Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 45 No. 1 (2024)

Cascade Wavelet Transform and Optimal Modified Deep Convolutional Neural Network for Alzheimer's Disease Detection in EEG

[1]Rashmi R. Nath, [2]Dr. S. Prabhu

[1]Research Scholar, PG and Research Department, Department of Computer Science, Park's CollegeTirupur, Tamilnadu , India

[2]Research Supervisor, Department of Computer Science, Government Arts & Science College, Thittamalai,Nambiyur, Tamilnadu, India

Abstract: Alzheimer's disease has become the most widespread neurological brain ailment and maybe diagnosed using a variety of medical approaches. EEG signals are frequently used to diagnose cognitive issues,especially if there is a discrepancy in the diagnosis following the first clinical investigations. Nonetheless, there is proof that the EEG may accurately detect Alzheimer's illness. The EEG diagnosis of Alzheimer's disease (AD)is growing in frequency. A novel resting-state EEG classification method for Alzheimer's disease (AD), moderatecognitive impairment (MCI), and healthy control (HC) is presented in this work. Early diagnosis of mild but noticeable cognitive impairments that do not significantly affect day-to-day functioning may lower the death andmorbidity rates associated with prodromal Alzheimer's condition. During such studies, disturbances and interference from the EEG dataset are eliminated using a band-pass elliptic digital filter (BEF). To get the properties of the EEG signal, the filtered data was separated into frequency bands using the Cascade wavelet transform (CWT) technique. Then, by including a variety of signal properties into feature vectors, the CWT technique was utilized to improve diagnostic performance. The improved EEG was categorized using an Optimal modified Deep Convolutional Neural Network (OMDCNN) once the necessary dataset was created. Here, the ideal MCNN hyperparameter is determined using the modified clonal selection algorithm (MCSA), which mightincrease the classifier's accuracy. Finally, the sensitiveness, accuracy, quality of diagnosis, and the area beneath the receiver operating characteristic (ROC) curve were additionally computed in order to compare and evaluate the performance of the various recommended approaches.

Keywords: *recommended, MCNN, approaches, categorized, diagnostic, frequency*

Introduction

Dementia is defined as a diminished capacity for memory, cognition, or making decisions which difficulties with daily responsibilities. It makes no reference to a specific ailment. Alzheimer's disease is the most common type of dementia. Alzheimer does not constitute a normal component of aging, despite the fact that it mostly affects the elderly. Dementia is predicted to impact 7.4% of persons aged 60 and over in India. Dementia affects 8.8 million Indians above the age of 60. According to the Dementia in India 2020 report36, there will be 5.3 million dementia cases among Indians over 60 in 2020, and that amount is predicted to climb to 14 million by 2050. [1].Medical evaluation of Alzheimer's disease is difficult, especially in its initial phases when symptoms are frequently dismissed as typical aging symptoms. Furthermore, there are some early-stage symptoms that are shared by Alzheimer's disease along with other illnesses, including dementia with Lewy body structures, dementiaof the frontal lobe, and vascular dementia.

Extensive testing is necessary to rule out any other possible causes of Alzheimer's disease. Comprehensive cognitive exams, neurological examinations, blood testing, brain imaging methods, and neurological fluidinvestigations are among the therapies provided [2]. Patients may benefit from better care if dementia can be identified using less invasive and costly approaches. A substantial amount of research has been conducted in recent years on the potential utility of electroencephalography (EEG) toward the detection of dementia disorders,such as Alzheimer's disease [3-

5]. Electroencephalography (EEG) is a harmless, low-cost, possibly portable technology for measuring time to the millisecond. The primary goal of this study was to identify Alzheimer's disease through the contrast of EEG recordings from AD patients to those from control participants, or healthy people [3-5]. The electroencephalogram (EEG) is one of the most often utilized diagnostic tools in modern medicine for identifying brain abnormalities. Simply explained, EEG involves an observation of the electrical reactions in the brain.

Adhesive electrodes placed on the scalp in certain locations are used to measure voltage fluctuations in different brain regions. The application of EEG signals has been advantageous due to the non-invasiveness of the technique. Additionally, EEG offers a high temporal resolution, often measured in milliseconds. The radiation risks associated with conventional diagnostic techniques like MRIs [6] and CT scans [7], as well as the invasive natureof therapeutic therapies, have led to an increase in the use of EEG. Alzheimer's is a long-term neurological condition. It's possible that this type of dementia affects seniors the most frequently. Among those over 65, it affects one in fifteen of them. Among those over 65, it affects one in fifteen of them. Alzheimer's disease is thoughtto be brought on by a dramatic shrinking of the hippocampal and cerebral cortex regions of the brain, while the exact cause of the illness is unknown. To reduce or even reverse the course of Alzheimer's disease, early detectionis essential. It is possible to comprehend the kinds of traits used in Alzheimer's disease research [3-5].

Duration domain, frequency realm, and duration-frequency domain characteristics are the most common forms of traits or biomarkers. Neural network models or classical machine learning are two more methods for categorizingraw EEG data. Time-frequency based methods like DWT (Discrete Wavelet Transform) are more appropriate forfeature extraction because EEG is a non-stationary signal. Acharya et al. classified Alzheimer's patients with an ccuracy rate of 91.70% by using estimated entropy and sample entropy extracted from EEG as input data for a support vector machine (SVM) classifier [8]. Tavares et al. achieved a typical classification success rate of 95.60%by using the power spectral density (PSD) of several EEG frequency bands as input parameters for eight differentmachine learning techniques [9].

Machine learning algorithms are utilized to scan EEG and diagnose AD, minimizing the span of time consumed monitoring and dealing with humans. As a result, the diagnostic procedure produces healthier and unbiased outcomes. Traditional machine learning algorithms usually need complicated feature engineering (FE) in order to attain acceptable performance, implying that researchers must devote significant time and effort in studying the original EEG and selecting appropriate features. Deep learning is a sort of machine learning that automates FE utilizing neural networks and hidden layers, saving researchers a significant amount of time and money [10]. Furthermore, artificially selected features risk losing vital details from EEG data, but deep learning can make better use of EEG data and boost program robustness. Through a system that uses deep learning for diagnosing Alzheimer's disease using EEG has been a promising topic in recent years. After gathering these features, the dataset could be utilized for training a classifier to figure out if the signal is from an AD patient as well as ordinary.An OMDCNN algorithm was employed in the suggested study. The primary inspiration for our effort stems fromprevious studies undertaken by AlSharabi et al [6]. The following is the work's key contribution:

In the current work, a BEF was employed to remove disturbance along with irregularities from an EEG dataset.

- The CWT approach was then used to divide the signal that was filtered into bands of frequency in order to retrieve EEG signal characteristics.

- Then, several signal properties were included into the CWT approach to build feature vectors and increase diagnostic performance.

- The OMDCNN was utilized to categorize the enhanced EEG after building the required dataset.

The following are the remaining portions of this document: The second part introduces relevant studies on the same issue. The third part explains the EEG data, data enhancement approaches, and deep learning model architecture employed in this work. The fifth section presents the study's findings and commentary. The sixth partpresents the findings, as well as recommendations for further research.

Literature Review

Khare and Acharya [11] introduced Adazd-Net, an adaptable and descriptive structure that enables the automated detection of AZD from EEG data. The suggested flexible adaptive analytic wavelet transformations automatically adjust to fluctuations in the EEG. This paper looks at both the most discriminating channel and the optimal feature count required for system performance. On the other hand, manual evaluation is laborious, biased, and dependenton the expertise of the individuals doing the evaluation.

Musaeus et al., [12] wanted to see whether ear-EEG-measured epileptiform discharges are more common in Alzheimer's patients than in age- and gender-matched healthy controls. Modifications throughout time and the association with cognitive function have been investigated to better understand the connection between epileptiform emissions and AD. It is unknown; however, which characteristics and channels produce recognizeddata for detection of AD.

Puri et al., [13] introduced a novel approach for separating AD and NC EEG signals into subbands (SBs) using low-complexity orthogonal wavelet banks of filters having features that the disappear (LCOWFBs-v). To reducethe computational expense of the original nonsensical wavelet filter banks (FBs), a bigger design technique is offered. EEG SBs were used to calculate Higuchi's fractal dimension (HFD) and Katz's fractal dimension (KFD). The significance of these recovered traits was assessed using the Kruskal-Wallis test. More study is needed, however, to discover when epileptiform emissions emerge everywhere as the illness progresses.

Imani [14] relied on Temporal sequences are analyzed by bidirectional long short-term memory (BiLSTM) networks, while EEG data gathered by several channels located in various brain regions is analyzed by convolutional neural networks (CNN). The temporal and geographical data is then integrated using a fully connected neural network. In addition, this study uses autoencoder networks for data augmentation and an entropymetric for channel selection in order to improve diagnostic accuracy. It is uncertain how many characteristics areneeded to get the optimal system performance, though.

Jia et al., [15] described a three-phase technique of decomposition-recombination to enhance the number of trialsin the training set. Initially, multivariate empirical decomposition of modes is applied to break down the initial signals in the training set into several intrinsic mode functions. Another arbitrary combination of those functionsof intrinsic mode from several trials comes next. Furthermore, the reconstructed function of the intrinsic mode is mixed as simulated trials to train the model. On a limited dataset, we tested the decomposition- recombination system utilizing functional connectivity matrix from each person as inputs. This benefit, however,is only evident for a short number of simulation trials utilizing the MCI datasets before the median efficiency begins to decrease.

Cejnek et al., [16] recommended an innovative technique for diagnosing Alzheimer's disease with the potential for moderate cognitive impairment using EEG data. The suggested technique uses the EEG's degree of signal novelty as a characteristic to classify EEG records. The EEG signal adaptive filtering parameters are used to calculate novelty. Under prediction settings, a linear neuron which is acclimated to a descent in gradient was employed as the filter. Next, using the acquired characteristic (novelty measure) categorization, Alzheimer's disease is diagnosed. The suggested method was cross validated using an EEG dataset that included 102 controls,seven individuals with significant cognitive impairment (MCI), and 59 patients with Alzheimer's disease. Cross-validation yielded findings with 89.51% sensitivity and 90.73% specificity.

Dogan et al., [17] created a unique directed network that collects local texture information for the purpose of automating the identification of AD using EEG data. The macroscopic connectome, a network of neural connections linking anatomo-functional brain regions of the monkey brain involved in the identification of visualobjects and physical response, was used to create the suggested graph. The evaluation of this PBP-based model was conducted with an openly accessible dataset of AD EEG recordings. But qualities are difficult to understandand need a great deal of effort to cultivate.

Chiang et al., [18] used fuzzy Petri net and associative methods to build a framework for an unbiased and successful Alzheimer's disease diagnosis. Alzheimer's disease prediction factors are developed using the variations in EEG

patterns between healthy individuals and Alzheimer sufferers. This might provide medical professionals a point of reference for a prompt diagnosis and allow for early intervention to prevent the illness from growing worse. On the other hand, training learning models requires a large number of trials.

Jiao et al., [19] suggested the use of a Key EEG indication to distinguish between individuals suffering from AD, MCI, and other neurological dementias. The indication was tested on people in good health. We also attempted to investigate the relationship between the identified CSF biomarkers and the EEG biomarker and individual decline in cognition. In addition, a machine learning method was developed and proven to assess patients' cognitive performance (MMSE and MoCA), age at onset of disease (ADO), and course of illness (COD). Its usein accurately diagnosing and evaluating AD and its preliminary stage, amnestic moderate cognitive impairment (MCI), is still up for debate.

Miltiadous et al., [20] developed an innovative technique that makes use of a Dual-Input Convolution Encoder Network (DICE-net) to categorize AD EEG data. Procedures: Information from age-matched healthy persons (CN) and 36 AD and 23 FTD were utilized. Following blurring, spectrum power and synchronization properties were obtained and added to DICE-net, which consists of convolution layers and a feed-forward transformer encoding. The principal results are: The results show that DICE-net performed better than other baseline models during Leave-One-Subject-Out validation and showed good generalization performance, with a prediction rate of83.28% regarding the AD-CN problem. This would be inaccurate, though, as there is no geographic relationship between any of the channels. Therefore, it is advised to use a depth-wise separable convolution.

I**nference:** In an attempt to address the issue of AD detection, a number of automatic approaches utilizing ML architectures have been put forth recently. However, the generalizability of these approaches' results has been hampered by their small sample sizes, dearth of published datasets, or inadequate validation techniques appropriate for selected datasets. Additionally, very few studies have used one of the most recent developments in Deep Learning—the Transformer architecture—in EEG-based dementia detection research. That being said, more accurate and effective deep learning diagnostic methods are required in order to effectively capitalize on the wealth of information provided by EEG recordings. Performance outcomes from such diagnostic tools should bereproducible and well-verified (by promoting the openness of the databases exploited) in order to assure their trustworthiness and usefulness in clinical practice.

Proposed Methodology

In the current work, contamination along with disturbance seen in the EEG dataset were substantially reduced by using the band-pass elliptic digital filter. Important characteristics might then be extracted from the EEG data by applying the CWT approach to separate the filtered signal into discrete frequency bands. After that, a lot of signal characteristics were added to the CWT methodology to build feature vectors and enhance the efficacy of diagnosis. After the required dataset was created, the MDCNN was used to classify the improved EEG.

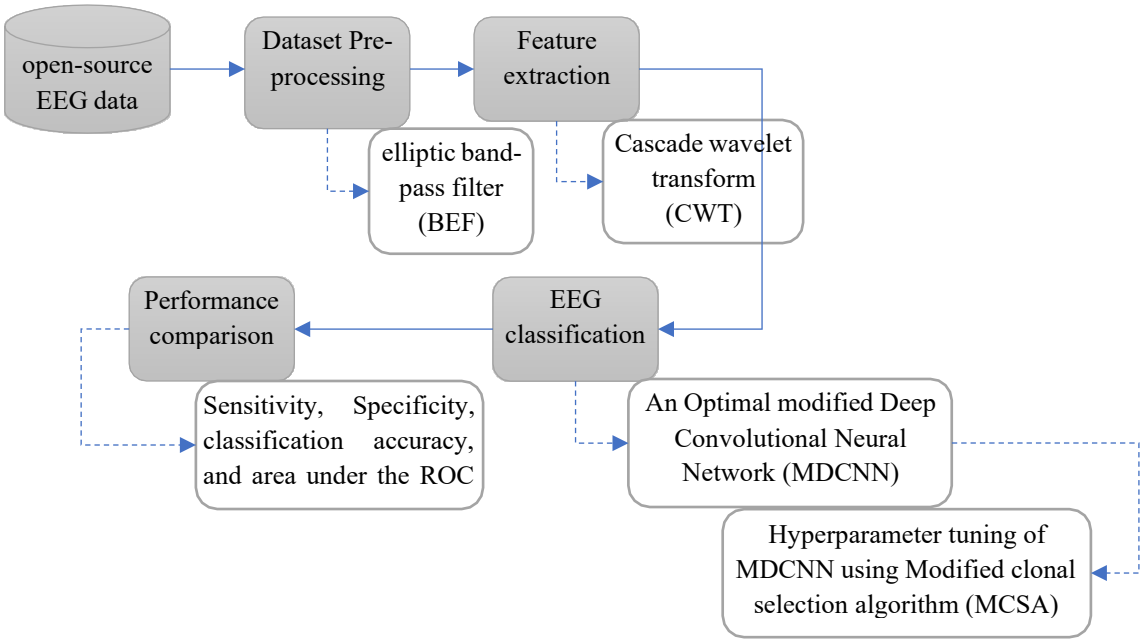

Fig.1. Blok Diagram of Proposed Methodology

Dataset Description

Datasets from the Behavioral and Cognitive Neurology Unit of the Department of Neurology including the Referral Center for Cognitive Disorders at the Hospital das Clinics in Sao Paulo, Brazil, were used in this study to gather information on AD patients and control volunteers. The Mini-Mental State Examination (MMSE) and the Brazilian version of the Clinical Dementia Rating - CDR scale was used by qualified neurologists to recognizeall AD individuals and Control volunteers and gather the necessary datasets [21]. 86 participants were split into three groups and provided the multi-channel EEG datasets. With a median age of 66.89 years and 8.18 StD, the first batch of 35 Control subjects (CS) is made up of 18 men and 19 females. Based on individual interviews, therequirements for being included for those in the cognitive normal subgroup included an MMSE of at least 25, a CDR score of 0, a median MMSE of 28, and a mean deviation of 2.2. Moreover, there was no indication of functional cognitive decline prior to recording. The second patient group met the NINCDS-ADRDA [22] along with DSM-IVTR [23] parameters. It consisted of 31 mild-AD patients (a total of 12 male and 19 female), havingthe median age of 75.23 years as well as 5.55 StDs. The criteria for inclusion that followed significantly reducedthe number of individuals with mild Alzheimer's disease: 0.5 CDR 1 as well as MMSE 24, which produced a median MMSE of 19.48 and a standard deviation of 3.16. Twenty-two individuals with intermediate Alzheimer'sdisease—seven men and fifteen women—make up the third group. Their average age is 73.77 years, and they have 10.16 StDs.

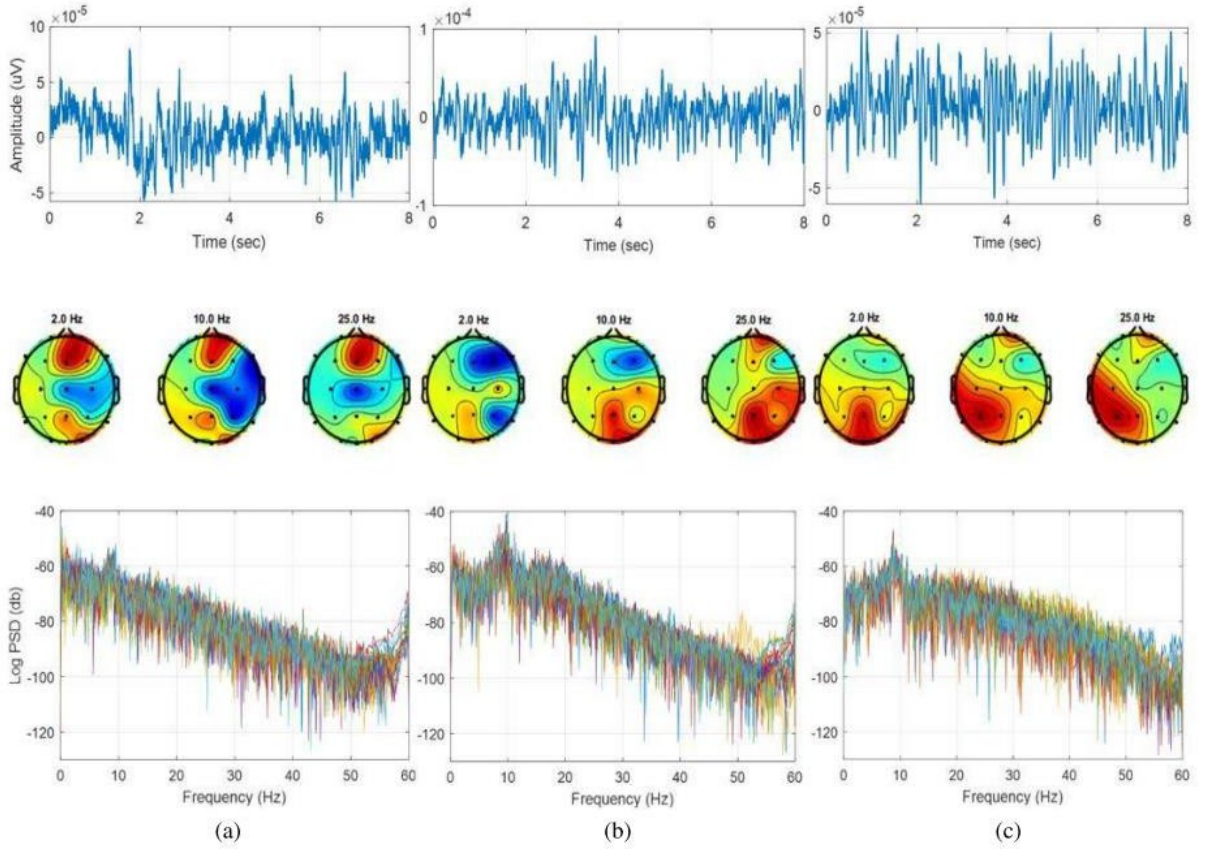

Fig 2. Samples of EEG data, electrode mappings, and power spectrum density pattern for three different EEGs: (a) control; (b) mild AD; and (c) moderate AD.

People with intermediate Alzheimer's disease needed to have an MMSE score of 20, with an average of 14.18 anda variance of 3.69, as well as a CDR score of 2. Inclusion in both AD cohorts (AD1 and AD2) was also necessarydue to the existence of cognitive and functional decline in the year prior, as long-term interviews with competent informants indicated. Several other diseases that may possibly contribute to cognitive loss were examined in bothAD groups, including B12 (vitamin) deficiency, diabetes, renal illness, thyroid disease, addiction to alcohol, liveror lung disease [24]. An example of electrode mappings, EEG power spectrum density, and EEG signals for threedifferent datasets—control, moderate Alzheimer's disease, as well as moderate Alzheimer's disease—are shown in Figure 2 on a logarithmic scale. The patterns of waves from the Fp1 electrodes collected from three different individuals from three different datasets make up the EEG signal sample. Three different arbitrary frequencies—2, 10, and 25 Hz—for which electrode maps are given in order to show how the three datasets differ from one another.

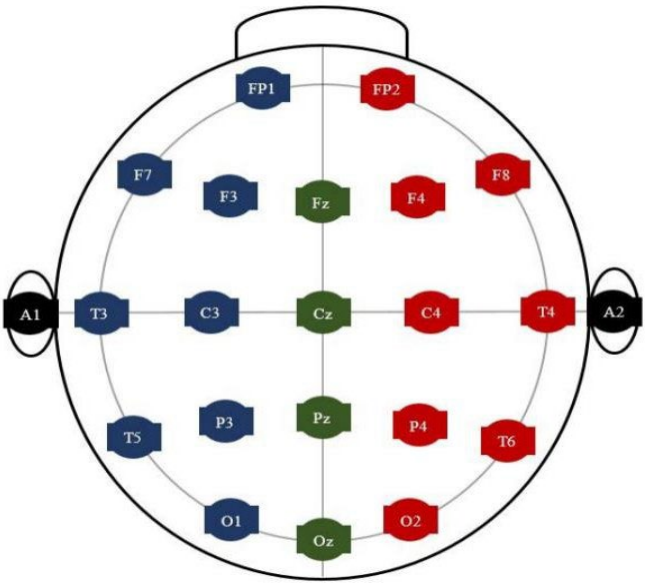

Fig 3. Distribution of the EEG acquisition system electrodes on the scalp

Using the Braintech 3.0 equipment collecting system (EMSA Medical Equipments Inc., Brazil), the EEG datasetwas acquired at a sampling rate of 200 Hz and an accuracy of 12 bits. The International 10-20 protocol was followed while positioning electrodes with the EEG data collection instrument. The subjects' earlobes, A1 on theleft and A2 on the right, as well as a pair of electrodes were used to collect EEG datasets for this study. The subjects' Fp1, Fp2, F3, F4, F7, F8, C3, C4, T3, T4, P3, P4, T5, T6, O1, O2, Fz, Cz, Pz, and Oz. Figure 3 depicts the dispersion of each electrode. Everyone was awake, calm, and wearing closed eyelids when the assessment wasdone. EEG artifacts such as muscle movements and blinking were meticulously eliminated from the data by two experienced neurophysiologists. Following that, a minimum of 28 epochs, that extending eight seconds, were selected by examining each person's eyes [25].

Data pre-processing

The research data is publicly available at https://github.com/tsyoshihara/Alzheimer-sClassification-EEG/tree/master/data. EEGs were obtained from 100 patients, including 14 HC subjects, 37 MCI patients, and 49AD patients. Prior to removing noise and interference from brainwave patterns, the recorded signals from the EEGare analyzed and filtered using a pre-processing block. An elliptic band-pass filter is used to successfully confinethe signals to a frequency range of 0.1 to 60 Hz. Beyond the EEG dataset, noises, interferences, and artifacts were recorded. The electrodes, magnetic fields from electronic devices, blood pressure, breathing, limb movement, eye blinking, and other people's actions were all responsible for these distortions, noises, and interferences [26]. During the preprocessing stage, the EEG signals were filtered using a band-pass filter to remove noise and disruptions caused by the EEG recording process. Numerous finite impulse response (FIR) and infinite impulse response (IIR) filters were in use. In this experiment, a band pass IIR elliptic digital filter with cutoff frequenciesof 0.1 and 60 Hz was employed. Two skilled neurophysiologists carefully removed EEG artifacts from the EEG data, such as blink and contractions of the muscles [27].

Band pass elliptic digital filter: Band-pass filters allow signals inside the band-pass area to flow through while attenuating frequencies outside of the band-pass zone. The passband and stopband ripple can be adjusted individually. If the stopband and passband ripples approach zero independently, the filter can change into a Chebyshev type I or type II filter. As a result, in order to have a thin transition band, both the passband and stopband ripples must be permitted. Elliptical (also known as Cauer) filters provide a smaller transition band thanother filters of the same order [28]. Discrete-time filters can be built in either IIR or FIR modes. The IIR filter is a feedback device with an infinite impulse response, as shown in Figure 2. Equations (1), (2), and (3) offer the mathematical difference equations that define the functioning of the IIR filter. For the IIR elliptic band-pass filter,we choose 100 kHz sampling frequency (f_{samp}), 2 kHz lower cut-off frequency (f_{lco}), 2.5 kHz upper cut-off frequency (f_{uco}), 1 dB

pass band ripple, and 80dB stop-band attenuation. Because the filter is of fourth order, it contains four zeros and four poles in the complex plane, as illustrated in [29]. The distance between the zeros andthe center is one. Using the suitable Bilinear transformation method, an analogue filter with a transfer function $H(sp)$ may be turned into a digital filter with a transfer function $H(zp)$. A projection between the s-plane and z-plane variables can be constructed if $sp = f(zp)$, This matching is utilized for the band-pass filter, and the equivalent frequency may be found by replacing $sp = j\omega$ and $zp = e^{j\omega}$, where ω is the angular frequency and $\omega = \frac{2\pi f}{f_{samp}}$. Equations (4) and (5) define the bilinear transformation employed in the band-pass filter.

$$y[n] = \frac{\sum_{k=1}^{M} rev_k x[n-k] - \sum_{k=1}^{N} frw_k y[n-k]}{\sum_{k=1}^{M} frw_k z^{-k}} \quad (1)$$

$$H(sp) = \frac{\sum_{k=1}^{N} rev_k z^{-k}}{\sum_{k=1}^{N}}, rev_0 = 1 \quad (2)$$

$$H(zp) = \frac{frw_0 \prod_{k=0}^{M} (1 - rev_k z^{-1})}{\prod_{k=1}^{N} (1 - frw_k z^{-1})} \quad (3)$$

$$s = \frac{1 - 2\cos\omega_0 z^{-1} + z^{-2}}{1 - z^{-2}} \quad (4)$$

$$\Omega = \frac{\cos\omega_0 - \cos\omega}{\sin\omega} \quad (5)$$

Where rev_k and frw_k are the reverse and forward filter coefficients, respectively, T_s represents the sampling period and ω_s corresponds to the center frequency of the band-pass/stop filter.

Feature Extraction

In order to handle EEG signals as efficiently as possible, the feature-extraction approach is essential. To deal withthe relatively small number of values that indicate the properties of the EEG signal, the signals are recorded and split into long time-series. These numbers are often referred as being features since they are merged into a vectorknown as the vector with features. Consequently, feature-extraction methods may be characterized by methods for converting signals into feature vectors. To extract features, a variety of feature-extraction methods are employed. The most popular and often applied approach, DWT, has been applied in this investigation. We proposeto use DWT with the cascade method (CWT) in this study to create vectors of attributes that include the parametersas follows: Logarithmic band power (LBP), Standard deviation (StD), Variance (VAR), Kurtosis (KUR), AverageEnergy (AE), Root mean square (RMS), and NO (Norm).

Using the STFT, non-stationary data analysis is not possible, such as EEG data. This can be attributed to STFT's consistent accuracy over the entire wavelength range. The technique of multi-resolution wavelet transform is employed to assess various frequencies at various resolutions. Additionally, the wavelet transform would be the ideal choice to avoid the related dimensionality problem because it may provide fewer attributes for the signal tobe analyzed. Wavelet transformations, for example, may analyze signal properties in the temporal along with frequency domains by decomposing signals across multiple steps via a single function [29]. This function is provided by and is known as the mother function ξ.

$$\xi(t) = \frac{1}{\sqrt{2}} \xi\left(\frac{t-y}{x}\right) \ldots x, y \in WS, x > 0, \quad (6)$$

where the wavelet space (WS) is represented by the parameter and the scaling and shifting parameters, respectively, by x and y. Wavelet transform is represented by the following equation.

$$F(x,y) = \frac{1}{\sqrt{x}} \int \xi \left(\frac{t-y}{x}\right) dt \quad (7)$$

DWT was utilized in this work because it provides a very useful wavelet representation. Low- and high-pass filtersare commonly employed in first-level decomposition to offer a depiction of the digital information as detail (D1)and estimate (A1) components. The following equation provides the DWT decomposition:

$$f(t) = \sum_{k=-\infty}^{k=+X\infty} DC_{n,k}\emptyset(2^{-n}t - k) + \sum_{k=-\infty}^{k=+X\infty} \frac{-j}{2^2} \times ac_{j,k} \times \psi(2^{-j}t - k) \quad (8)$$

where the detail coefficients are represented by $DC_{n,k}$ and the approximation coefficients by $ac_{j,k}$, respectively; n indicates the level and $\emptyset$ is the function of scale. The process is repeated once the first estimate is deconstructed.At the end of the process, n+1 deconstructed signals are present. The mother wavelet function utilized in this study is Daubechies 4 (db4); level 4 was selected since it provides the best qualities for correctly detected signal features.The following techniques were employed to produce the extra feature vectors: There are $n = 1, 2, \ldots, N$, discretesignal samples in S(n), where N is the number of signal samples.

The signal's variance

$$\underline{Var}_s = \frac{1}{N} \sum_{n=1}^{N} (S(n) - \mu_s) \quad (9)$$

$$\sigma_s = \sqrt{\frac{1}{N} \sum_{n=1}^{N} (S(n) - \mu_s)^2} \quad (10)$$

where the signal sample mean is denoted by μ_s [29]. The signal's Standard Deviation (σ)

The kurtosis of the signal

$$kurt_{\sigma_s} = E\left[\left(\frac{S(n)-\mu_s}{\sigma_s}\right)\right] \quad (11)$$

where the anticipated value of the signal samples is denoted by $E[\,]$ [32]. The NNSE (non-normalized SE) [33]

$$NNSE = \sum_{n=1}^{N} |S(n)|^2 \log|S(n)|^2 \quad (12)$$

Cascade algorithm:

Wavelets built such that the family of wavelets produced through both scale and translating by any parameter of either form is an orthonormal family are advantageous for the discrete wavelet transform (DWT) [26]. The DWT method splits all approximate values for a signal f through two parts: the details of the coefficients (also called wavelet coefficients) of the coarser estimation a, which produce a reduced version of the signal f filtered inside the high-pass filter. This procedure is done through the use of wavelets as well their corresponding scaling function. An orthonormal basis of $L^2(\mathbb{R})$. is the invertible decomposition of every function experiencing such a transformation. Furthermore, because there is a factor two in scale between the levels as well as coefficient translation, it can be shown that the final decomposition at level l may be stated as the product for the original non-dilated wavelet with the preceding estimated value, which had been sub-sampled with a factor of two. In a similar vein, the reconstruction at every level may be quantized as the average of the conjugate originating wavelet's convolution with the level before

it's reconstruction, where interpolated zeros are added in between eachsample. With the help of this characteristic, DWT is able to downscale and interpolate the signals at every stage of the reconstructions and decompose with zeros, respectively, using a single wavelet to conduct a cascaded procedure that is the same at every level. With the help of this characteristic, DWT is able to downscale and interpolate the signals at every stage during reconstruction as well as decomposition with zeros, respectively, using a single wavelet to conduct a cascaded operation that is identical at every level. We call it the cascade algorithm. A convolution neural network with two two L layers, L encoded blocks, and L decoded blocks might be used to mimic the DWT. There are two outputs on each encoding block, a single which is connected via the skip connection to the matching decoding block. In this scenario, we suggest training the pertinent filters for making the network learnable.

Modified Convolutional Neural Network for EEG Classification

CNN [30] has a weight sharing network topology that resembles a real brain network, making it a bionic structure. Reducing the weight and complexity of the network has shown to be successful. This device handles multidimensional pictures with remarkable power. Since the entire network can be thought of as an end-to-end system, the procedure involving gathering and categorizing features requires far less sophisticated data generation. Owing to CNN's potent capacity to learn spatial properties from two-dimensional input, it can effectively extract detailed spectral information from various brain areas when used to EEG processing. Convolution and pooling are the two fundamental working processes of the CNN structure, which significantly streamlines the preprocessing procedure. In order to obtain the appropriate feature maps, the convolutional kernel in the convolutional layer functions as a filter during image processing. Finally, a multi-layer convolution process allowsus to extract picture characteristics of varying hierarchies. An example of the convolutional computation is as follows:

where M_j represents the input layer's receptive field, f stands for the activation function, l for the number of layers, and ck and b for the convolutional kernel and bias, accordingly.
The term "pooling layer" also refers to the sub-sampling layer, which mostly consists of overlapping, mean pooling, and max pooling. In order to reduce the feature maps' dimensionality, this layer is often positioned below the convolutional layer. In this work, CNN will employ maximum pooling, while the pooling layer's expression is displayed as follows:

$$x_j^l = f\left(\sum_{i\in M} x_i^{l-1} * ck_{ij}^l + bias_j^l\right) \quad (13)$$

$$x_j^l = f\left(\beta_j^l down\left(cx_j^{l-1}\right) + bias_j^l\right) + \quad (14)$$

where β and b stand for the coefficient and bias, respectively, and $down(.)$ for the pooling function. The MCNN structure in this article was constructed using the Inception module. It is primarily motivated by GoogLeNet [30], which shows that network expansion will greatly improve the network's capacity for feature extraction. Figure 2 shows the inception module that we will be using.

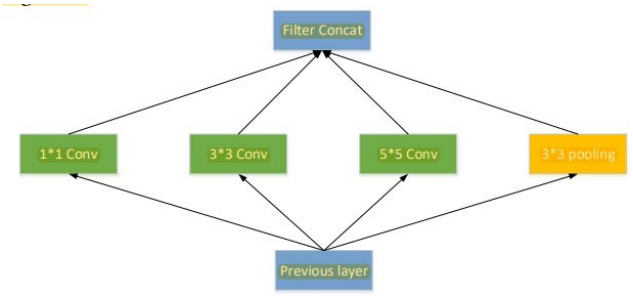

Fig.2. Network structure of Inception module

Convolutional kernel stacking of 1*1, 3*3, 5*5, and 3*3 pooling gives the inception module a stronger ability to adapt to input size. Additionally, the network gains a multi-scale fusion feature and a range of sensory fields, which

significantly enhances the model's generalization power. We have used an inception module which will efficiently record the distribution rule and characteristic information within the original EEG spectral pictures. Onthe other hand, the implementation of the inception module raises the information density of the network, which ultimately drives up the computation cost. Consequently, in order to reduce computation complexity, we use ReLU as the activation function. Figure 3 depicts the MCNN, the proposed CNN structure.

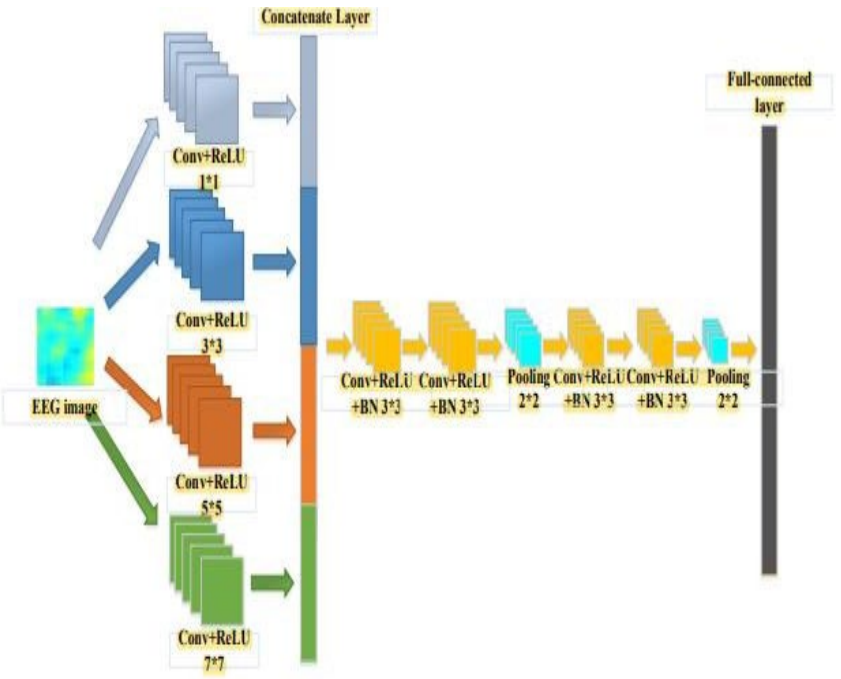

Fig.3. Architecture of MDCNN

As seen in Figure 3, the Inception module, which has four layers of convolutional kernels, serves as the foundationfor the first layer of CNN. The first layer's input is designated by $h(x)$, where x stands for the input pictures, $h_m(x)$ for the mth layer's output, Θ for the convolutional process, and $\beta_m$and $bias_m$ for the convolutional kernel'scoefficient and bias, respectively. The output of the inception module may then be shown as follows:

1 4 $m=1$ $1m$ $1m$

Features of several spatial scales can be obtained by using the multi-scale convolution process of the inception module. Using convolutional operations, the activation values of distinct features from different positions in the provided image may be calculated. Every neuron in this layer may then have its output value obtained using ReLU. The resulting data is finally concatenated using the concatenate layer. Using the same 128*3*3 convolutional kernel setup, the remaining four convolutional layers in the MCNN were comprised. Features withmore abstract and detailed information can be extracted using the multi-layer convolution process. With 1024 neurons in the full-connected layer, the input feature maps would be converted into a 1024*1 one-dimensional vector. Additionally, a batch normal (BN) [31] technique was implemented to speed up network training and normalizing. Between the convolution and activation operations is where BN is operated. It successfully preventsinternal covariate shift. The remaining layers' output is shown as:

$$H_m(x) = \max(0, \hbar_{m-1}(x)\Theta\beta_m + bias_m) \quad (16)$$

Based on the quantity of EEG datasets, the following five categorization issues have been studied: Two classes of characteristics: Control in comparison to symptoms of mild and moderate AD Three classes of AD characteristics were compared: mild against moderate, regulate versus weak and moderate, and controls vs weak vs intermediate(3-class). B; m[][pl9ygu90—09uhgv 6The hyperparameter definition has a significant impact on the network's efficiency. The most significant issue is that there is no accepted paradigm for determining the ideal set of MPCNN hyperparameters and creating the architecture of the network. Using the MCSA algorithm in this work, the authors created an automated method for hyperparameter optimization and structural design to solve this problem. The MDCNN model's chosen hyperparameters are batch size, number of epochs, number of filters, pooling size, and kernel size.

MCSA Method: The primary concept behind this work's MCSA implementation—which draws inspiration fromthe forgetting mechanism—is to substitute a special forgetting mechanism for the receptor editing mechanism [24] in the CSA. Replacing the CSA's receptor editing mechanism with a special forgetting mechanism is the main idea behind its implementation. Here is the precise way of implementation: The suitable memory strength and life duration of every antibody candidate set are noted throughout every algorithm iteration. The algorithm's determination of antibody forgetting was based on whether protein activity crossed the threshold after several repetitions. A new, clever algorithm called MCSA can also efficiently overcome prematurity and has a quick rateof convergence, ensuring both local and global search capabilities and improving algorithmic performance.

Affinity Calculation. The function value, which can be written as, ab is the target test function of antibody ab affinity aff to antigen in order to streamline the computation.

$$aff_{ab_i} = f(x_1, x_2, \ldots, x_{dim}) \quad (17)$$

where dim is the antibody's dimension and $ab = \{x_i | i = 1,2, \ldots, dim\}$ is the antibody.

Cloning Method: The cloning approach selects the hyperparameter value for the antibody cloning based on the affinity that matches the antigen and antibody. The number of antibodies which might be cloned increases with affinity. The particular cloning formula that is used is

$$ab_{clone} = \{ab_{ij}|_{i \in [0, ps-1]}\}_j = \max(1, int(^{aff_{ab_i}+ab}_{ab} * max\ clone)), ab > 0 \quad (18)$$

The number of antibodies in the probable set of antibodies is denoted by ps population size, the affinity betweenthe antibody ab_i and the antigen is represented by aff_{ab_i}, the starting clone number is indicated by max clone, and the clone number of the antibody ab_i is shown by j.

Variation Method: The purpose of the mutation approach is to determine the antibody's level of mutation basedon the affinity of the cloned antibody for the antigen. The likelihood and severity of a mutation decrease with increasing affinity. The particular variant formula that is
When $aff > 0$, max affinity is the maximal affinity of the concentrated antibody; mr is the mutation rate, and an is the variation range.

Forgetting Method: Based on the antibody's life duration, the right memory strength, and the Rac1 protein's activity, the forgetting technique establishes whether antibody forgetting is necessary. The precise formula for forgetting is

where the suitable memory strength is ab_s, the th protein activity threshold is γ, and the antibody survival duration is ab_t. Algorithm 1 illustrates how the enhanced method suggested in this study will operate. It is possibleto modify the suspension conditions of Algorithm 1's algorithms to suit certain requirements. A common termination condition is when the function evaluation reaches its maximum value or when the number of generations reaches its limit. th Protein activity, which is determined by the algorithm based on the antibody's survival duration and suitable memory strength at the time of initial selection into the candidate set, is an intrinsic feature of every candidate antibody. Additionally, it varies dynamically while the algorithm is run. The antibody is not adequately competing relative to other candidate antibodies when its property value hits the threshold, indicating that it has never mutated in an improved direction during the anticipated time frame. Accordingly, the algorithm will carry out the forgetting operation for the antibody that satisfies the threshold value.

Experimental Results and Discussion

The sensitivity, specificity, receiver operating characteristic curve area, and classifier accuracy may all be used to evaluate a diagnostic system. A multitude of classifiers were developed implemented and evaluated in order to obtain the greatest diagnostic efficiency and accuracy in classification. In this study, the proposed MCNN is compared with various methods, such as RF, ELM, QDA, KNN, SVM, ANN, and classifiers, that produced the best outcomes. The k-fold cross-validation approach is used in the classification process, which randomly dividesall EEG features into k equal groups [38]. The balance of subsets is used for training, and one subset is chosen fortesting (validation). A selected number of each has been used once in the k repetitions of this procedure. We used10-fold cross-validation in the current study, putting all of the detected EEG signal characteristics into a feature matrix that was obtained using feature extraction techniques, and then applying 10-fold a cross-valid to the featurematrix. Following that, these qualities were split into two distinct subsets: 10% for testing and 90% for training. The examination subset was given into the model after it was successfully trained each time, and the training subset was used for training the classifier with the aim to build and preserve the database of features of the learnedclassifier.

As mentioned before, the EEG data utilized in this study were divided into 3 categories. Thirty-one moderate AD patients, twenty mild AD patients, as well as 30 Normal subjects submitted the EEG records for the initial, next, and final groups, respectively. The electroencephalogram (EEG) dataset was processed utilizing a band-pass IIRelliptic digital filter having the threshold frequency of 0 and 60 Hz in order to improve the signal-to-noise ratio. The CWT technique was then used to derive the features from the filtered EEG data. Next, various statistical parameters including LBP, StD, Var, etc. have been combined with the CWT technique to improve diagnostic performance and generate the EEG feature vectors. Ultimately, several classifier types were employed to divide the EEG data into groups based on these classifications, and the classification accuracy of each group was calculated and contrasted. An in-depth evaluation of the recommended techniques was made possible by plottingthe receiver's operating characteristic curves and calculating the regions beneath them.

Results Evaluation: Table 2 shows that the classifiers with the greatest overall accuracy in classification are theRF, ELM, QDA, KNN, SVM, ANN, and ones. The median accuracy rates of these classifiers are 99.7, 99.8, 99.85,99.9,

and 99.98%, in that order. For a more comprehensive evaluation of the recommended approaches, Figure 5 demonstrates the ROC curves. According to which features provide the maximum accuracy, Table 3 shows the outcome for the calculations of the threshold, particulars, accuracy of classification, and the area of the ROC curves of each classifier. It is clear that ANN, SVM, RF, ELM, QDA, and KNN were the classifiers that performedthe best.

Its definition is a ratio of correctly segmented samples to the total number of samples. This represents one of themost frequently utilized metrics of the effectiveness of classification, as seen below.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (22)$$

Where True Positive (TP): This denotes the accurate detection of AD. True Negative (TN): This status shows thata feature was appropriately identified as non-AD. FP: It means that a feature was mistakenly identified as AD. False Negative (FN): This signal means that feature was missed and classified as an AD.

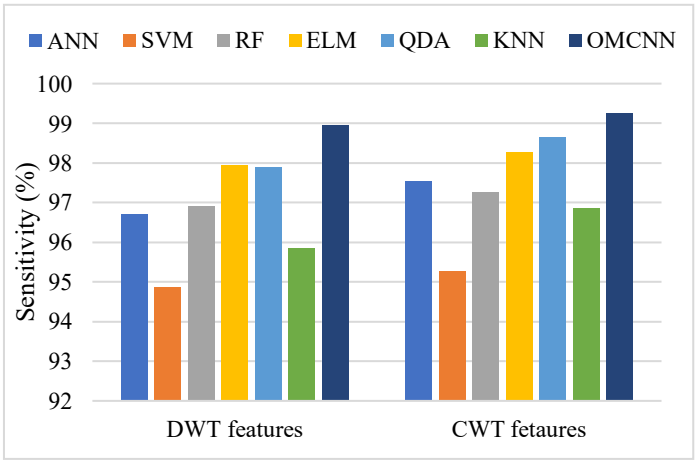

Fig.4. Sensitivity performance comparison

Figure 4 compares the sensitivity of the existing RF, ELM, QDA, KNN, SVM, ANN, and that of the proposed OMCNN. Figures comparing the proposed strategy to existing methods indicate that it can achieve a high sensitivity rate when used for DWT and CWT based feature extraction. When compared to other existing approaches such as RF, ELM, QDA, KNN, SVM, ANN, and, OMCNN for DWT achieves a high sensitivity rateof 98.96%, indicating its effectiveness in AD detection. By obtaining a sensitivity rate of 99.25% for CWT, OMCNN enhances the issue. The recommended solution displays a significantly lower end error and a substantially faster and less variable training curve when compared to the preceding strategies. These results provide additional evidence that the suggested method has a greater rate of AD detection.

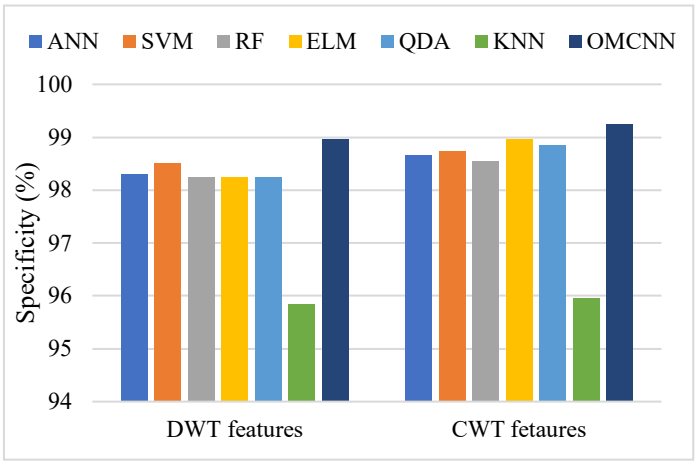

Fig.5. Specificity performance comparison

The results of a specificity comparison between the present RF, ELM, QDA, KNN, SVM, ANN, and KNN the proposed OMCNN are shown in Figure 5. According to the results, the suggested OMCNN obtains specificity rates of 99.25% for IWT-based feature extraction techniques and 98.96% for DWT-based feature extraction approaches. Based on a specificity rate comparison between OMCNN and ANN, SVM, RF, ELM, QDA, and KNN, which yields lower rates, the suggested study might yield superior findings for AD identification than current methods. Compared to RF, ELM, QDA, KNN, SVM, ANN, and KNN networks, the OMCNN network trains more quickly. It also has a more effective CWT feature extraction procedure, which raises the specificity value.

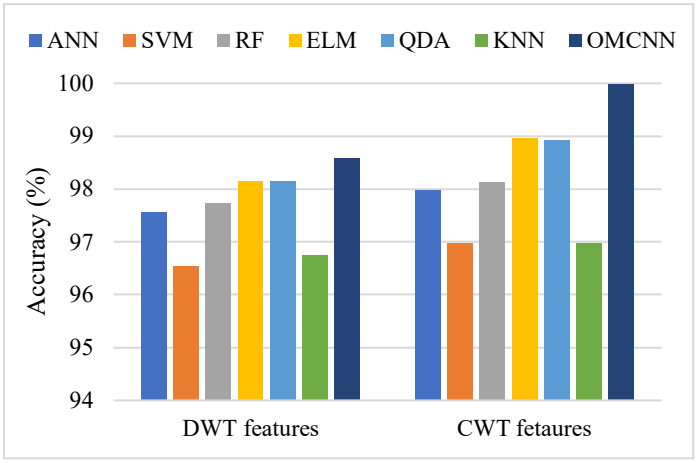

Fig.10. Accuracy performance comparison

Fig. 6 shows the accuracy comparison of the proposed OMCNN with the current RF, ELM, QDA, KNN, SVM, ANN. Based on the data, a high accuracy rate value was obtained by the suggested OMCNN, suggesting a high rate of AD detection. With an efficient prediction model, the training time would be shortened by the filtering andfeature extraction based on CWT and DWT in the suggested scheme. Accuracy values of 98.59 and 99.99% for DWT along with CWT, respectively, show that the suggested study can provide superior AD detection findings than current approaches. Medical professionals and physicians can benefit from the suggested method's automated, quick, simple, accurate, and efficient identification of Alzheimer's disease. The suggested method maycut down on the restricted number of neurologists, speed up diagnosis, and increase diagnosis precision.

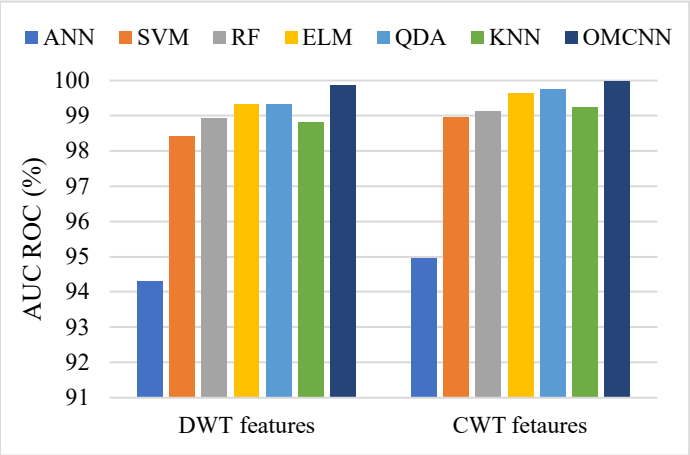

Fig.7. AUCROC performance comparison

The AUCROC comparisons for AD prediction is shown in Fig. 7 above. Methods like DCCNN-SMO, OMCNN Classifier, VELA, ML-IDS, and others are applied. As the number of photos rises, the AUCROC value grows linearly. The suggested method has a high AUCROC rate, with DWT and CWT-based feature extraction achieving 98.59% and 99.12%, respectively. The findings demonstrate that, in terms of improved liquid particle identification results with high accuracy rates, the suggested OMCNN technique outperforms the current algorithms. OMCNN learning approaches improve accuracy without causing the local optima issue because theyare very resilient to distortion in data used for training. These results imply that the suggested method is better capable of reliably and effectively differentiating not just AD however their sorts.

Conclusion and Future work

The main goal of this project was to develop an Alzheimer's disease diagnosis tool based on EEG data processing.The creation of an automated diagnostic system which can interpret signals from the brain on itself would quickenup and boost the reliability of the diagnosis process. The gathered EEG datasets during this investigation were filtered using BEF. After that, many signal features were combined using the DWT methodology to improve diagnosis performance, and the CWT approach was created to divide the signal that was filtered into frequency bands. OMCNN then looked into how EEG characteristics were categorized into groups. This study aims to evaluate many approaches and determine the best combination approach for Alzheimer's disease diagnosis. Datasets related to manageable, minor, and moderate Alzheimer's illnesses were utilized. The proposed system obtained 99.99% precision for classification annually using CWT. Additionally, the accuracy of the present OMCNN was enhanced by 0.01% to 2% when compared to the traditional model prior to the implementation of the CWT function. This makes the recommended approach for processing EEG signals very important, and the best classification would result from using a customized hyper-parameter model to predict the severity range. Thisstudy used PCA to remove unnecessary and redundant information from EEG data and signal processing techniques to identify characteristics that might be helpful for future early Alzheimer's disease detection.

References

1. Ravindranath, V., & Sundarakumar, J. S. (2021). Changing demography and the challenge of dementia in India. *Nature Reviews Neurology*, *17*(12), 747-758.
2. Gallagher, D., Coen, R., Kilroy, D., Belinski, K., Bruce, I., Coakley, D., ... & Lawlor, B. A. (2011). Anxiety and behavioural disturbance as markers of prodromal Alzheimer's disease in patients with mildcognitive impairment. *International journal of geriatric psychiatry*, *26*(2), 166-172.
3. Dauwels J, Srinivasan K, Reddy MR, Musha T, Vialatte FB, Latchoumane C, et al. Slowing and Loss ofComplexity in Alzheimer's EEG: Two Sides of the Same Coin? *International Journal of Alzheimer's Disease*. 2011. Vol. 2011, Article ID 539621, 10 pages.
4. Alberdi A, Aztiria A, Basarab A. On the early diagnosis of Alzheimer's disease from multimodal signals:A survey.

Artificial Intelligence in Medicine. 2016; 71:1–29.

5. Malek N, Baker MR, Mann C, Greene J. Electroencephalographic markers in dementia. *Acta Neurologica Scandinavica*. 2017; 135(4):388–393.
6. AlSharabi, K., Salamah, Y. B., Abdurraqeeb, A. M., Aljalal, M., & Alturki, F. A. (2022). EEG signal processing for Alzheimer's disorders using discrete wavelet transform and machine learning approaches. *IEEE Access*, *10*, 89781-89797.
7. Falahati, F., Westman, E., & Simmons, A. (2014). Multivariate data analysis and machine learning in Alzheimer's disease with a focus on structural magnetic resonance imaging. *Journal of Alzheimer's disease*, *41*(3), 685-708.
8. Acharya U.R., Sree S.V., Chattopadhyay S., et al. Automated diagnosis of normal and alcoholic EEG signals. *Int. J. Neural Syst.* 2012;22(3).
9. Tavares G., San-Martin R., Ianof J.N., et al. *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)* 2019. Improvement in the automatic classification of Alzheimer's disease using EEGafter feature selection.
10. Rivera M.J., Teruel M.A., Maté A., Trujillo, et al. Diagnosis and prognosis of mental disorders by meansof EEG and deep learning: a systematic mapping study. *Artif. Intell. Rev.* 2021;55(12):1–43.
11. Khare, S. K., & Acharya, U. R. (2023). Adazd-Net: Automated adaptive and explainable Alzheimer's disease detection system using EEG signals. *Knowledge-Based Systems*, *278*, 110858.
12. Musaeus, C. S., Frederiksen, K. S., Andersen, B. B., Høgh, P., Kidmose, P., Fabricius, M., ... & Kjær, T. W. (2023). Detection of subclinical epileptiform discharges in Alzheimer's disease using long-term outpatient EEG monitoring. *Neurobiology of Disease*, *183*, 106149.
13. Puri, D. V., Nalbalwar, S. L., Nandgaonkar, A. B., Gawande, J. P., & Wagh, A. (2023). Automatic detection of Alzheimer's disease from EEG signals using low-complexity orthogonal wavelet filter banks. *Biomedical Signal Processing and Control*, *81*, 104439.
14. Imani, M. (2023). Alzheimer's diseases diagnosis using fusion of high informative BiLSTM and CNN features of EEG signal. *Biomedical Signal Processing and Control*, *86*, 105298.
15. Jia, H., Huang, Z., Caiafa, C. F., Duan, F., Zhang, Y., Sun, Z., & Solé-Casals, J. (2023). Assessing the Potential of Data Augmentation in EEG Functional Connectivity for Early Detection of Alzheimer's Disease. *Cognitive Computation*, 1-14.
16. Cejnek, M., Vysata, O., Valis, M., & Bukovsky, I. (2021). Novelty detection-based approach for Alzheimer's disease and mild cognitive impairment diagnosis from EEG. *Medical & Biological Engineering & Computing*, *59*, 2287-2296.
17. Dogan, S., Baygin, M., Tasci, B., Loh, H. W., Barua, P. D., Tuncer, T., ... & Acharya, U. R. (2023). Primate brain pattern-based automated Alzheimer's disease detection model using EEG signals. *Cognitive Neurodynamics*, *17*(3), 647-659.
18. Chiang, H. S., & Pao, S. C. (2016). An EEG-based fuzzy probability model for early diagnosis of Alzheimer's disease. *Journal of medical systems*, *40*, 1-9.
19. Jiao, B., Li, R., Zhou, H., Qing, K., Liu, H., Pan, H., ... & Shen, L. (2023). Neural biomarker diagnosis and prediction to mild cognitive impairment and Alzheimer's disease using EEG technology. *Alzheimer'sresearch & therapy*, *15*(1), 1-14.
20. Miltiadous, A., Gionanidis, E., Tzimourta, K. D., Giannakeas, N., & Tzallas, A. T. (2023). DICE-net: a novel convolution-transformer architecture for Alzheimer detection in EEG signals. *IEEE Access*.
21. S. Brucki, R. Nitrini, P. Caramelli, P. Bertolucci, and I. H. Okamoto, ''Suggestions for utilization of themini-mental state examination in Brazil,'' Arquivos de Neuro-Psiquiatria, vol. 61, no. 3B, pp. 777–781,2003.
22. G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, and E. M. Stadlan, ''Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA work group* under the auspices of Departmentof Health and Human services task force on Alzheimer's disease,'' Neurol., vol. 34, no. 7, p. 939, 1984.
23. Diagnostic and Statistical Manual of Mental Disorders, American Psychiatric Association, Washington,DC, USA, 1980, vol. 3.
24. P. A. M. Kanda, L. R. Trambaiolli, A. C. Lorena, F. J. Fraga, L. F. I. Basile, R. Nitrini, and R. Anghinah,''Clinician's road map to wavelet EEG as an Alzheimer's disease biomarker,'' Clin. EEG Neurosci., vol.45, no. 2, pp. 104–112, 2014
25. R. Cassani, T. H. Falk, F. J. Fraga, P. A. Kanda, and R. Anghinah, ''Towards automated EEG-based Alzheimer's disease diagnosis using relevance vector machines,'' in Proc. 5th ISSNIP-IEEE Biosignals Biorobot. Conf., Biosignals Robot. Better Safer Living (BRC), May 2014, pp. 1–6.

26. X. Perrin, ‘‘Semi-autonomous navigation of an assistive robot using low throughput interfaces,’’ Ph.D. dissertation, ETH Zurich, Zürich, Switzerland, 2009.
27. F. J. Fraga, T. H. Falk, L. R. Trambaiolli, E. F. Oliveira, W. H. L. Pinaya, P. A. M. Kanda, and R. Anghinah, ‘‘Towards an EEG-based biomarker for Alzheimer’s disease: Improving amplitude modulation analysis features,’’ in Proc. IEEE Int. Conf. Acoust., Speech Signal Process., May 2013, pp.1207–1211.
28. Prommee, P., Karawanich, K., Khateb, F., & Kulej, T. (2021). Voltage-mode elliptic band-pass filter based on multiple-input transconductor. *IEEE Access*, *9*, 32582-32590.
29. Chavan, M. S., Agarwala, R. A., & Uplane, M. D. (2006). Digital elliptic filter application for noise reduction in ECG signal. *WSEAS transactions on Electronics*, *3*(1), 65-70.
30. Morabito, F. C., Campolo, M., Ieracitano, C., Ebadi, J. M., Bonanno, L., Bramanti, A., ... & Bramanti, P. (2016, September). Deep convolutional neural networks for classification of mild cognitive impaired and Alzheimer's disease patients from scalp EEG recordings. In *2016 IEEE 2nd International Forum onResearch and Technologies for Society and Industry Leveraging a better tomorrow (RTSI)* (pp. 1-6). IEEE.