

# Identification of Alzheimer's Disease Using Novel Dual Decomposition Technique and Machine Learning Algorithms from EEG Signals

Digambar Puri<sup>a,b,\*</sup>, Sanjay Nalbalwar<sup>a</sup>, Anil Nandgaonkar<sup>a</sup>, Jaswantsing Rajput<sup>a,b</sup>, Abhay Wagh<sup>b,c</sup>

<sup>a</sup> Department of Electronics and Telecommunication, Dr. Babasaheb Ambedkar Technological University, Lonere-402103, India

<sup>b</sup> Department of Electronics and Telecommunication, RAIT, D.Y. Patil deemed to be University, Navi Mumbai- 400706, India

<sup>c</sup> Directorate of Technical Education, Mumbai, 400001, Maharashtra, India

Corresponding author: \*digambarpuri@dbatu.ac.in; digambar.puri@rait.ac.in

**Abstract**— Alzheimer's disease (AD) is one of the neurodegenerative disorders. The rate of AD prevalence is rapidly increasing worldwide. The existing clinical invasive methods and neuro-imaging techniques to detect AD are time-consuming, subjective, and expensive. To overcome these issues, we proposed a new automatic framework for detecting AD at an early stage based on the dual decomposition method. Initially, EEG signals of mild cognitive impairment (MCI), AD, and normal control (NC) patients are divided into five subbands by employing discrete wavelet transform (DWT). Subsequently, a Variational mode decomposition (VMD) is applied to these five EEG subbands for further decomposition into various intrinsic mode functions (IMFs). Afterward, three different multiscale permutation entropy (PE) features, namely Shannon PE (SPE), Tsalli's PE (TPE), and Renyi PE (RPE), have been measured from each IMF. Later, these features have been used to train and test ensemble bagged tree (EBT), k-nearest neighbor, support vector machine (SVM), decision tree (DT), and neural networks with a 10-fold cross-validation scheme. The proposed method has been verified using EEG signals of 59-AD, 7-MCI, and 102-NC subjects. The results obtained from the proposed DWT-VMD method provide 95.20% accuracy for three-class and 97.70% for two-class classification using an EBT classifier with 10-fold cross-validation. It shows a significant ability to distinguish AD from MCI. The proposed dual decomposition method can employ for other neurodegenerative disorders such as Parkinson's disease, epilepsy, various sleep disorders, and major depressive disorders.

**Keywords**—Alzheimer's disease (AD); support vector machine (SVM); electroencephalogram (EEG); ensemble bagged tree (EBT).

Manuscript received 22 Aug. 2022; revised 7 Nov. 2022; accepted 27 Dec. 2022. Date of publication 30 Apr. 2023.  
IJASEIT is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.



## I. INTRODUCTION

Alzheimer's disease (AD) is one of the irreversible neurodegenerative disorders. The rate of AD prevalence is rapidly increasing worldwide. It shows symptoms like memory loss, cognitive dysfunctions, poor judgment, aphasia, difficulty in daily activities, etc. [1]. There is no cure for AD, but its rate can be reduced by detecting it early. The initial AD stage is Mild Cognitive Impairment (MCI) [2]. MCI problems occur in 5-20% of people older than 60; however, the symptoms of MCI are often not perceptible or visible, allowing patients to continue their daily routine. The detection of AD at the MCI stage is challenging due to overlapping symptoms with normal aging-related decline. As stated by Ruiz-Gómez *et al.* [3], 60 million people were living with AD in 2020, which is projected to become more than 130 million

by 2050 [3]. Thus, it is an urgent need to develop an automatic detection system for AD in its early stage. Such early detection of AD can help the dependent family members of AD patients to prepare mentally, physically, and financially. Hence, early diagnosis can help patients to maintain their independence for a long time and delay psychiatric symptoms like depression and social costs. Despite all of these, detecting MCI is a challenge for experienced neurologists [3]. AD and MCI detection are usually performed by a combination of blood tests, neuropsychological tests, psychiatric examinations, physical, and cerebrospinal fluid investigation [4]. In addition, various invasive clinical methods and neuroimaging techniques like computed tomography, fMRI, PET are widely used for detection of AD. However, estimating the results of all these tests are time consuming, subjective and expensive [5]. Therefore, researchers are

focusing on the development of non-invasive, cost effective, rapid and more reliable techniques for early diagnosis of AD [6]. Recently, researchers and clinicians have used the high temporal resolution feature of electroencephalograms (EEG) to investigate the early-stage detection of AD. It records the electrical variations from electrodes placed on the different locations of the human brain scalp. The brain dynamics of AD patients are described by five EEG rhythms 0-4 Hz ( $\delta$ ), 4-8Hz ( $\theta$ ), 8-16Hz ( $\alpha$ ), 16-32Hz ( $\beta$ ), and 16-32Hz ( $\gamma$ ) [7]. These bands show different characteristics in AD patients [8].

In the literature, various efforts have been taken to detect AD using time, frequency, and time-frequency domain methods from EEG [9]. The performance of various time-domain complexity and entropy-based neuro markers has been evaluated by Simon *et al* [10] and achieved a maximum classification accuracy of 96%. However, parameter selection is a major issue with complexity and entropy-based methods. Recently, Ding *et al.* [11] used spectral, complexity, and functional connectivity-based features with various machine learning (ML) classifiers to detect AD and achieved 80.08% accuracy. Similarly, Rodrigues *et al.* [12] proposed Lacsral and cepstral-based features for the three-class classification of AD, MCI, and NC subjects. They employed a support vector machine (SVM) and neural network (NN) to verify the effect on the AD performance of Lacsogram. The best accuracy of 98.06% has been obtained for NC vs. MCI and 96% for the three-class classification by ANN. Recently, a novel detection approach has been proposed by Cejnek *et al.*[13] for AD and MCI detection. They employed the adaptive gradient descent method to extract the neuro markers from EEG signals and reached 90.73% accuracy. In another study [14], multiple signals are classified with empirical wavelet transform (MUSIC-EWT) and Hurst exponent to detect AD and MCI. The enhanced probabilistic neural networks (EPNN) provided a 90.3% classification rate for MUSIC-EWT-based features. Safi *et al.* [15] investigated the performance of Horth parameters (HP) with discrete wavelet transform (DWT) to classify the EEG signals. They also used empirical mode decomposition (EMD) based HP features to detect MCI and AD. The HP-based features have been used to train and test ML algorithms such as SVM, K-nearest neighbor (KNN), and refined linear discriminant analysis. They reported that KNN achieved a maximum accuracy of 97.64% from 10-fold cross-validation (CV) [15]. Toural *et al.* [16] has proposed wavelet-based entropy, theta power, and Pearson's correlation coefficient features to detect MCI and AD patients from normal controlled (NC) subjects. They reported 94.44% accuracy using NN.

From the above-discussed state-of-art methods, it has been noted that the non-linear features failed to capture the hidden information present in the spectrum of EEG signals. The fast

Fourier Transform (FFT) assumes the EEG signal is stationary. It provides spectral information about the signal but also fails to provide the corresponding location in the time domain. In Short Time Fourier Transform (STFT), the long window provides better frequency resolution, and the short window gives a good time resolution. The selection of window length to define the time-frequency resolution is the major drawback of STFT. In most state-of-art AD decomposition methods, DWT or Tunable Q-wavelet transform [17], or EMD [18] has been used to extract features from EEG signals. Intrinsic mode functions (IMFs) of EMD are failed to understand low-frequency EEG signals [18]. EMD lacks a strong mathematical foundation and suffers from mode mixing, noise sensitivity, and endpoint effect. To address this issue, we developed an automatic AD detection technique based on the dual decomposition of EEG signals using DWT and Variational mode decomposition (VMD) [19]. The DWT has been used for the decomposition of EEG signals into a sequence of narrow frequency bands. These subbands are further decomposed into IMFs using VMD. The adaptive filter bank has been used in the VMD for noise robustness. The permutation entropy-based features have been measured from these subbands. Afterward, three different multiscale permutation entropy features were measured from each IMF. Later, these features were used to train and test various ML algorithms such as SVM, ensemble bagged tree (EBT), KNN, NN, and decision tree (DT) with a 10-fold CV scheme. The major contributions of the present work are as follows:

- Evaluation of dual-decomposition method to detect AD from EEG recordings.
- To explore the performance of the DWT-VMD decomposition method with ML models by generating a model for three-class classification.
- To discover an appropriate feature set through feature selection analysis and find a robust classifier.
- To enhance the performance of AD detection compared to existing methods.

## II. MATERIALS AND METHOD

The flow of the proposed method has presented in Fig. 1. The proposed method has four different stages. In stage-1, pre-processing the EEG signal and first decomposition using DWT is performed. The different subbands obtained from stage-1 are reconstructed and decomposed into six IMFs using VMD in stage-2. Afterward, different MPEs are calculated from IMFs. The significance of these features was evaluated using Kruskal Walli's (KW) test in stage-3. Different ML algorithms are trained and tested in the final stage of the proposed method to validate the performance of the proposed DWT-VMD-based features.

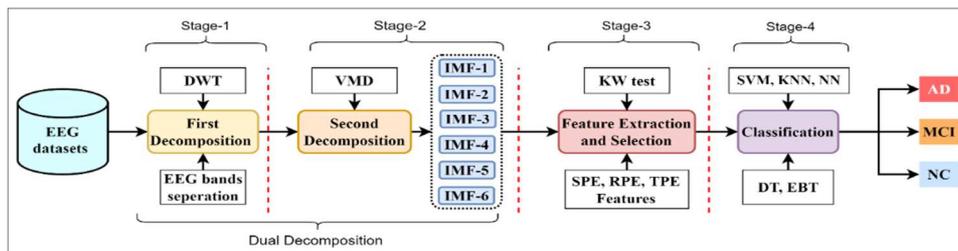


Fig. 1 Proposed methodology of the AD detection system using dual decomposition and machine learning algorithms

The details of the EEG dataset used for the experimentation are discussed below.

#### A. EEG Dataset Details

The present study has used the publicly available and previously utilized EEG dataset [13] of AD, MCI, and NC. In this dataset, the EEG signals are recorded from three different groups: AD, MCI, and NC. A group of 59 AD patients (28-male and 31-female) with ages ranging from  $(67 \pm 7.6)$  years has been used for EEG recording after the necessary physical and neurological examinations by the neurophysiologist. It has been confirmed that the MMSE (Mini-mental State Examination) points of the above-mentioned AD patients range from 10 to 19 with a mean value of 14.9. It is confirmed that all these patients are AD patients. Similarly, a group of seven MCI patients (3-male and 4-female) with age groups ranging from  $(70.5 \pm 4.9)$  has been used for EEG recordings. Similarly, a group of seven MCI patients (3-male and 4-female) with an age group ranging from  $(70.5 \pm 4.9)$  has been used for EEG recordings. Similarly, a group of seven MCI patients (3-male and 4-female) with an age group ranging from  $(70.5 \pm 4.9)$  has been used for EEG recordings. In addition, the other group of 102 NC subjects (43-male and 59-female) had also been used to record the EEG signals. All EEG signals were recorded in standard conditions with the eye closed and in the rest position. These signals were recorded using a 21-channel digital EEG (EEG PL-231) setup, according to a 10-20 electrode placement system. All EEG signals are then sampled at the 128 Hz sampling rate. More details about AD, MCI, and NC EEG datasets have been provided in [13]. Further, this dataset has been used to decompose using dual decomposition.

#### A. Dual-Decomposition: DWT-VMD Combination

The present work proposes the application of DWT and VMD separately for the decomposition of EEG signals. Hence, it is coined as dual decomposition. Recently, efforts have been made to detect epilepsy using the dual decomposition technique from EEG signals. Ji et al. [20] utilized the EMD-DWT approach with various spectral features. In our proposed method, all EEG signals of AD, MCI, and NC subjects are decomposed into four levels using Daubechies (db4) wavelets. The DWT for input  $x(t)$  is defined as follows [20]:

$$W_{a,b}(t) = \int x(t) \psi_{a,b}^*(t) dt \quad (1)$$

Where,  $\psi_{a,b}^*(t)$  is the conjugate of the mother wavelet function, and that is expressed as follows:

$$\psi_{a,b}(t) = 2^{-a/2} \psi(2^{-a}t - b) \quad (2)$$

Where,  $a$  and  $b$  are scale and translation constants. The result of four-level DWT is a series of  $(c_{a,b})$ , and detailed coefficients  $(d_{a,b})$  are expressed as follows:

$$c_{a,b} = \langle x(t), \phi_{a,b}(t) \rangle = \int 2^{-a/2} x(t) \phi(2^{-a}t - b) dt \quad (3)$$

$$d_{a,b} = \langle x(t), \psi_{a,b}(t) \rangle = \int 2^{-a/2} x(t) \psi(2^{-a}t - b) dt \quad (4)$$

Where,  $\phi_{a,b}(t)$  is called the scaling function. It is expressed as follows:

$$\phi_{a,b}(t) = 2^{-a/2} \phi(2^{-a}t - b) \quad (5)$$

The detailed and approximated subbands obtained from this decomposition are used to reconstruct the original time-domain EEG signals into significant EEG bands, such as  $\delta$ ,  $\theta$ ,  $\alpha$ ,  $\beta$ , and  $\gamma$ . These bands are extracted from inverse DWT [20] as expressed in equation (6) below.

$$x(t) = \sum_b c_{a,b} \phi_{a,b}(t) + \sum_b c_{a,b} \psi_{a,b}(t) \quad (6)$$

More details of DWT are available in [15-17].

Further decomposition has been carried out using VMD from individual reconstructed EEG signals into various frequency and amplitude-modulated signals, named as IMFs (intrinsic mode functions) [19]. IMFs have high noise immunity. IMFs ( $v_k$ ) has been achieved by the decomposition of input signal  $f$  into a required number of discrete sub-signals. Each IMF is reaching towards the central frequency ( $\omega_k$ ), obtained from decomposition. The sparsity of each IMF must be calculated by its bandwidth, in frequency domain before performing the decomposition [19]. The constrained Variational problem is described by the following equation.

$$\min_{\{v_k\}, \{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[ \partial_t \left( \frac{j}{\pi t} v_k(t) \right) e^{-j\omega_k t} \right] \right\|^2 \right\} \quad (7)$$

Where,  $v_k = \{v_1 \dots v_k\}$ ,  $\omega_k = \{\omega_1 \dots \omega_k\}$  and  $\sum_k v_k = f$ . The more details of VMD are provided in the [19]. The application of VMD is also evaluated.

#### B. Feature Extraction

Feature extraction is used to reduce the dimensions of IMFs into manageable feature sets, and a large dataset requires more computational time. Significant features are extracted to describe original IMFs coefficients [19]. Firstly, Band and Pompe proposed and successfully analyzed the Permutation entropy (PE) [21]. In the proposed method, three different multiscale PE (MPE) features namely Tsallis PE (TPE), Renyi PE (RPE), and Shannon PE (SPE) have been used for feature extraction [21].

Let  $K$  be the signal length, is divided into various vectors having  $N$  consecutive data extracted by the moving window method. The resultant vectors are named as  $y(i)$ :  $1 \leq i \leq N$ . Then vectors are reconstructed as follows:

$$Y_t = [y_t, y_{t+\tau}, \dots, y_{t+(m-1)\tau}] \quad (8)$$

Where, embedding dimension  $m$  with lag  $\tau$  by rearranging the  $Y_t$  in increasing order. There will be  $m!$  permutation patterns  $\pi_i$ . To every pattern  $\pi_i$ , there will be  $f(\pi_i)$  frequency of occurrence in a given signal. The permutation probability distribution  $p(\pi_i)$  is,

$$p(\pi_i) = \frac{f(\pi_i)}{N - (m-1)\tau} \quad (9)$$

Where, length is denoted by  $N$ . The normalized SPE is given by the following equation.

$$SPE = \frac{-\sum_{i=1}^{m!} p(\pi_i) \log(p(\pi_i))}{\log(m!)} \quad (10)$$

The RPE is expressed in equation (11),

$$RPE = \frac{\log \sum_{i=1}^{m!} (p(\pi_i))^a}{(1-a) \log(m!)} \quad (11)$$

Similarly, Zunino has proposed normalized TPE based on Tsallis entropy,

$$TPE = \frac{1}{(1-m)^{(1-q)}} \sum_{i=1}^m (p(\pi_i) - p(\pi_i)^q) \quad (12)$$

The forward elimination process selects the optimal parameters of SPE, RPE, and TPE. It is noted that  $\tau = 1$  and  $m = 6$  are best values for SPE, whereas  $m = 5$ ,  $a = 2$  and  $\tau = 1$  are chosen for RPE. Similarly,  $m=5$ ,  $q=0.1$  and  $\tau = 1$  have performs better in TPE measurement.

### C. Statistical Analysis

In the present work, the statistical analysis has carried out using the KW test. It is a non-parametric method that checks whether samples originate from the same distribution or not. It returns the probabilistic (p) value. If the  $p < 0.05$ , that indicates feature sets belong to two different distributions [17].

### D. Classification Algorithms

The selection of the most suitable ML algorithms is essential for the robustness of the proposed technique. In the present work, the four most popular ML models including DT [22], SVM [23], EBT, KNN [24], NN with different kernels have used for the detection of AD and MCI. DTs are tree-like structures that include internal nodes with multiple leaves with branches. Internal nodes show the features of a dataset, branches indicate decision rules, and outputs are classified [25]. The SVM is a powerful supervised ML-model developed for binary and multiclass classification [26]. It separates data into two different groups by obtaining a proper hyperplane in higher dimensional space that will maximize the margin of the two groups. It could reduce the over-fitting problem. KNN is a non-parametric supervised classifier. It does not need any prior assumption about the training datasets. For new testing data points, KNN classifies it by calculating the distance between the K number of training data points and test data points. KNN has only two parameters to tune, K-value and distance. The details of EBT are provided by Wang *et al.* [24]. These four classifiers are powerful models of ML. They have been successfully used for various neurological applications. The present work uses 10-fold CV for all ML models. The classification tasks have been performed 30 times to ensure the robustness of ML models. Sensitivity (SENS), specificity (SPEC), Accuracy (ACC), F1-score, kappa ( $\kappa$ ), and precision (PREC) parameters were evaluated to explore the proposed method with machine learning models [27].

## III. RESULTS AND DISCUSSION

The present study used the EEG datasets of three groups, namely AD, MCI, and NC subjects, to validate the proposed work. Initially, all the EEG signals have decomposed using DWT with 'db4' kernel at the fourth level. This yields five different subbands. These subbands have been reconstructed to get significant EEG bands namely, delta, theta, alpha, beta, and gamma. Further, these reconstructed subbands have decomposed using VMD into six IMFs. The first three IMFs out of six IMFs for AD and NC class are shown in Fig. 2 and Fig. 3. The three MPEs, namely, RPE, SPE, and TPE have been extracted from each IMF. The significance of the extracted features has been tested using the KW test. The box plot of SPE, RPE, and TPE features for MCI vs. AD, NC vs. AD, NC vs. MCI, and MCI vs. NC vs. AD classes are shown in Fig. 4. It has been observed that all three features are significant and have the high discriminant ability. In the

literature, various classifiers have utilized to classify EEG signals. The most promising classifiers that have obtained maximum accuracy are SVM, KNN, DT, and EBT. Hence, in the present study, SVM, KNN, DT, and EBT with different kernel functions have been used to evaluate the proposed method. The optimal hyper-parameters have been selected for all ML algorithms and are presented in Table I. To avoid the overfitting of the data, the training and testing data were splitted using 10-fold CV method.

TABLE I  
OPTIMAL HYPER-PARAMETERS FOR DIFFERENT CLASSIFIERS OBTAINED USING VARIOUS ITERATIONS

Model	Parameter settings
DT	Number of splits =30
SVM	Kernel scale: automatic
KNN	Distance weight = square inverse, Number of neighbors = 9, Distance metric=Euclidean
EBT	Max split = 50, Number of learners = 90, LR = 0.2, SD=11
NN	First layer size: 10, Activation: ReLU, Iteration limit: 900

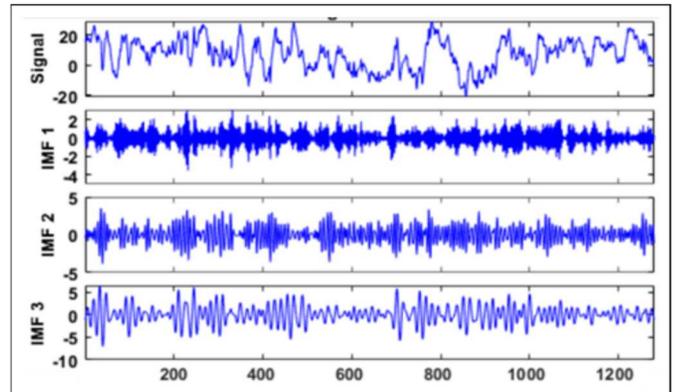


Fig. 2 IMFs plot for AD EEG signal obtained from VMD

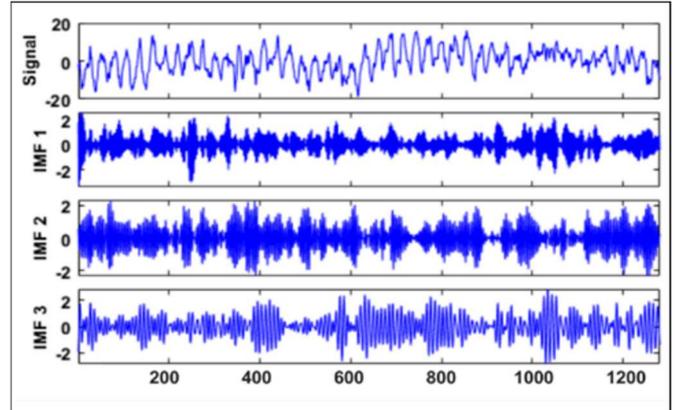


Fig. 3 IMFs plot for NC EEG signal obtained from VMD

### A. DWT-VMD Feature Based Classification

The three binary and one multiclass classification performances using proposed methods with different ML algorithms are presented in Table II. The different scenarios have been discussed below.

1) *AD vs. NC*: Table II shows that the EBT and SVM reported very good classification accuracy, of 97.7% and 97.4%, respectively. The F1-scores are 97.84% and 96.52% for EBT and SVM, respectively. On the other hand, DT

provides lower accuracy of 92.30%. Hence, EBT shows the highest AD discriminating power compared to other ML models.

2) *AD vs. MCI*: The EBT and SVM reported good performance compared to other classifiers, with 94.7% and 94.3% accuracy. However, the F1-score is higher in NN than EBT and SVM classifiers. The DT performed poorly with 86.3% accuracy and 83.6% F1-score. The other parameters of different classifiers are presented in Table II.

3) *MCI vs. NC*: In this scenario, a maximum of 94.9% accuracy has been provided by the EBT with a 92.36% F1-score. However, SVM reported poor accuracy of 90.5% compared to DT, NN, and KNN. The NN achieved comparable accuracy of 94% with EBT. The performance of EBT has reduced in MCI vs. NC classification compared to AD vs. NC due to the PSD closeness of EEG signals in MCI and NC.

4) *AD vs. MCI vs. NC*: Table II reports that the EBT has the highest discriminating ability of three different classes with an accuracy of 94.3%. The SVM performs slightly less than EBT, with 92.6% accuracy. The performance of all the ML models is lower in AD vs. MCI vs. NC compared to AD vs. NC. Out of five ML models, the EBT model provided higher accuracy in 2-way and 3-way classification than other ML models.

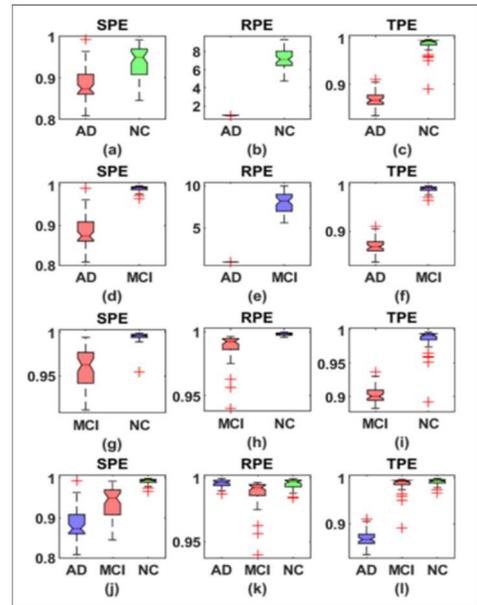


Fig. 4 Box plots of the three different features, including SPE, RPE, and TPE, for three binary classes: AD vs. NC, AD vs. MCI, MCI vs. NC, and AD vs. MCI vs. NC. SPE, RPE, and TPE for AD vs. NC are shown in (a), (b), and (c). SPE, RPE, and TPE for AD vs. MCI have shown in (d), (e), and (f). SPE, RPE, and TPE for MCI vs. NC are shown in (g), (h), and (i). SPE, RPE, and TPE for AD vs. MCI are shown in (j), (k), and (l).

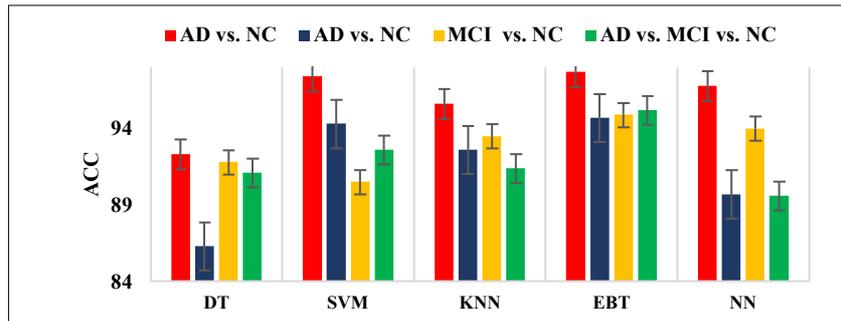


Fig. 5 Performance of different classifiers for three binary classes and one 3-way class.

TABLE II  
PERFORMANCE PARAMETERS OBTAINED FROM VARIOUS CLASSIFIERS FOR TWO AND THREE-CLASS CLASSIFICATION

Class	Model	Kernel	ACC	SENS	SPES	F1-score	PREC	MCC	K-value
AD vs. NC	DT	Course	92.30	85.93	99.75	91.51	95.45	88.34	87.74
	SVM	Cubic	97.40	96.20	98.25	96.52	97.88	94.64	94.63
	KNN	Medium	95.60	89.35	99.75	93.44	96.49	91.04	90.70
	EBT	Bagged tree	<b>97.70</b>	95.06	99.50	97.84	98.15	95.30	95.24
AD vs. MCI	NN	Wide	96.80	96.20	97.25	97.49	97.37	93.39	93.39
	DT	Course	86.30	71.48	96.00	83.66	89.41	71.53	70.18
	SVM	Cubic	94.30	85.93	99.75	91.51	95.45	88.34	87.74
	KNN	Medium	92.60	84.79	97.75	90.72	94.10	84.66	84.24
MCI vs. NC	EBT	Bagged tree	<b>94.70</b>	86.69	100	91.95	95.81	89.28	88.71
	NN	Wide	89.70	89.35	90.00	92.78	91.37	78.79	78.74
	DT	Course	91.80	85.93	99.00	91.45	95.08	87.29	86.80
	SVM	Cubic	90.50	89.35	91.25	92.88	92.06	80.26	80.24
AD vs. MCI vs. NC	KNN	Medium	93.50	90.11	95.75	92.64	94.68	86.41	86.37
	EBT	Bagged tree	<b>94.90</b>	87.45	99.75	93.36	95.91	89.53	89.06
	NN	Wide	94.00	90.87	96.00	94.12	95.05	87.36	87.33
	DT	Course	91.10	89.73	92.00	93.16	92.58	81.48	81.47
AD vs. MCI vs. NC	SVM	Cubic	92.60	84.79	97.75	90.72	94.10	84.66	84.24
	KNN	Medium	91.40	91.25	91.50	94.09	92.78	82.21	82.17
	EBT	Bagged tree	<b>95.20</b>	85.93	99.75	91.51	95.45	88.34	87.74
	NN	Wide	89.60	81.37	95.00	88.58	91.68	78.18	77.84

TABLE III  
ACCURACIES OBTAINED FROM VARIOUS CLASSIFIERS FOR TWO CLASS AND THREE CLASS CLASSIFICATION

Model	Kernel	AD vs. NC	AD vs. MCI	MCI vs. NC	AD vs. MCI vs. NC
DT	Course	92.3	86.30	91.80	91.10
SVM	Cubic	97.4	94.30	90.50	92.60
KNN	Medium	95.6	92.60	93.50	91.40
EBT	Bagged tree	97.7	94.70	94.90	95.20
NN	Wide	96.8	89.70	94.00	89.60

The maximum accuracies obtained for binary and multiclass datasets are depicted in Fig. 5 and Table III. It has been noticed that the performance of DT is poor compared to other classifiers. The performance of all the models has been degraded for 3-way classification due to the closeness of the MPE features in MCI and NC.

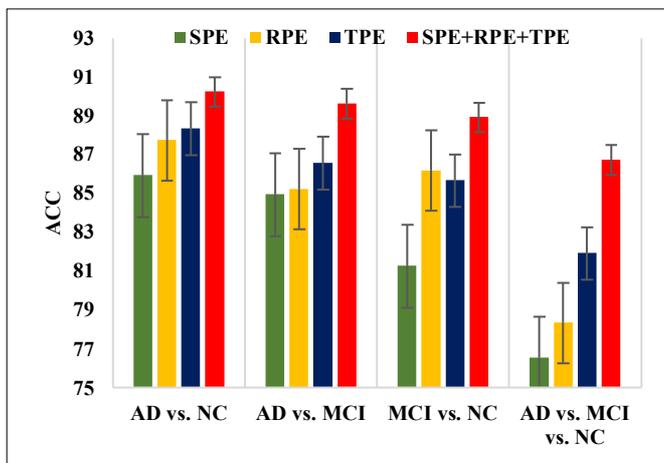


Fig. 6 Performance of the time domain features SPE, TPE, RPE and combination of all MPEs (SPE+RPE+TPE) for three binary classes and one three class.

### B. MPEs Feature Based Classification

The SPE, TPE, and RPE features are calculated from the original EEG datasets of AD, MCI, and NC subjects without decomposition. The performance of these features has been evaluated and presented in Fig. 6. A combination of all the features, including SPE, RPE, and TPE provides 90.2% accuracy for AD vs. NC. However, its performance has been reduced to AD vs. MCI vs. NC. If we compare the classification accuracy for all MPE features with (DWT-VMD) and without decomposition, the DWT-VMD based features achieved higher accuracy (7%) than the MPE based features.

### C. Comparison of DWT, EMD, and DWT-VMD Based Classification

The performance of the proposed DWT-VMD based dual decomposition method has been compared with the EMD, DWT, and DWT-EMD dual decomposition methods. The classification rate for these methods is presented in Table IV and Fig. 7. The EMD based feature set provided 92.9% accuracy, whereas DWT-based features obtained 93.6% classification accuracy. The DWT-EMD surpassed the EMD and DWT, which achieved 94.5% accuracy for AD vs. NC.

This performance has been reduced for the three-way classification (AD vs. MCI vs. NC).

TABLE IV  
ACCURACIES OBTAINED FROM VARIOUS CLASSIFIERS FOR TWO CLASS AND THREE CLASS CLASSIFICATION USING EBT ALGORITHM

Model	AD vs. NC	AD vs. MCI	MCI vs. NC	AD vs. MCI vs. NC
SPE+RPE+TPE	90.2	89.6	88.9	82.0
EMD	92.9	90.3	89.4	87.6
DWT	93.6	91.2	90.1	88.4
DWT-EMD	94.5	93.2	92.8	90.5
DWT-VMD	97.6	94.4	94.6	95.2

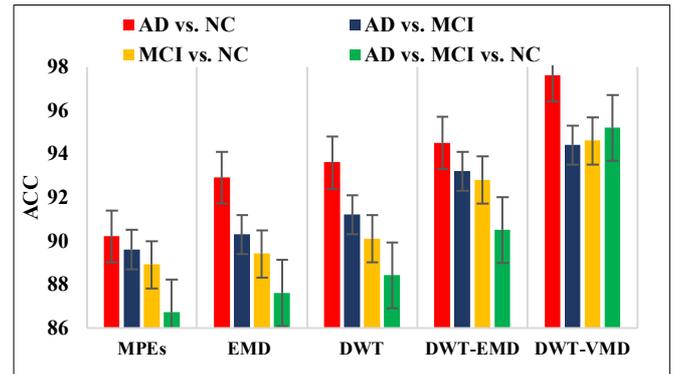


Fig. 7 Performance of the MPEs, EMD, DWT, DWT-EMD, and DWT-VMD for 2-way and 3-way classification.

Hence, the proposed method enhances the accuracy by 3% compared to DWT-EMD based method. The accuracy of the proposed method for AD vs. MCI vs. NC has enhanced only by 4% compared to other existing algorithms.

### D. Benchmarking with Previous Studies

The existing work has exploited different features for the detection of AD and MCI. They also concluded that the EEG features for AD detection might not perform well for detecting MCI. Various trends can be accepted from the literature to detect AD: (1) decrease in EEG complexity of AD patients compared to NC and MCI. (2) The shift of power spectrum from beta and gamma frequency bands to delta and theta frequency bands.

The present study has compared the recent state-of-the-art techniques using various decomposition methods. The comparison of this work with recent studies has been presented in Table V. Recently, Safi *et al.* [15] proposed DWT based Hjorth parameters to detect the AD with LDA, SVM, and KNN to achieve 95.79% classification accuracy. However, they achieved very low accuracy using the leave-one-out CV method. Puri *et al.* [28] has used wavelet packet-based features with SVM to classify the AD EEG signals and achieved 95.2% accuracy. Similarly, Oltu *et al.* [29] have proposed a DWT based PSD features with EEG bands coherence values and obtained 96.5% accuracy using EBT. However, they obtained the results without selecting the optimal hyper-parameters of classifiers. Toural *et al.* [16] used DWT based entropy (WE) and PSD feature with SVM to achieve 94.4% accuracy. However, their performance is poor for the three-class classification. Ieracitano *et al.* [30] proposed Bi-spectrum (BiS) based higher order statistical features with multi-layer perceptron (MLP) to detect the AD

from NC. They reported 96.7% accuracy [30], [31]. In other studies [32]–[36] authors used DWT and FFT based PSD and magnitude features with KNN and SVM model to get almost 87% accuracy. Moreover, iterative filtering (IF) based statistical features have been used to detect AD from NC. However, the FFT-based features fail to capture the small changes in EEG. Moreover, the selection of the number of levels and mother wavelet is a major issue with DWT. On the other hand, the proposed method overcomes these issues with dual decomposition (DWT-VMD). The MPEs have the high capability to capture the complexity present in non-stationary signals. The present study achieved maximum ACC of 97.7% for NC vs. AD and 95.2% for AD vs. NC vs. MCI. Also, this study improved the accuracy by 1-10% from the existing methods. The main features of the proposed method are listed below,

- The proposed method uses the dual decomposition technique (DWT-VMD) to overcome the limitation of DWT and EMD.
- The MPEs capture the hidden information in EEG signals due to their better anti-noise ability.
- The most powerful and popular ML models are trained and tested to achieve maximum accuracy.
- The proposed two- and three-class classification techniques achieved 97.7% and 95.2% accuracy, respectively.
- The MPEs features were evaluated using with and without decomposition, and their performance has been compared on AD, MCI, and NC EEG datasets.
- The proposed model has been compared with state-of-the-art methods.

#### E. Limitations and Future Research Work

The dataset used in the present work is quite small from the 16 channels. The proposed method needs to validate the big EEG datasets with maximum EEG channels used for EEG

acquisition. We used only three MPEs features to train and test the ML models for AD detection. In the future, authors intend to extend the proposed system for the detection of other diseases like epilepsy [37], Schizophrenia, mental depressive disorder [38], Parkinson’s disease [39], and alcoholism [40]. This work can also be extended using deep learning models on the same EEG databases.

#### IV. CONCLUSION

EEG biomarkers play an important role in diagnosing AD at its early stage. In the present study, we explore the potential of a dual decomposition technique that combines DWT and VMD with multiscale permutation entropies (SPE, RPE, and TPE) to detect AD. It helps to understand the development and disease progress in an earlier stage. It has been noticed that the MPEs values for AD are lower than MCI and NC. Moreover, significant differences have been observed between MCI and NC. Hence, the EEG signals of AD and MCI patients are more regular compared to NC subjects. However, there is no clear boundary between the normal aging effect and MCI. The proposed method focused on achieving higher classification accuracy, i.e., 97.70% for AD vs. NC and 95.20% for AD vs. MCI vs. NC with the EBT algorithm.

On the other hand, the DWT-VMD decomposition provides significant features from AD and MCI EEG signals. Moreover, the proposed technique surpassed the state-of-the-art techniques and improved AD detection accuracy by 2%. Further, this proposed method can be validated for other neurological disorders like epilepsy, sleep disorders, hypertension, and Parkinson’s disease. In the future, deep learning modules can be applied to the present EEG dataset. The use of new filter banks to find the different EEG-based features can be checked to enhance the detection and classification of different diseases.

TABLE V  
COMPARISON OF PROPOSED METHOD EXISTING METHODS FOR TWO AND THREE CLASS CLASSIFICATION

References	Year	Method	Features used	Classifier	Class	CV	ACC (%)	SENS (%)	SPEC (%)
Safi <i>et al.</i> [15]	2021	DWT	Hjorth parameters	SVM	NC vs. AD vs. MCI	10	95.79	91.93	97.85
Oltu <i>et al.</i> [29]	2021	DWT	Amplitude and PSD	EBT	NC vs. AD vs. MCI	5	96.50	96.21	97.96
Toural <i>et al.</i> [16]	2021	DWT	WE, theta power	NN	NC vs. AD vs. MCI	10	94.44	98.92	97.21
Ieracitano <i>et al.</i> [30]	2020	CWT, BiS	Higher order feature	MLP	NC vs. AD	10	96.70	94.56	96.24
Sharma <i>et al.</i> [31]	2020	IF	PSD, FD, and TE	KNN	NC vs. AD vs. MCI	10	92.00	86.77	94.89
Durongbhan <i>et al.</i> [32]	2019	CWT, FFT	Average magnitude	KNN	NC vs. AD vs. MCI	10	83.32	72.57	87.52
This work		DWT-VMD	SPE, TPE, and RPE	EBT	NC vs. AD	10	<b>97.70</b>	96.06	99.50
				EBT	AD vs. MCI	10	94.70	86.69	98.50
				EBT	NC vs. AD vs. MCI	10	95.20	85.93	99.75

#### REFERENCES

- [1] A. Atri, “The Alzheimer’s Disease Clinical Spectrum: Diagnosis and Management,” *Medical Clinics of North America*, vol. 103, no. 2. W.B. Saunders, pp. 263–293, Mar. 01, 2019. doi: 10.1016/j.mcna.2018.10.009.
- [2] “2020 Alzheimer’s disease facts and figures,” *Alzheimer’s and Dementia*, vol. 16, no. 3, pp. 391–460, Mar. 2020, doi: 10.1002/alz.12068.
- [3] S. J. Ruiz-Gómez *et al.*, “Automated multiclass classification of spontaneous EEG activity in Alzheimer’s disease and mild cognitive impairment,” *Entropy*, vol. 20, no. 1, Jan. 2018, doi: 10.3390/e20010035.
- [4] K. Oppedal, K. Engan, T. Eftestøl, M. Beyer, and D. Aarsland, “Classifying Alzheimer’s disease, Lewy body dementia, and normal controls using 3D texture analysis in magnetic resonance images,” *Biomed Signal Process Control*, vol. 33, pp. 19–29, Mar. 2017, doi: 10.1016/j.bspc.2016.10.007.
- [5] X. Bi and H. Wang, “Early Alzheimer’s disease diagnosis based on EEG spectral images using deep learning,” *Neural Networks*, vol. 114, pp. 119–135, Jun. 2019, doi: 10.1016/j.neunet.2019.02.005.
- [6] E. E. Tülay, B. Güntekin, G. Yener, A. Bayram, C. Başar-Eroğlu, and T. Demiralp, “Evoked and induced EEG oscillations to visual targets reveal a differential pattern of change along the spectrum of cognitive decline in Alzheimer’s Disease,” *International Journal of Psychophysiology*, vol. 155, pp. 41–48, Sep. 2020, doi: 10.1016/j.ijpsycho.2020.06.001.

- [7] G. Fison *et al.*, "Combining EEG signal processing with supervised methods for Alzheimer's patients classification," *BMC Med Inform Decis Mak*, vol. 18, no. 1, May 2018, doi: 10.1186/s12911-018-0613-y.
- [8] A. Khosla, P. Khandnor, and T. Chand, "A comparative analysis of signal processing and classification methods for different applications based on EEG signals," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 2. Elsevier Sp. z o.o., pp. 649–690, Apr. 01, 2020. doi: 10.1016/j.bbe.2020.02.002.
- [9] S. Siuly *et al.*, "A New Framework for Automatic Detection of Patients with Mild Cognitive Impairment Using Resting-State EEG Signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 9, pp. 1966–1976, Sep. 2020, doi: 10.1109/TNSRE.2020.3013429.
- [10] S. Simons, P. Espino, and D. Abásolo, "Fuzzy Entropy analysis of the electroencephalogram in patients with Alzheimer's disease: Is the method superior to Sample Entropy?," *Entropy*, vol. 20, no. 1, Jan. 2018, doi: 10.3390/e20010021.
- [11] Y. Ding *et al.*, "Fully automated discrimination of Alzheimer's disease using resting-state electroencephalography signals," *Quant Imaging Med Surg*, vol. 12, no. 2, pp. 1063–1078, Feb. 2022, doi: 10.21037/qims-21-430.
- [12] P. M. Rodrigues, B. C. Bispo, C. Garrett, D. Alves, J. P. Teixeira, and D. Freitas, "Lacsogram: A New EEG Tool to Diagnose Alzheimer's Disease," *IEEE J Biomed Health Inform*, vol. 25, no. 9, pp. 3384–3395, Sep. 2021, doi: 10.1109/JBHI.2021.3069789.
- [13] M. Cejnek, O. Vysata, M. Valis, and I. Bukovsky, "Novelty detection-based approach for Alzheimer's disease and mild cognitive impairment diagnosis from EEG," *Med Biol Eng Comput*, vol. 59, no. 11–12, pp. 2287–2296, Nov. 2021, doi: 10.1007/s11517-021-02427-6.
- [14] J. P. Amezquita-Sanchez, N. Mammone, F. C. Morabito, S. Marino, and H. Adeli, "A novel methodology for automated differential diagnosis of mild cognitive impairment and the Alzheimer's disease using EEG signals," *J Neurosci Methods*, vol. 322, pp. 88–95, Jul. 2019, doi: 10.1016/j.jneumeth.2019.04.013.
- [15] 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, Mar. 2021, doi: 10.1016/j.bspc.2020.102338.
- [16] J. E. Santos Tournal, A. Montoya Pedrón, and E. J. Marañón Reyes, "Classification among healthy, mild cognitive impairment and Alzheimer's disease subjects based on wavelet entropy and relative beta and theta power," *Pattern Analysis and Applications*, vol. 24, no. 2, pp. 413–422, May 2021, doi: 10.1007/s10044-020-00910-8.
- [17] V. Bajaj, S. Taran, S. K. Khare, and A. Sengur, "Feature extraction method for classification of alertness and drowsiness states EEG signals," *Applied Acoustics*, vol. 163, Jun. 2020, doi: 10.1016/j.apacoust.2020.107224.
- [18] D. Puri, S. Nalbalwar, A. Nandgaonkar, P. Kachare, J. Rajput, and A. Wagh, "Alzheimer's Disease Detection using Empirical Mode Decomposition and Hjorth parameters of EEG signal," in *2022 International Conference on Decision Aid Sciences and Applications, DASA 2022*, 2022, pp. 23–28. doi: 10.1109/DASA54658.2022.9765111.
- [19] Y. Miao, M. Zhao, and J. Lin, "Identification of mechanical compound-fault based on the improved parameter-adaptive variational mode decomposition," *ISA Trans*, vol. 84, pp. 82–95, Jan. 2019, doi: 10.1016/j.isatra.2018.10.008.
- [20] N. Ji, L. Ma, H. Dong, and X. Zhang, "EEG signals feature extraction based on DWT and EMD combined with approximate entropy," *Brain Sci*, vol. 9, no. 8, Aug. 2019, doi: 10.3390/brainsci9080201.
- [21] Z. Huo, Y. Zhang, L. Shu, and M. Gallimore, "A New Bearing Fault Diagnosis Method Based on Fine-to-Coarse Multiscale Permutation Entropy, Laplacian Score and SVM," *IEEE Access*, vol. 7, pp. 17050–17066, 2019, doi: 10.1109/ACCESS.2019.2893497.
- [22] S. Uddin, A. Khan, M. E. Hossain, and M. A. Moni, "Comparing different supervised machine learning algorithms for disease prediction," *BMC Med Inform Decis Mak*, vol. 19, no. 1, Dec. 2019, doi: 10.1186/s12911-019-1004-8.
- [23] J. Panyavaraporn and P. Horkaew, "Classification of Alzheimer's Disease in PET Scans using MFCC and SVM," vol. 8, no. 5, 2018.
- [24] Z. Wang, J. Na, and B. Zheng, "An Improved kNN Classifier for Epilepsy Diagnosis," *IEEE Access*, vol. 8, pp. 100022–100030, 2020, doi: 10.1109/ACCESS.2020.2996946.
- [25] J. Rabcan, V. Levashenko, E. Zaitseva, and M. Kvassay, "Review of methods for EEG signal classification and development of new fuzzy classification-based approach," *IEEE Access*, vol. 8. Institute of Electrical and Electronics Engineers Inc., pp. 189720–189734, 2020. doi: 10.1109/ACCESS.2020.3031447.
- [26] H. Sujaini, "Image Classification of Tourist Attractions with K-Nearest Neighbor, Logistic Regression, Random Forest, and Support Vector Machine," vol. 10, no. 6, 2020.
- [27] D. Wang *et al.*, "Classification, experimental assessment, modeling methods and evaluation metrics of Trombe walls," *Renewable and Sustainable Energy Reviews*, vol. 124. Elsevier Ltd, May 01, 2020. doi: 10.1016/j.rser.2020.109772.
- [28] D. Puri, S. Nalbalwar, A. Nandgaonkar, and A. Wagh, "Alzheimer's disease detection with Optimal EEG channel selection using Wavelet Transform," in *2022 International Conference on Decision Aid Sciences and Applications, DASA 2022*, 2022, pp. 443–448. doi: 10.1109/DASA54658.2022.9765166.
- [29] B. Oltu, M. F. Akşahin, and S. Kibaroglu, "A novel electroencephalography based approach for Alzheimer's disease and mild cognitive impairment detection," *Biomed Signal Process Control*, vol. 63, Jan. 2021, doi: 10.1016/j.bspc.2020.102223.
- [30] 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 Networks*, vol. 123, pp. 176–190, Mar. 2020, doi: 10.1016/j.neunet.2019.12.006.
- [31] D. Puri, S. Nalbalwar, A. Nandgaonkar, and A. Wagh, "Alzheimer's disease detection from optimal electroencephalogram channels and tunable Q-wavelet transform," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 25, no. 3, pp. 1420–1428, Mar. 2022, doi: 10.11591/ijeecs.v25.i3.pp1420-1428.
- [32] N. Sharma, M. H. Kolekar, and K. Jha, "Iterative Filtering Decomposition Based Early Dementia Diagnosis Using EEG with Cognitive Tests," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 9, pp. 1890–1898, Sep. 2020, doi: 10.1109/TNSRE.2020.3007860.
- [33] P. Durongbhan *et al.*, "A dementia classification framework using frequency and time-frequency features based on EEG signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 5, pp. 826–835, May 2019, doi: 10.1109/TNSRE.2019.2909100.
- [34] K. D. Tzamourta *et al.*, "EEG window length evaluation for the detection of Alzheimer's disease over different brain regions," *Brain Sci*, vol. 9, no. 4, Apr. 2019, doi: 10.3390/brainsci9040081.
- [35] A. M. Pineda, F. M. Ramos, L. E. Betting, and A. S. L. O. Campanharo, "Quantile graphs for EEG-based diagnosis of Alzheimer's disease," *PLoS One*, vol. 15, no. 6, Jun. 2020, doi: 10.1371/journal.pone.0231169.
- [36] L. Tylová, J. Kukal, V. Hubata-Vacek, and O. Vyšata, "Unbiased estimation of permutation entropy in EEG analysis for Alzheimer's disease classification," *Biomed Signal Process Control*, vol. 39, pp. 424–430, Jan. 2018, doi: 10.1016/j.bspc.2017.08.012.
- [37] D. Puri, R. Chudiwal, and P. Kachare, *Detection of Epilepsy using Wavelet Packet Sub-bands from EEG Signals*, vol. 303 SIST. 2022. doi: 10.1007/978-981-19-2719-5\_28.
- [38] J. S. Rajput, M. Sharma, R. S. Tan, and U. R. Acharya, "Automated detection of severity of hypertension ECG signals using an optimal bi-orthogonal wavelet filter bank," *Comput Biol Med*, vol. 123, Aug. 2020, doi: 10.1016/j.combiomed.2020.103924.
- [39] M. Zhu, A. HajiHosseini, T. R. Baumeister, S. Garg, S. Appel-Cresswell, and M. J. McKeown, "Altered EEG alpha and theta oscillations characterize apathy in Parkinson's disease during incentivized movement," *Neuroimage Clin*, vol. 23, Jan. 2019, doi: 10.1016/j.nicl.2019.101922.
- [40] D. Puri, R. Ingle, P. Kachare, M. Patil, and R. Awale, "Wavelet Packet Sub-band Based Classification of Alcoholic and Controlled State EEG Signals," *Proceedings of the International Conference on Communication and Signal Processing*, Atlantis Press, pp. 562–567, 2017. doi: 10.2991/iccasp-16.2017.82.