at epoch 25. The fact that there isn't much difference throughout the training and validation accuracy levels suggests that the model is operating effectively in this situation and has minimal overfitting. This is further supported by the test set results, which demonstrate a high accuracy of 0.9456 and a low loss of 0.2162, closely mirroring the validation performance. The fact that there isn't much difference throughout the training and testing accuracy levels suggests that the model operates effectively in this situation and is likely to perform well on unseen data. The time taken for each ninth epoch is between 4 seconds in the beginning, then stabilizes at 1-2 seconds for most epochs. The first epoch only experienced a 4-second delay due to the short loading and processing times, and then the time took longer.

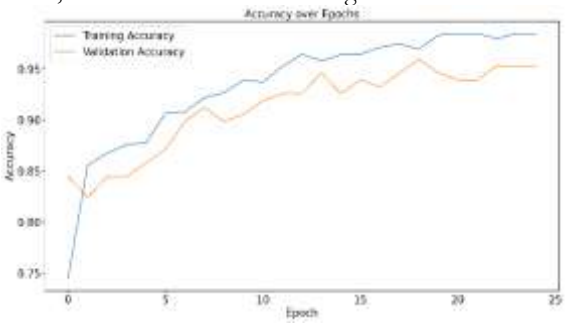

Figure (5): Accuracy curves of the CNN model.

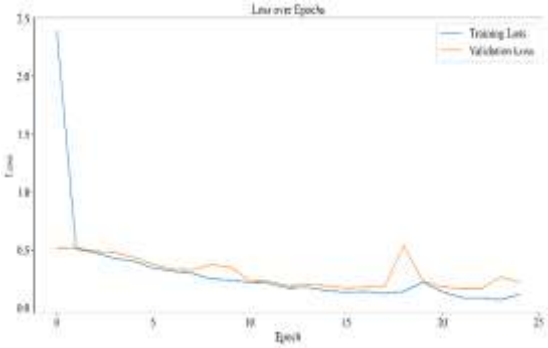

Figure (6): loss curves of CNN model.

After analyzing the plot graphs, we conducted the model evaluation using a confusion matrix, as illustrated in Figure 7. The confusion matrix represents a comprehensive evaluation of the model's classification efficacy, which functions as an essential

tool for enhancing machine learning models. It offers an intensive comprehension of the model's predicting accuracy by illustrating the correlations between actual and anticipated classes. This allows a more refined assessment of the performance of the model in accurately identifying data. The configuration of the confusion matrix is of a 2 x 2 dimensionality, corresponding to the total number of classes encompassed within the study.

Figure (7): Confusion matrix of CNN.

It can also be spotted that the model correctly predicted the model to the Alzheimer's class was 157 patients while misclassifying 3 patients as belonging to the healthy class. Similarly, the model correctly classified 17 patients in the healthy class, with 7 patients misclassified as Alzheimer's. Moreover, the ROC curve is also plotted for the CNN models for the performance assessment using an alternative metric of AUC as shown in Figure 8. However, the CNN model provided the highest AUC value of 0.93828, a better classification capability.

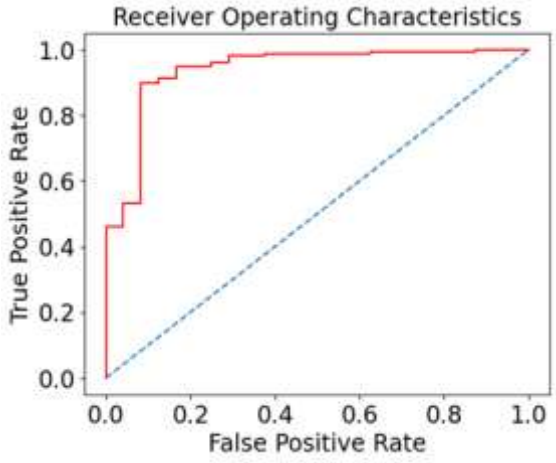

Figure (8): ROC curve of the CNN model**.**

The performance of the CNN model, regarding some key metrics such as accuracy, recall, precision, and F1-score, has been summarized in Table 2. These metrics provide an overall evaluation of the efficiency of the model classification for both classes. The range values are between 0 and 1, and a higher value indicates a lesser error.

Table (2): The CNN model's performance.

Class	Precision	Recall	F1 score	Accuracy
0	0.85	0.71	0.77	0.9456
1	0.96	0.98	0.97	

By comparing other studies with our proposed model, this study notably differs from previous research in its approach to data size, focusing on a subset of electrodes, and providing comprehensive evaluation across several performance metrics. Table 3 shows the previous methods used to categorize AD using EEG datasets compared to the suggested approach.

Table (3): shows previous methods in comparison to the suggested approach.

Author	Data	Method	Accuracy
Ieracitano et al. [16].	63 AD, 63 HC, 63 MCI	CNN	83.33%.
You et al. [26].	35 HC, 35 MCI, 17 AD	Cascade Neural Network	91.70%
Amini et al. [28].	64 AD, 64 HC, 64 MCI	CNN	82.30%
Gkenos et al. [29].	18 AD, 18 MCI, 18 CS	RNN	75.00%
Proposed method	80 AD, 12 HC	CNN	94.56%

The outcomes of our study provide empirical evidence supporting the ability of the CNN model we developed to extract intricate features. This model exhibits a remarkable level of accuracy, reaching 94.5%, exceeding the performance of other intricate models. This achievement is especially noteworthy due to its successful application on a limited clinical dataset. These findings demonstrate that the performance of the model, whether good or bad, is based on variables such as data conditions, the number of layers, the ratio of data separation, input quantity, and data balance.

While the results are promising, this study is subject to certain limitations including the dataset was relatively small, comprising EEG recordings from 80 AD and 12 CN participants. The system needs better patient data variation and simpler models to expand into more health conditions and be faster to run.

8. Conclusions

AD is a neurological disorder that typically manifests in later life, leading to significant impairments such as memory loss and confusion. Early detection is essential to mitigate its effects and improve patient outcomes. This study explored a CNN model coupled with an Adam optimizer, designed for precise identification and categorization of AD using EEG datasets. CNN model achieved superior performance with an accuracy score of 0.9456. The findings highlight the ability of CNNs to extract meaningful spatial and structural features from EEG data, offering a robust method for automated AD diagnosis.

The proposed method offers several practical advantages, including real-time classification capabilities, reduced computational complexity, and improved diagnostic accuracy. By leveraging EEG signals, this approach provides a cost-effective and

non-invasive alternative to traditional AD diagnostic techniques, potentially reducing the reliance on expensive imaging techniques and enabling timely interventions.

Despite the promising results, this study has some limitations. The dataset size remains relatively limited, which may affect the generalizability of the findings. Additionally, EEG signals are inherently noisy and subject to variability among individuals, which may introduce challenges in model performance across diverse populations. Lastly, the model's reliance on supervised learning requires large amounts of labeled data, which can be labor-intensive to obtain.

To further enhance the effectiveness of EEG-based AD classification, future work could explore using real-world data, expanding the dataset, integrating additional data modalities, and developing multi-class classification models to distinguish between AD, Mild Cognitive impairment (MCI), and HC, facilitating more precise disease staging.

9. References

[1] H. Jia and H. Lao, "Deep learning and multimodal feature fusion for the aided diagnosis of Alzheimer's disease," Neural Comput. Appl., vol. 34, no. 22, pp. 19585-19598, 2022. https://doi.org/10.1007/s00521-022-07501-0

[2] S. A. Ajagbe, K. A. Amuda, M. A. Oladipupo, F. A. Oluwaseyi, and K. I. Okesola, "Multi-classification of Alzheimer disease on magnetic resonance images (MRI) using deep convolutional neural network (DCNN) approaches," Int. J. Adv. Comput. Res., vol. 11, no. 53, p. 51, 2021. https://doi.org/10.19101/IJACR.2021.1152001

[3] B. Jesus Jr, R. Cassani, W. J. McGeown, M. Cecchi, K. Fadem, and T. H. Falk, "Multimodal prediction of Alzheimer's disease severity level based on resting-state EEG and structural MRI," Front. Hum. Neurosci., vol. 15, p. 700627, 2021. https://doi.org/10.3389/fnhum.2021.700627

[4] C.-N. Jiao, Y.-L. Gao, D.-H. Ge, J. Shang, and J.-X. Liu, "Multi-modal imaging genetics data fusion by deep auto-encoder and self-representation network for Alzheimer's disease diagnosis and biomarkers extraction," Eng. Appl. Artif. Intell., vol. 130, p. 107782, 2024. https://doi.org/10.1016/j.engappai.2023.107782

[5] M. A. Scelsi, R. R. Khan, M. Lorenzi, L. Christopher, M. D. Greicius, J. M. Schott, S. Ourselin, and A. Altmann, "Genetic study of multimodal imaging Alzheimer's disease progression score implicates novel loci," Brain, vol. 141, no. 7, pp. 2167-2180, 2018. https://doi.org/10.1093/brain/awy141

[6] A. Balasundaram, S. Srinivasan, A. Prasad, J. Malik, and A. Kumar, "Hippocampus segmentation-based Alzheimer's disease diagnosis and classification of MRI images," Arab. J. Sci. Eng., vol. 48, no. 8, pp. 10249-10265, 2023. https://doi.org/10.1007/s13369-022-07538-2

[7] D. Cheng and M. Liu, "CNNs based multi-modality classification for AD diagnosis," in Proc. 10th Int. Congr. Image Signal Process., Biomed. Eng. Inform. (CISP-BMEI), 2017, pp. 1-5. https://doi.org/10.1109/CISP-BMEI.2017.8302281

[8] A. El-Assy, H. M. Amer, H. Ibrahim, and M. Mohamed, "A novel CNN architecture for accurate early detection and classification of Alzheimer's disease using MRI data," Sci. Rep., vol. 14, no. 1, p. 3463, 2024. https://doi.org/10.1038/s41598-024-53733-6

[9] C. Ieracitano, N. Mammone, A. Hussain, and F. C. Morabito, "A novel multi-modal machine learning based approach for automatic classification of EEG recordings in dementia," Neural Netw., vol. 123, pp. 176-190, 2020. https://doi.org/10.1016/j.neunet.2019.12.006

[10] M. Nour, U. Senturk, and K. Polat, "A novel hybrid model in the diagnosis and classification of Alzheimer's disease using EEG signals: Deep ensemble learning (DEL) approach," Biomed. Signal Process. Control, vol. 89, p. 105751, 2024. https://doi.org/10.1016/j.bspc.2023.105751

[11] N. Kulkarni and V. Bairagi, "EEG-based diagnosis of Alzheimer disease: a review and novel approaches for feature extraction and classification techniques," 2018.

[12] A. Horvath, A. Szucs, G. Csukly, A. Sakovics, G. Stefanics, and A. Kamondi, "EEG and ERP biomarkers of Alzheimer's disease: a critical review," Front. Biosci. (Landmark Ed.), vol. 23, pp. 183-220, 2018. https://doi.org/10.2741/4587

[13] M. Breakspear, "Dynamic models of large-scale brain activity," Nat. Neurosci., vol. 20, no. 3, pp. 340-352, 2017. https://doi.org/10.1038/nn.4497

[14] P. M. Rossini, R. Di Iorio, F. Vecchio, M. Anfossi, C. Babiloni, M. Bozzali, A. C. Bruni, S. F. Cappa, J. Escudero, and F. J. Fraga, "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., vol. 131, no. 6, pp. 1287-1310, 2020. https://doi.org/10.1016/j.clinph.2020.03.003

[15] D. V. Puri, S. L. Nalbalwar, and P. P. Ingle, "EEG-Based Systematic Explainable Alzheimer's Disease and Mild Cognitive Impairment Identification Using Novel Rational Dyadic Biorthogonal Wavelet Filter Banks," Circ. Syst. Signal Process., vol. 43, no. 3, pp. 1792-1822, 2024. https://doi.org/10.1007/s00034-023-02540-x

[16] C. Ieracitano, N. Mammone, A. Bramanti, A. Hussain, and F. C. Morabito, "A convolutional neural network approach for classification of dementia stages based on 2D-spectral representation of EEG recordings," Neurocomputing, vol. 323, pp. 96-107, 2019. https://doi.org/10.1016/j.neucom.2018.09.071

[17] H. Azami, C. Zrenner, H. Brooks, R. Zomorrodi, D. M. Blumberger, C. E. Fischer, A. Flint, N. Herrmann, S. Kumar, and K. Lanctôt, "Beta to theta power ratio in EEG periodic components as a potential biomarker in mild cognitive impairment and Alzheimer's dementia," Alzheimer's Res. Ther., vol. 15, no. 1, p. 133, 2023. https://doi.org/10.1186/s13195-023-01280-z

[18] D. Kim and K. Kim, "Detection of early stage Alzheimer's disease using EEG relative power with deep neural network," in Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), 2018, pp. 352-355. https://doi.org/10.1109/EMBC.2018.8512231

[19] P. Khan, M. F. Kader, S. R. Islam, A. B. Rahman, M. S. Kamal, M. U. Toha, and K.-S. Kwak, "Machine learning and deep learning approaches for brain disease diagnosis: principles and recent advances," IEEE Access, vol. 9, pp. 37622-37655, 2021. https://doi.org/10.1109/ACCESS.2021.3062484

[20] M. S. Safi and S. M. M. Safi, "Early detection of Alzheimer's disease from EEG signals using Hjorth parameters," Biomed. Signal Process. Control, vol. 65, p. 102338, 2021. https://doi.org/10.1016/j.bspc.2020.102338

[21] S. Dodge and L. Karam, "A study and comparison of human and deep learning recognition performance under visual distortions," in Proc. 26th Int. Conf. Comput. Commun. Netw. (ICCCN), 2017, pp. 1-7. https://doi.org/10.1109/ICCCN.2017.8038465

[22] S. Murugan, C. Venkatesan, M. Sumithra, X.-Z. Gao, B. Elakkiya, M. Akila, and S. Manoharan, "DEMNET: A deep learning model for early diagnosis of Alzheimer diseases and dementia from MR images," IEEE Access, vol. 9, pp. 90319-90329, 2021. https://doi.org/10.1109/ACCESS.2021.3090474

[23] C. J. Huggins, J. Escudero, M. A. Parra, B. Scally, R. Anghinah, A. V. L. De Araújo, L. F. Basile, and D. Abasolo, "Deep learning of resting-state electroencephalogram signals for three-class classification of Alzheimer's disease, mild cognitive impairment and healthy ageing," J. Neural Eng., vol. 18, no. 4, p. 046087, 2021. https://doi.org/10.1088/1741-2552/ac05d8

[24] M. L. Vicchietti, F. M. Ramos, L. E. Betting, and A. S. Campanharo, "Computational methods of EEG signals analysis for Alzheimer's disease classification," Sci. Rep., vol. 13, no. 1, p. 8184, 2023. https://doi.org/10.1038/s41598-023-32664-8

[25] C. Ieracitano, N. Mammone, A. Hussain, and F. C. Morabito, "A Convolutional Neural Network based self-learning approach for classifying neurodegenerative states from EEG signals in dementia," in Proc. Int. Joint Conf. Neural Netw. (IJCNN), 2020, pp. 1-8. https://doi.org/10.1109/IJCNN48605.2020.9207167

[26] Z. You, R. Zeng, X. Lan, H. Ren, Z. You, X. Shi, S. Zhao, Y. Guo, X. Jiang, and X. Hu, "Alzheimer's disease classification with a cascade neural network," Front. Public Health, vol. 8, p. 584387, 2020. https://doi.org/10.3389/fpubh.2020.584387

[27] E. M. Rad, M. Azarnoosh, M. Ghoshuni, and M. M. Khalilzadeh, "Diagnosis of mild Alzheimer's disease by EEG and ERP signals using linear and nonlinear classifiers," Biomed. Signal Process. Control, vol. 70, p. 103049, 2021. https://doi.org/10.1016/j.bspc.2021.103049

[28] M. Amini, M. M. Pedram, A. Moradi, and M. Ouchani, "Diagnosis of Alzheimer's Disease by Time-Dependent Power Spectrum Descriptors and Convolutional Neural Network Using EEG Signal," Comput. Math. Methods Med., vol. 2021, no. 1, p. 5511922, 2021. https://doi.org/10.1155/2021/5511922

[29] G. Gkenios, K. Latsiou, K. Diamantaras, I. Chouvarda, and M. Tsolaki, "Diagnosis of Alzheimer's disease and mild cognitive impairment using EEG and recurrent neural networks," in Proc. 44th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), 2022, pp. 3179-3182. https://doi.org/10.1109/EMBC48229.2022.9871302

[30] M.-j. Kim, Y. C. Youn, and J. Paik, "Deep learning-based EEG analysis to classify normal, mild cognitive impairment, and dementia: Algorithms and dataset," NeuroImage, vol. 272, p. 120054, 2023. https://doi.org/10.1016/j.neuroimage.2023.120054

[31] P. Kachare, D. Puri, S. B. Sangle, I. Al-Shourbaji, A. Jabbari, R. Kirner, A. Alameen, H. Migdady, and L. Abualigah, "LCADNet: a novel light CNN architecture for EEG-based Alzheimer disease detection," Phys. Eng. Sci. Med., pp. 1-14, 2024. https://doi.org/10.1007/s13246-024-01425-w

[32] W. S. Pritchard, D. W. Duke, K. L. Coburn, N. C. Moore, K. A. Tucker, M. W. Jann, and R. M. Hostetler, "EEG-based, neural-net predictive classification of Alzheimer's disease versus control subjects is augmented by non-linear EEG measures," Electroencephalogr. Clin. Neurophysiol., vol. 91, no. 2, pp. 118-130, 1994. https://doi.org/10.1016/0013-4694(94)90033-7

[33] M. A. Al-Hashem, A. M. Alqudah, and Q. Qananwah, "Performance evaluation of different machine learning classification algorithms for disease diagnosis," Int. J. E-Health Med. Commun. (IJEHMC), vol. 12, no. 6, pp. 1-28, 2021. https://doi.org/10.4018/IJEHMC.20211101.oa5

[34] R. A. Hazarika, D. Kandar, and A. K. Maji, "An experimental analysis of different deep learning based models for Alzheimer's disease classification using brain magnetic resonance images," J. King Saud Univ.-Comput. Inf. Sci., vol. 34, no. 10, pp. 8576-8598, 2022. https://doi.org/10.1016/j.jksuci.2021.09.003

[35] X. Xu, J. Li, Z. Zhu, L. Zhao, H. Wang, C. Song, Y. Chen, Q. Zhao, J. Yang, and Y. Pei, "A Comprehensive Review on Synergy of Multi-Modal Data and AI Technologies in Medical Diagnosis," Bioengineering, vol. 11, no. 3, p. 219, 2024. https://doi.org/10.3390/bioengineering11030219

[36] S.-K. Kim, H. Kim, S. H. Kim, J. B. Kim, and L. Kim, "Electroencephalography-based classification of Alzheimer's disease spectrum during computer-based cognitive testing," Sci. Rep., vol. 14, no. 1, p. 5252, 2024. https://doi.org/10.1038/s41598-024-55656-8

[37] O. Caelen, "A Bayesian interpretation of the confusion matrix," Ann. Math. Artif. Intell., vol. 81, no. 3, pp. 429-450, 2017. https://doi.org/10.1007/s10472-017-9564-8

[38] M. Aljalal, S. A. Aldosari, M. Molinas, and F. A. Alturki, "Selecting EEG channels and features using multi-objective optimization for accurate MCI detection: validation using leave-one-subject-out strategy," Sci. Rep., vol. 14, no. 1, p. 12483, 2024. https://doi.org/10.1038/s41598-024-63180-y